Dynamics and Heterogeneity of mRNA Translation and RNA Virus Infections

Sanne Boersma

Dynamics and Heterogeneity of mRNA Translation and RNA Virus Infections

Dynamiek en variaties tijdens mRNA translatie en RNA virus infecties
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

woensdag 18 mei 2022 des middags te 2.15 uur

door

Sanne Boersma

geboren op 3 augustus 1993
te Amsterdam

Promotor:
Prof. dr. W.L. de Laat

Copromotor:
Dr. M.E. Tanenbaum

TABLE of CONTENT

Chapter 1

Introduction

SUMMARY

This thesis shows that the development of single-molecule imaging assays can lead to insights into dynamic processes related to cellular and viral gene expression. In this chapter, I provide background information on the processes that we studied: 1) cellular gene expression, in particular mRNA translation initiation, 2) viral protein synthesis and replication, specifically the early events during the lifecycle of +RNA viruses *Picornavirales* and -RNA viruses *Mononegavirales*, and 3) innate antiviral response, focusing on key pathways responsible for detection and inhibition of RNA viruses. Throughout this introduction, I highlight some of the open questions that can be addressed by live-cell single-molecule experiments, such as the assays described in this thesis.

STUDY DYNAMIC PROCESSES USING LIVE-CELL SINGLE-MOLECULE IMAGING

Biologists often depict biological processes as simplified cartoons. They illustrate a transition from start to end with a single arrow, implying that the transition is robust and unidirectional. However, most biological processes are far from static. Instead, processes are dynamic and comprise transitions from one state into another that are highly influenceable by their environment (Fig. 1). An example is the transition of a pluripotent stem cell into a differentiated cell with its highly specified function. Another example is the binding of a signaling molecule to its receptor causing activation of a signaling cascade and resulting in expression of genes. Although the simplified cartoon of these processes implies that the differentiation or signaling cascades are very efficient, a spectrum of factors may contribute to the success-rate of the process. For instance, most processes are not a single step, but rather occur through many intermediate steps, each of which may be subjected to regulation. For example, the concentration of regulatory factors can affect the success-rate of each step. Additionally, a process can be either unidirectional or may be reversable, reducing the overall success-rate of a process. Furthermore, factors controlling the efficiency of a of a dynamic process depend on the spatiotemporal context of the process: circumstances are not uniform at all locations at any time. For example, translation on a single mRNA requires the presence of regulatory proteins on that mRNA at a specific time-point. Only measuring the general abundance of such a regulatory protein in a large collection of cells may be insufficient to understand the translation of an individual mRNA in a particular cell.
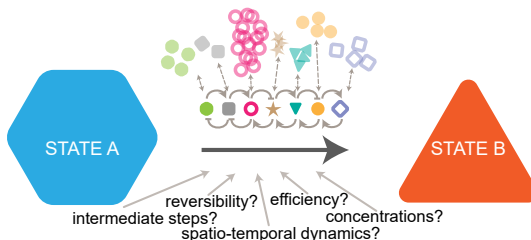

Figure 1. Many factors can contribute to dynamics of a biological process.

Schematic illustrating classical 'textbook' depiction of a biological process, including examples of factors that may complicate the simplified depiction.

A multitude of experimental assays can be employed to study a biological process and each of these assays have their advantages and limitations. An ideal assay provides detailed information about the stages of the dynamic process by analyzing the process at different time-points. Furthermore, one aims for a set-up that involves minimal perturbations, and facilitates direct examination in a relevant context. Such a context entails a native environment (e.g., a live cell) and minimal destruction of the sample (e.g., through lysis or fixation). Live-cell single-molecule imaging meets most of the above-mentioned advantages and is thus ideally suited to study dynamic processes. Throughout this thesis, I therefore employ live-cell single-molecule imaging assays to investigate dynamic processes. Specifically, I focus on three different processes: cellular protein production at the level of mRNA translation initiation, viral replication and protein synthesis, and the competition between a host cell and a virus during early infection.

I.   CELLULAR GENE EXPRESSION

CENTRAL DOGMA

Proteins are essential for all processes in life. They are required for building cells, maintaining cellular structures and communication within and between cells. Protein synthesis depends on proper transfer of genetic information from DNA in the nucleus of mammalian cells into functional polypeptides throughout a cell, collectively referred to as The central dogma of molecular biology (Fig. 2; Crick, 1970). In eukaryotes, protein synthesis involves transcription of a gene into mRNA in the nucleus, which is subsequently translated by ribosomes in the cytoplasm. During translation, the nucleotide sequence of mRNA is converted into a polypeptide string of amino acids. This polypeptide then folds into a 3D structure and forms a functional protein. The process of gene expression to produce proteins is subjected to multiple types of regulation to control the type and quantity of proteins in cells at different time points. These regulations must not only be very robust to ensure synthesis of the correct proteins, but they must also be dynamic to enable inducible and rapid changes in protein levels. Cell-to-cell variations in protein levels cannot be explained by heterogeneity in mRNA levels alone, but an important step in shaping the proteome takes place at the level of mRNA translation (Azimifar et al., 2014; Frei et al., 2016; Mair et al., 2020; Schwanhäusser et al., 2011).
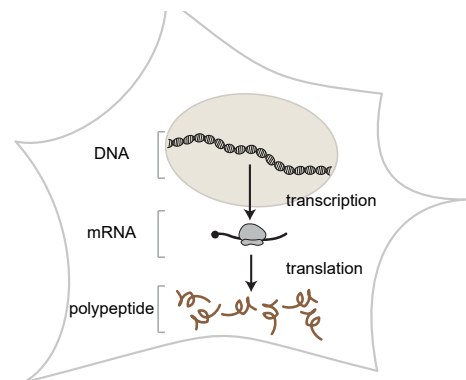

Figure 2. Central dogma of gene expression.

Schematic illustrating the flow from genetic information stored as DNA in the nucleus, transcribed into mRNAs that can be translated into polypeptides in the cytoplasm by ribosomes.

TRANSLATION INITIATION IS AN IMPORTANT STEP DURING GENE EXPRESSION.

Translation of mRNAs into polypeptides comprises of three main steps: initiation, elongation, and termination. Initiation involves recruiting the small ribosomal subunit (40S) to the 5' end of an mRNA molecule and scanning the 5' untranslated region (UTR) to identify the start site, i.e. initiation codon (Hinnebusch et al., 2016; Jackson et al., 2010). Recognition of the start site depends on base pairing between the start codon in the mRNA and a tRNA in the scanning ribosome, leading to the assembly of the large ribosomal subunit (60S) to form a full (80S) ribosome that then continues with elongation. During elongation, the ribosome decodes the mRNA sequence into a polypeptide sequence. tRNA molecules loaded with aminoacyls are first loaded into the aminoacyl -site (A-site) of the ribosomes. This loading is orchestrated by base pairing between the mRNA codon and the aminoacyl-tRNA anticodon. Next, the aminoacyl from the tRNA is transferred to the peptidyl-tRNA in the peptidyl-site (P-site) of the ribosome, resulting in coupling of the aminoacyl group from the incoming tRNA to the nascent chain. Finally, the ribosome translocates causing the deactylated tRNA to occupy the exit-site (E-site) from where it can dissociate from the ribosome and a new aminoacyl-tRNA can enter the A-site. The elongation cycle to decode the mRNA codon sequence into a polypeptide sequence continues until a stop codon is positioned in the A-site. Once a ribosome reaches a stop codon, a termination protein is loaded into the ribosome, causing the release of the polypeptide and dissociation of the small and large ribosomal subunit. The ribosomal subunits may be recruited to the same mRNA or another to translate again.

The number of proteins that are formed from an mRNA molecule depends on the number of ribosomes recruited, which is determined by the initiation efficiency. Additionally, the amino acid sequence of the polyprotein product is defined by the exact start codon from which translation is initiated. Namely, selection of an alternative start codon would result in a protein isoform with a different N-terminus. If the alternative start codon leads to translation of a different open reading frame (ORF) it leads to a completely different polypeptide sequence. Therefore, translation initiation is a key process during gene expression and most regulation of translation is exerted at the level of initiation.

All steps that lead to the assembly of an 80S ribosome with a methionine-loaded tRNA at the start of the coding sequence are collectively referred to as 'translation initiation' (Fig. 3). In this section, I explain the main events during canonical translation initiation as well as regulatory pathways controlling translation initiation and non-canonical deviations that may affect mRNA translation efficiency.

1. Pre-initiation complex formation

The first step of translation initiation is the assembly of a 43S pre-initiation complex (PIC), consisting of a 40S, several eukaryotic initiation factors (eIFs), and a tRNA molecule (Fig. 3). The large multi-subunit complex eIF3 acts as a scaffold to stabilize the PIC (Erzberger et al., 2014; Fraser et al., 2004). A specialized initiating tRNA (Met-tRNAi) complementary to the start codon and loaded with methionine forms a complex with a GTP-coupled eIF2, collectively referred to as the ternary complex

(TC). The TC is loaded into the 40S so that the Met-tRNAi is loaded in the prospective P-site of 40S (Sokabe and Fraser, 2014). The two other pockets of the 40S (E-site and P-site) are loaded with eIF1 and eIF1A respectively (Aylett et al., 2015). The final subunit making up the PIC is the GTPase activating protein (GAP) eIF5, that facilitates GTP hydrolysis of the Met-tRNAi-coupled GTP upon start codon recognition. This hydrolysis induces Met-tRNAi release from eIF2 (Das et al., 1997; Singh et al., 2004).

2. Ribosome recruitment to the cap

A handful of eIFs facilitates the recruitment of a PIC to an mRNA molecule, specifically to the 5' cap of mRNA. The scaffold eIF4G binds the cap-binding protein eIF4E, the helicase eIF4A, and can also bind to eIF4A-stimulating proteins, such as eIF4B or eIF4H (Marintchev et al., 2009). The entire cap-assembly of eIF4s is jointly called eIF4F (Hinnebusch et al., 2016; Merrick, 2015). In addition to scaffolding eIF4F, eIF4G can also bind eIF3, thereby facilitating the recruitment of a PIC to the 5' end of an mRNA and forming a 48S PIC-eIF4F assembly (Eliseev et al., 2018). eIF4A contributes to PIC recruitment by 'opening' the cap-proximal region of an mRNA, so that the PIC can 'land' on the mRNA (Merrick, 2015).

3. Scanning of the 5'UTR

During the scanning step of initiation, the 48S complex moves along the 5'UTR in a 5'-to-3' direction to identify the start codon (Alekhina and Vassilenko, 2012). In addition to opening the most 5' part of the mRNA in order for the 43S to land (step 2), helicase eIF4A is important for the scanning step as well, specifically scanning of 5'UTRs with structures (Pisareva et al., 2008; Svitkin et al., 2001). In the absence of eIF4A, translation initiation is strongly impaired, although some mRNAs with short unstructured 5'UTRs can still be translated (Pestova and Kolupaeva, 2002).

The PIC interaction with the 5' end of an mRNA is established through interactions between cap(mRNA)-eIF4E-eIF4G-eIF3-40S(ribosome), but the stability of these interactions during scanning is unclear. As a result, the mechanism and regulation of scanning are elusive. Based on
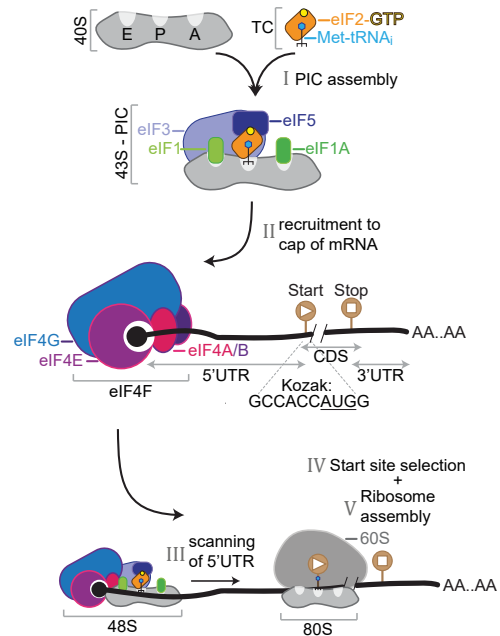
Figure 3. Translation initiation steps.

Schematic illustrating the key steps required for translation initiation and the main elements in an mRNA molecule. See main text for details. Components are not depicted at scale. Abbreviations: exit site (E); peptidyl site (P); aminoacyl site (A); initiator transfer RNA coupled to a methionine Met-tRNAᵢ; eukaryotic initiations factor (eIF); pre-initiation complex (PIC); untranslated region (UTR); coding sequence (CDS).

ribosomal profiling on initiating ribosomes, eIF3 remains associated with 40S during scanning and even during the first couple of elongation cycles (Archer et al., 2016; Lin et al., 2020; Wagner et al., 2020), suggesting that the 40S-eIF3 interaction is stable during scanning. Whether, when, and in which order the other interactions between the cap and the 40S are released during scanning is only partially understood (Bohlen et al., 2020; Gandin et al., 2021; Jackson et al., 2010; Shirokikh and Preiss, 2018). According to the cap-tethered model, the cap-to-40S interactions remain intact during scanning, whereas the cap-severed model suggests the connection is lost. In the cap-tethered model, only a single scanning event can take place at the time, as the cap remains occupied during scanning. Consequently, the cap-tethered model predicts that mRNAs with long 5'UTRs are translated inefficiently. Indeed, mRNAs with long 5'UTRs are typically translated at low efficiency (Chappell et al., 2006; Paek et al., 2015), although long 5'UTRs are also associated with other features that reduce translation efficiency (e.g. alternative start sites and uORFs). In contrast, in the cap-severed model, consecutive 43Ss can be recruited to an mRNA molecule by a single cap-eIF4E-eIF4G complex. One of the main mechanisms of controlling translation initiation is modulating the availability of eIF4E (see section 'eIF4E availability downstream of mTOR). The effect of limited eIF4E availability is different between the cap-tethered and the cap-severed model. Therefore, the composition of the scanning complex has implications for the exact mechanism of translation initiation regulation.

Importantly, data supporting either the cap-tethered or cap-severed model is based on structural analysis of eIFs, or based on ribosomal profiling of scanning ribosomes (Archer et al., 2016; Bohlen et al., 2020; Giess et al., 2020; Kolupaeva et al., 2003; Shirokikh and Preiss, 2018). Although these assays indicate whether a certain eIF has the ability to interact and whether the eIF has been detected during scanning or start codon recognition, they lack the sensitivity or resolution to conclusively proof or disprove a model. For example, the initiation ribosomal profiling indicates that some eIFs can bind to the scanning complex (Archer et al., 2016; Bohlen et al., 2020; Lin et al., 2020; Wagner et al., 2020). However, it is unclear if these are the same molecules that facilitate the recruitment of that 43S. It is possible that a stable cap-eIF4E-eIF4G complex recruits a 43S, which may be bound by another eIF4G molecule during scanning. Therefore, sophisticated assays are required to study the exact composition of the ribosome and determine the contribution of each eIF to the regulation of translation initiation during each step of initiation.

4+5. Start site selection & ribosome assembly

Start site selection depends on complementarity between the start codon (AUG) and the anticodon in the TC. Furthermore, the sequence surrounding the start codon, i.e. the Kozak sequence, is crucial for start site selection as well. The optimal Kozak sequence is GCCACC<u>AUG</u>G and the most important sites are the -3 nucleotide, the start codon (underlined), and the +4 nucleotide (Kozak, 1986; Pisarev et al., 2006). The efficiency of start site selection depends on the Kozak sequence. Start codons in a suboptimal Kozak sequence can also be used as a start site, albeit less efficiently (Noderer et al., 2014). Similarly, complementarity between the start codon and TC (AUG vs CAU) causes most efficient translation initiation, but alternative codons can also be used as a start site (Peabody, 1989). Particularly, the non-canonical start

codons CUG, GUG, and UUG are frequently used alternative start sites (Ingolia et al., 2011), and initiation from an alternative start codon also depends on its Kozak sequence (Hussain et al., 2014; Kearse and Wilusz, 2017; Martin-Marcos et al., 2011).

Non-start site AUGs and non-canonical start codons are abundant in the 5'UTRs and are widely used as alternative translation initiation sites (Ingolia et al., 2011). Selection of an alternative initiation site can have profound effects on the protein encoded by an mRNA. For example, mRNAs encoding the proto-oncogenic transcription factor c-Myc contain multiple frequently used alternative start sites (Ingolia et al., 2011). Translation from one of these non-canonical start sites results in a c-Myc isoform with altered transcription activity and is associated with Burkitt's lymphomas (Hann et al., 1988, 1994). Translation initiation from an alternative start codon also affects the protein encoded by the mRNA indirectly, by reducing the number of ribosomes that produce the encoded protein. In sum, the presence of alternative start sites impacts the proteome and start site selection contributes to the control of protein synthesis.

Codon-anticodon complementarity induces a cascade of rearrangements of the scanning complex. First, the 40S changes from an open and scanning-competent conformation into a closed conformation. During scanning, eIF1 and eIF1A enable the open conformation (Passmore et al., 2007). But the rearrangement of 40S, due to codon-anticodon complementarity, causes the release of eIF1 and increases the interaction between 40S and eIF1A (Maag et al., 2005; Unbehaun et al., 2004). Since eIF1 is required for the open conformation of 40S, recognition of the start codon and eIF1 release cause an irreversible commitment to translation. Second, the release of eIF1 allows eIF5 to induce hydrolysis of the GTP-coupled TC, which in turn induces the dissociation of GDP-eIF2 (Kapp and Lorsch, 2004). Third, the subunit joining factor GTP-eIF5B is recruited to 40S, through binding to eIF1A. This recruitment is not exclusive to 40S at the start codon, but GTP-eIF5B favors binding to eIF1A in the rearranged ('closed') conformation (Acker et al., 2006; Nag et al., 2016). Fourth, 60S is recruited to the start site via GTP-eIF5B, which in turn induces GTP hydrolysis and release of eIF1A-GDP-eIF5B (Nag et al., 2016). Finally, eIF3, eIF5, and possibly members of the eIF4F complex dissociate from the ribosome and the ribosome continues with elongation cycles, although the exact timing and mechanisms are unknown.

Taken together, start site selection involves several steps and the identification of a start site in a preferably strong Kozak context. Most of these steps are induced by changes in affinity upon release of another factor during a previous step. Translation initiation is therefore an irreversible process. Once begun, the process cannot be reversed unless the dissociated factors are recruited again. The individual steps during start site selection may have variable success-rates and can result in alternative translation initiation.

I

II

III

IV

V
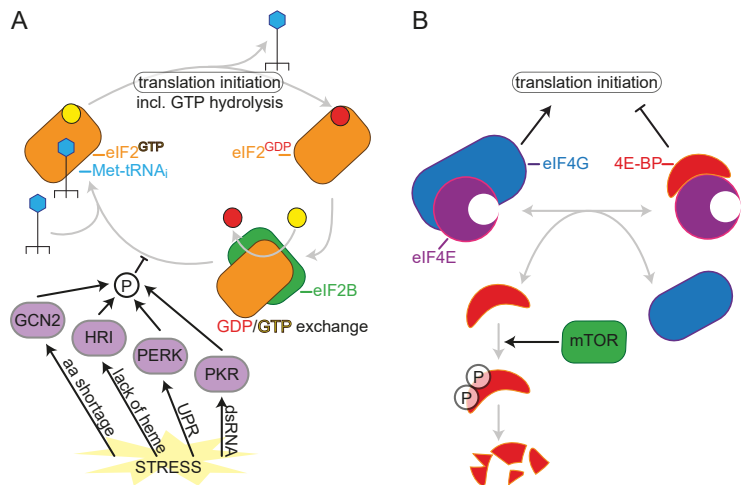
VI

&

REGULATION OF TRANSLATION INITIATION

Regulation of translation initiation mainly depends on two pathways that control the availability of key eIFs and that can thereby modulate initiation rates (Fig. 4).

eIF2 and integrated stress response

Formation of a TC to deposit a Met-tRNAi at a start codon requires a GTP-coupled eIF2, which is hydrolyzed during this deposition. To produce a next TC, GDP-eIF2 is recycled into GTP-eIF2 by its Guanine exchange factor (GEF) eIF2B (Fig. 4A). eIF2B levels in cells are limited and continuous translation with efficient eIF2 recycling therefore relies on rapid GDP/GTP release. Several stressed conditions can cause rapid reduction of eIF2B availability, resulting in a strong repression of almost all translation in the stressed cells. Stress pathways lead to the activation of four kinases: GCN2 upon amino acid shortage, HRI if heme levels are low, PERK as a result of the unfolded protein response, and PKR after recognition of dsRNA indicative of a viral infection (Taniuchi et al., 2016). Activation of any of these kinases results in phosphorylation of eIF2. This phosphorylation strongly increases the affinity between eIF2 and eIF2B (Kashiwagi et al., 2016). Consequently, eIF2 release from eIF2B is impaired, leading to a shortage of eIF2B and lack of GTP-eIF2. Although translation of some mRNAs can be enhanced under stressed conditions (e.g. ATF4, see below) or some mRNAs less sensitive to translation repression by eIF2 phosphorylation (e.g. some ISGs see section 'Antiviral response and inhibition of virus' (Rath et al., 2019)), stress pathways generally cause a rapid and potent inhibition of almost all translation.

Figure 4. Regulation of mRNA translation initiation.

(A) During translation initiation, GTP is hydrolyzed and releases the initiator tRNA. GEF eIF2B replaces GDP with GTP so that a new TC can be formed. However, eIF2^a phosphorylation downstream of stress-induced kinases prevents release of eIF2 from eIF2B, causing general inhibition of translation. (B) Recruitment of PIC to an mRNA depends on eIF4E-eIF4G and 4E-BPs compete with eIF4G. Under homeostatic conditions, mTOR is active and phosphorylates 4E-BP, which prevents eIF4E binding and phosphorylated 4E-BP is degraded. Active mTOR thus supports translation initiation. Abbreviations: General Control Non-derepressible 2 (GCN2); Heme Responsive Inhibitor (HRI; PKR-like ER-resident kinase (PERK); Protein Kinase R (PKR); amino acid (aa); unfolded protein response (UPR); double-stranded RNA (dsRNA); eIF4E binding protein (4E-BP); mammalian target of rapamycin (mTOR).

eIF4E availability downstream of mTOR

Recruitment of the 43S PIC relies on eIF4E-cap and eIF4G-eIF4E binding. The availability of the eIF4G-pocket on eIF4E can be controlled to regulate translation initiation (Fig. 4B). Inhibitory proteins, called eIF4E binding proteins (4E-BPs), compete with eIF4G for binding to eIF4E. If an eIF4E molecule is bound by a 4E-BP, the eIF4E cannot be bound by eIF4G. Thereby, 4E-BPs interfere with the recruitment of 43S to mRNAs and prevent translation initiation (Peter et al., 2015). Under homeostatic conditions, the mTOR pathway is active. 4E-BPs are among the mTOR targets. Consequently, phosphorylation by mTOR prevents 4E-BP binding to eIF4E and mTOR prevents translation inhibition by 4E-BPs (Gingras et al., 1999; Thoreen et al., 2012). Additionally, phosphorylated 4E-BPs are rapidly degraded, further supporting the mTOR role in facilitating mRNA translation (Yanagiya et al., 2012). Imbalances in eIF4E levels can drive tumorigenesis (Siddiqui and Sonenberg, 2015; Truitt et al., 2015) and the eIF4E/4E-BP ratio can be used as a biomarker to predict the efficiency of some oncology therapies (Alain et al., 2012; Smit et al., 2020), illustrating the importance of proper control of eIF4E through 4E-BPs.

Although most mRNA translation depends on eIF4E-dependent 43S recruitment to mRNAs, the translation of some genes is particularly sensitive to mTOR inhibition. Specifically, translation of mRNAs with 5' terminal oligo-pyrimidine tracts (TOP mRNAs) is strongly suppressed upon mTOR inhibition (Damgaard and Lykke-Andersen, 2011). TOP genes include some highly expressed housekeeping genes, whose expression is essential to execute basic cellular functions, such as ribosomal proteins, translation factors, and cytoskeletal proteins (Iadevaia et al., 2008; Thoreen et al., 2012). In contrast to TOP mRNAs, some other mRNAs are resistant to mTOR-mediated regulation of translation, for example mRNAs encoding histones or certain transcription factors (Thoreen et al., 2012).

Collectively, translation initiation is generally subjected to regulation at the level of ribosome assembly (through eIF2) and ribosome recruitment (through eIF4E). Although these regulatory steps target universal translation initiation events, the consequences of inhibiting these steps can vary for different genes. Insight into the dynamics of translation initiation of single mRNAs of various genes under different conditions may help to understand the exact mechanisms controlling mRNA translation initiation and reveal conditions that enable mRNAs the escape global regulation.

Additional elements affecting translation initiation

In addition to regulation of eIFs to control translation initiation, several features can affect initiation efficiency at the single gene or single mRNA level. These features may be either a permanent mRNA feature (i.e. the mRNA sequence) or a transient feature (i.e. may be lost or acquired over time and/or not present on all mRNAs from a gene) (Sonneveld et al., 2020).

The strength of the Kozak sequence strongly determines the translation efficiency of mRNA. If the sequence surrounding the start codon deviates a lot from the optimal Kozak sequence, the translation initiation on that start codon is inefficient (Kozak, 1978, 1987; Noderer et al., 2014). In addition to the Kozak sequence (-6 to +1 nt from AUG), several other nucleotides can impact start site selection as well. During start

codon selection, nucleotide -17 to -8 interface with eIF3 and the stability between eIF3 and the mRNA can contribute to successful start codon selection (Pisarev et al., 2008). Another permanent mRNA feature contributing to translation efficiency is the distance between the cap and the start codon. Start codon recognition is impaired on mRNAs with a short 5'UTR (<20 nt) (Kozak, 1991), as the 43S may have already 'passed' the start codon upon recruitment to the mRNA. In contrast, if the 5'UTR is very long, the odds of premature initiation on a non-canonical start codon may increase. Additionally, mRNA isoforms with long 5'UTRs may harbor regulatory features that are absent in the short isoform (e.g. Emi1 uORFs, (Yan et al., 2016)). The length of the 5'UTR can thus contribute to variations in initiation efficiencies between genes.

Post-transcriptional modifications of an mRNA molecule are examples of transient mRNA features that can contribute to initiation efficiency. For instance, one of the most studied RNA modifications - N6-methyladenosine (m6A) – has been shown to stimulate translation indirectly through stabilizing mRNAs and directly by facilitating 43S recruitment (Coots et al., 2017; Mauer et al., 2017; Meyer et al., 2015). Another example of a transient mRNA feature contributing to heterogeneity in translation is the ability of the 5'UTR to form structures. mRNAs with extensive 5'UTR structures are generally translated inefficiently and their translation depends heavily on eIF4A (Pisareva et al., 2008; Svitkin et al., 2001).

In addition to eIFs, mRNA can be bound by many RNA binding proteins (RBPs). RBP binding may influence initiation efficiency. The binding of an RBP to an mRNA depends on a certain binding motive in the mRNA, the presence of a co-factor that (de)stabilizes the RBP-mRNA interaction, and the subcellular localization of both the RBP and the mRNA. For example, translation of caudal mRNA is limited to a subcellular region in the *Drosophila* embryo due to the presence of its translational repressor in the rest of the embryo (Niessing et al., 2002).

Heterogeneity in the exact composition of ribosomes translating an mRNA is another transient feature that can impact initiation efficiency. Although the ribosome traditionally was considered to have a fixed composition, recent studies have indicated that the exact composition of ribosomal proteins is variable (Genuth and Barna, 2018). Presence or absence of a certain ribosomal protein can influence translation differentially. The ribosome composition can differ between tissues as some ribosomal proteins are differentially expressed. Furthermore, various ribosome compositions can be present simultaneously in the same cell (Guimaraes and Zavolan, 2016; Shi et al., 2017b). Translation of mRNAs with any of the aforementioned features may be more susceptible to changes in the ribosome composition, adding yet another layer of complexity to regulation of translation.

NON-CANONICAL TRANSLATION INITIATION
The canonical translation initiation model explains the different steps required to bring a ribosome to a start codon (Fig. 3). Additionally, the regulatory pathways and various mRNA features help explaining how mRNA translation is regulated and why some mRNAs are more sensitive to regulation than others. However, several exceptions to the initiation model exist in eukaryotic cells which complicate the understanding of translation initiation and its regulation.

Upstream open reading frames (uORFs)

According to the canonical model, translation is initiated on the first start codon, leading to the translation of its ORF and termination at the stop codon. Consequently, only a single ORF per mRNA can be translated. However, approximately half of all mammalian genes encode at least two ORFs, and the majority of the alternative ORFs are located upstream of the canonical ORF (upstream ORF-uORF) (Calvo et al., 2009; Johnstone et al., 2016; Lee et al., 2012) and can be translated efficiently (Ingolia et al., 2011). Although the presence of uORFs generally results in reduced translation of the main ORF, the main ORF can still be translated. A single gene can thus encode multiple polypeptides, raising many questions about translation initiation on uORF-containing mRNAs.

Translation of the canonical ORF downstream of an uORF depends either on leaky scanning of the uORF or reinitiation after uORF translation. The leaky scanning model predicts that uORF start site selection is inefficient. Scanning ribosomes often fail to initiate on an uORF, they rather continue scanning to the next start site (i.e. the canonical ORF's start site). Indeed, non-canonical start codons are abundant among uORFs and uORFs are translated with variable efficiency (Ingolia et al., 2011).

Alternatively, translation of the mainORF downstream of an uORF can take place due to reinitiation. Instead of dissociating from the mRNA after completing uORF translation, the small ribosomal subunit remains associated with the mRNA and resumes scanning to then translate the next ORF. To be able to scan, some of the eIF may have to be recruited or remain present during the uORF translation. For example, eIF3 release takes place during the first cycles of translation elongation and could facilitate the recruitment of other eIFs after translation of an uORF (Archer et al., 2016; Jackson et al., 2012; Wagner et al., 2020; Zhang et al., 2008). MainORF translation is more efficient downstream of a short uORF, supporting a role for slowly released eIFs in rerecruiting other eIFs (Calvo et al., 2009; Johnstone et al., 2016). Recruitment of a new TC is crucial for selection of a next start site, since a scanning complex without a Met-tRNAi cannot induce initiation. Based on analyses of mRNAs with multiple uORFs, reinitiation is limited by the speed of recruiting a new TC. For example, ATF4 mRNAs contain multiple uORFs and the first uORF is translated efficiently (Ferreira et al., 2014; Palam et al., 2011). Under homeostatic conditions, a new TC is recruited after uORF1 translation, leading to the translation of the second uORF. In stressed conditions, however, TC levels are low and recruitment of a new TC is inefficient. Consequently, the reinitiating scanning ribosome cannot initiate translation on the start site of uORF2, as the uORF2's start site cannot be recognized in the absence of the new TC. Instead, the scanning continues, providing more time to recruit a new TC, which then supports start site selection of the next start site: the mainORF's start site. Under stressed conditions, the translation of ATF4's mainORF is therefore increased, because the translation of the second uORF is decreased. The ATF4 example illustrates that reinitiation can take place downstream of uORFs. It also illustrates how the presence of uORFs can affect translation efficiency of the mainORF

Both leaky start site selection of the uORF and reinitiation downstream of an uORF demonstrate that the canonical translation model is incomplete. For example, the frequency of either mechanism may vary between genes, between mRNAs of the same gene, and over time. Detailed analysis of uORF and mainORF translation is

I

II

III

IV

V

VI

&

required to understand the role of uORFs in controlling translation.

eIF4E-independent ribosome recruitment

In contrast to the canonical translation initiation model, some mRNAs are translated in an eIF4E-independent manner. Some mRNAs escape repression of translation by 4E-BPs by recruiting 43S through alternative cap-recognition. For example, translation of c-JUN is insensitive to 4E-BPs and 43S is recruited to the mRNAs through interaction between eIF3d and the cap (Lee et al., 2016).

Another type of eIF4E-independent ribosome recruitment relies on internal ribosome entry sites (IRESs): RNA structures that support the recruitment of a ribosome. Many viruses depend on IRESs to express their transcripts, allowing the virus to bypass eIF4E/4E-BP regulation and facilitating the translation of multiple ORFs from a single transcript. The recruitment of 40S by an IRES differs among viruses. Some viruses, like CVB3, only bypass eIF4E's requirement, but eIF4G is still necessary for 43S recruitment (see section 'Viral replication and gene expression'). The IRES of some other viruses, such as Cricket Paralysis Virus, can recruit 43S independently of any eIF4F members by direct interaction with eIF3 and the 43S is deposited on the start codon directly (Neupane et al., 2020). IRES-mediated recruitment is also prevalent in human cells. Thousands of sequences can function as an IRES by recruiting 43S in an cap-independent manner (Weingarten-Gabbay et al., 2016).

The uORF and eIF4E-independent examples illustrate that translation initiation can deviate from the canonical initiation model and that alternative initiation can affect the regulation of mRNA translation.

APPROACHES TO STUDY MRNA TRANSLATION

Many approaches can be employed to study mRNA translation initiation, including structural analysis of the various complexes involved in initiation, measuring the protein output by luciferase assays, analyzing the ribosome density by polysome gradients, or using ribosomal profiling to analyze the distribution of ribosomes along mRNAs. However, these assays are limitingly suited to study the dynamics and regulation of translation initiation in detail. First, most assays require fixation or lysis of the sample, thereby preventing any analysis of the same sample at multiple time points. Second, most assays are based on pooling many cells together, which complicates analysis of heterogenous processes; a relatively infrequent process may be hard to detect. Even assays that enable single-cell analysis mostly measure the average behavior of many mRNAs in that cell and cannot distinguish individual mRNA molecules. Third, *in vitro* assays to study mRNA translation cannot mimic the *in vivo* conditions and variations to those conditions and are therefore ill-suited to investigate the entire complex process of translation initiation.

To overcome these limitations, several groups have developed single-molecule live-cell imaging approaches to study translation of single mRNAs (Morisaki et al., 2016; Pichon et al., 2016; Wang et al., 2016; Wu et al., 2016; Yan et al., 2016). These approaches are based on labeling of nascent polypeptides by systems such as the SunTag fluorescence labeling strategy (Tanenbaum et al., 2014). This strategy is based on stable expression of a genetically encoded, fluorescently-labeled intrabody

(called SunTag antibody or STAb) and a peptide that can be bound by STAb. By introducing multiple copies of the peptide at the amino terminus of a protein of interest, translation of the mRNA encoding the protein of interest can be observed by STAb fluorescence on the mRNA. As multiple ribosomes can translate a single mRNA simultaneously, many STAbs can accumulate on a single mRNA, resulting in readily detectable fluorescence foci. The brightness of these foci corresponds to the number of ribosomes engaged in translating the mRNA molecule. Additionally, fluctuations in mRNA translation result in temporal changes in fluorescence. Using live-cell single-molecule imaging of mRNA translation has uncovered many aspects of mRNAs and their translation, including mRNA-to-mRNA variations in translation efficiency, the effect of uORFs on translation, differences in translation in different subcellular locations, and various mechanisms involved in mRNA decay (e.g. Boersma et al., 2019; Hoek et al., 2019; Mateju et al., 2020; Moon et al., 2019; Ruijtenberg et al., 2020; Wang et al., 2016; Yan et al., 2016). Collectively, these findings illustrate how live-cell single-molecule imaging contributes to understanding dynamic processes, such as mRNA translation.

In summary, mRNA translation, and specifically initiation, is imperative in producing proteins and in shaping the proteome. Translation initiation depends on many proteins and various features of the mRNA affect the efficiency of initiation. Over the course of the different processes involved with initiation, the composition of complexes or the relative contribution of a factor or feature may change. Even though the model of translation initiation thoroughly describes how canonical translation initiation may take place, the details of many aspects of initiation remain elusive. Most assays used to study translation rely on bulk or fixed analyses, and thus lack the resolution to study how translation initiation happens on individual mRNAs by single ribosomes over time in live cells. Live-cell single-molecule imaging assay can help to decipher translation initiation dynamics, efficiencies, and modes of regulation.

II. VIRAL REPLICATION AND GENE EXPRESSION

THE DOGMA CHALLENGE OF A VIRUS

Viruses are typically small and do not contain all components to enable expression of their genes. To produce viral proteins, viruses therefore hijack cellular dogma components (e.g. cellular ribosomes). Concomitantly, the viral genome also serves as template for viral genome replication. Both synthesis of viral proteins and replication of genomes are required to assemble new virions. These two processes are mutually exclusive on the same genome for most viruses. However, these processes also support each other: high replication rates result in the production of many genomes that can then serve as template for viral protein synthesis. However, replication often relies on viral proteins that need to be produced prior to replication. Therefore, viruses must coordinate gene expression and replication for optimal production of new virions.

Viruses are subdivided into families. Members of each family have evolved different

ways to enable viral protein production and use different approaches to replicate their genome. Here, I focus on the order of the positive-sensed single-stranded RNA (+ssRNA) *Picornavirales,* using enterovirus Coxsackievirus B3 (CVB3) as an example, and the order of negative-sensed single-stranded RNA (-ssRNA) *Mononegavirales*, exemplified by the paramyxovirus respiratory syncytial virus (RSV).

PICORNAVIRALES

The order of *Picornavirales* includes well-known pathogens, such as poliomyelitis-inducing poliovirus, foot-and-mooth-disease causing EV- A71, common-cold causing rhinoviruses, and cardiomyopathy inducing CVB3 (Baggen et al., 2018). As the name suggests, picornaviruses are small RNA viruses. The virion comprises an unenveloped capsule of approximately 30 nm that surrounds the viral genome. The genome consists of a single stranded, positive-sensed RNA molecule (+ssRNA; referred to as viral RNA (vRNA)) of approximately 7.5 kb and

Figure 5. Schematic picornavirus virion.

A single picornavirus virion (left) and the composition of its genome (right). The picornavirus genome is coupled to a Vpg at the 5' end and is polyadenylated at the 3' end. The genome contains an IRES in the 5'UTR and it encodes one polyprotein, divided into P1, P2, and P3, which are further subdivided into the individual proteins into which the polyprotein is processed. Genome is not depicted at scale. See main text for further details.

encodes a single polyprotein. Upon production, the polyprotein is proteolytically processed by viral proteases to form ~11 proteins, including 4 capsid proteins (VP1-4), 2 proteases ($2A^{pro}$ and $3C^{pro}$), and an RNA dependent RNA polymerase ($3D^{pol}$) (Fig. 5).

Picornavirus life cycle

A picornavirus infection includes several steps, starting with the entry into a host cell, followed by translation and replication. Ultimately, new virions are assembled and released to infect other cells. The individual steps in the picornaviral life cycle are explained below (also see Fig. 6), focusing on the early events in infection.

1. Entry

The life cycle of a picornavirus infection is very similar for all members of the family. The main difference between members is the receptor that the virions can bind to, which is one of the determinants of the viral tropism. For example, CVB3 utilizes the Coxsackie and Adenovirus receptor (CAR) to enter a host cell (Lonberg-Holm et al., 1976). CAR contributes to the stability of tight junctions: protein complexes that form a tight interaction between cells such as polarized epithelial cell layers

and cardiac muscles (Cohen et al., 2001). CVB3 infection is initiated by entry of CVB3 into epithelial cells along the gastrointestinal tract. Subsequently, CVB3 infects other CAR-positive cells, such as cardiomyocytes which can lead to CVB3-induced cardiomyopathy (Garmaroudi et al., 2015; Sin et al., 2015). Poliovirus entry into host cells depends on attachment to the poliovirus receptor (PVR; also called CD155)(Mendelsohn et al., 1989), expressed on cells along the gastrointestinal tract and on cells of the nervous system. This explains how polioviruses can spread between individuals and can cause poliomyelitis (Bowers et al., 2017). Interestingly, PVR upregulation has been detected in many tumors, raising the possibility to use poliovirus as an oncolytic virus (Brown and Gromeier, 2015; Takai et al., 2008). An oncolytic viral therapy is based on a virus that specifically infects and kills cancer cells with minimal effect on non-cancerous cells. The upregulation of PVR on tumor cells offers the possibility to use poliovirus's tropism to target a tumor. To avoid an infection in non-cancerous cells, which may lead to poliomyelitis, a chimeric poliovirus strain is used that cannot be expressed in neuronal cells (Brown and Gromeier, 2015; Gromeier and Nair, 2018).

Upon attaching to the viral receptor, picornaviruses enter host cells through receptor-mediated endocytosis. Next, a pore in the endocytotic vesicle is formed and the vRNA is released into the host cell's cytoplasm (Baggen et al., 2018). Genome release takes place near the plasma membrane (100-200nm from membrane) and rapidly after infection (~10-20 minutes) (Brandenburg et al., 2007). The released viral genome is naked: except for the 5'coupled Vpg, no viral proteins associate with the genome upon entry into a host cell.

2. Translation

Translation of the vRNA into the viral polyprotein depends on the recruitment of host cell's translation machinery. In contrast to translation initiation on host cell's mRNAs, recruitment of eIFs and ribosomes to the vRNAs is not cap-dependent but is mediated by an IRES in the 5'UTR. The picornavirus' IRES bypasses the requirement of eIF4E by directly recruiting eIF4G and other eIFs (Lee et al., 2017; Souii et al., 2013). Additionally, trans-acting IRES factors, such as PCBP2, bind to non-IRES structures in the vRNA and facilitate ribosome recruitment (Sean et al., 2009). Mutations in loop V of the IRES can potently inhibit translation initiation, as 3 point mutations are responsible for the attenuation of poliovirus in the Sabin vaccine (Ben M'hadheb-Gharbi et al., 2006; Svitkin et al., 1990). This indicates that efficient vRNA translation is essential for a picornavirus to establish a successful infection.

3. Proteolytic processing of polyprotein

Translation of vRNA results in a polyprotein that is proteolytically processed into the individual viral proteins. This processing is performed by the viral proteases 2A[pro], 3C[pro], and the processing intermediate 3CD. The exact order of processing differs between genera of picornaviruses. For enteroviruses, 2A[pro] is responsible for cleaving between P1 and P2 and 3C[pro]/3CD separates P2 and P3. Further processing of P1, P2, and P3 into the viral proteins is then mostly done by 3C[pro]/3CD (Sun et al., 2016). Processing of VP4-VP2 only takes place during the assembly step (see step 6), illustrating that the timing of processing is heterogeneous. Furthermore, throughout

I

II

III

IV

V

VI

&

Figure 6. Picornavirales lifecycle.

Schematic illustrating the key steps during a picornavirus infection. Note that some picornaviruses contain a leader protein at the N-terminus of the polyprotein, which is not depicted in this schematic, as the exemplifying CVB3 does not encode an L protein. See main text for details on each step.
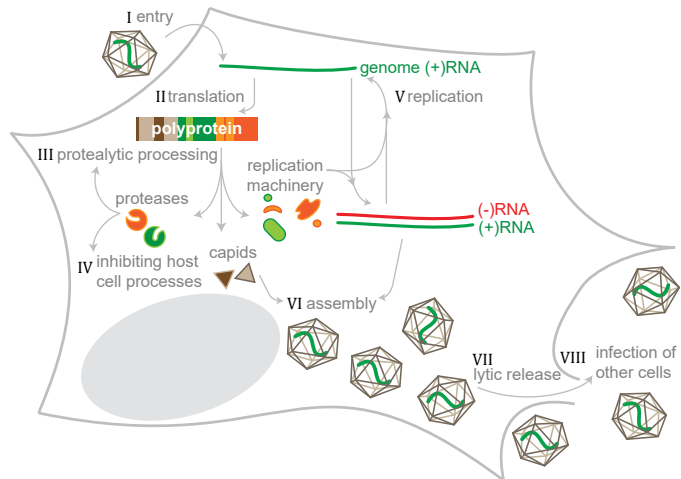
the course of an infection, nascent polypeptides, cleavage intermediates and fully processed single viral proteins can be present in a cell. It is likely that the exact timing and efficiency of polyprotein processing may differ during various stages of an infection as the concentration of proteases and protease-targets changes. Currently, knowledge on the details of polyprotein processing during different moments in the viral life cycle is limited.

4. Proteolytic processing of host cell's proteins

In addition to separating viral proteins, viral proteases can also cleave host cell's proteins. During a CVB3 infection, many host cell proteins are cleaved, including gene expression proteins and innate immune regulators (Saeed et al., 2020). For example, picornavirus' proteases cleave several nuclear pore complex proteins, leading to impaired nuclear transport and disruption of host cell's gene expression (Flather and Semler, 2015). Similarly, CVB3's 2A[pro] efficiently cleaves eIF4G and the cleavage product is no longer capable of engaging in cap-dependent translation initiation, but it does support IRES-dependent translation (Lamphear et al., 1995). In this way, vRNAs can be translated, while host cell's mRNAs translation is impaired. Another group of protease targets are proteins involved in innate antiviral response. For example, CVB3's 3C[pro] cleaves MAVS, which results in attenuation of the host cell's ability to launch an antiviral response (Mukherjee et al., 2011).

Interestingly, the timing of host cell's protein cleavage is not uniform for all targets. Some proteins are cleaved rapidly during infection, whereas other protease targets are cleaved slowly (Saeed et al., 2020). As timing of hijacking host cell's gene expression and repressing antiviral components likely impacts the outcome of an infection, insights into the dynamics of host cell's proteins cleavage may help to understand viral infection efficiency.

5. Replication & the translation-to-replication switch

To generate new virions, new vRNAs need to be produced during an infection. During

a picornavirus infection, dynamic membranous structures, so called replication organelles, are formed that aid in replication by creating an optimal replication environment (Schaar et al., 2016). Replication of picornaviruses' genome depends on the 3D^{pol}, which requires vRNA translation and polyprotein processing. The first step in replication is the synthesis of a complementary negative-sensed (-RNA) molecule, using the incoming vRNA as a template. Next, this -RNA serves as a template to produce new +RNAs, which can then be translated to form more viral proteins and serve as another template for replication to generate more vRNAs (Baggen et al., 2018). Next to 3D^{pol}, another viral protein (3B or Vpg), is required for replication. Vpg is coupled to a Uridyl-group by interacting with a cis-acting element (CRE) and the Uridyl-Vpg serves as a primer for both -RNA and +RNA synthesis (Paul and Wimmer, 2015). In addition to priming replication, Vpg binding to the 5' end of a vRNA also serves to escape decay; RNA molecules without an cap are targeted for degradation, which is circumvented by Vpg (Ullmer and Semler, 2016).

Although the core elements required for replication have been identified, many questions regarding the dynamics and efficiency of replication remain unanswered. For example, how a virus controls the switch from translating an vRNA to replicating an vRNA is unclear. As vRNA in the infecting particle is naked (i.e. does not carry viral proteins), translation of the vRNA and processing of the polyprotein is required to form 3D^{pol}, before replication can be initiated. At some point, however, the vRNA must switch from translation to replication. Some factors have been identified that can contribute to this switch. For example, binding of PCBP2 to an RNA structure neighboring the IRES supports translation and PCBP2 binding is reduced by 3CD (Gamarnik and Andino, 1998; Sean et al., 2009). Treatment with translation inhibitor puromycin induces -RNA production, suggesting that replication is induced if translation is impaired (Barton et al., 1999). In this way, accumulation of 3CD can support the translation-to-replication switch by preventing translation initiation, leading to initiation of replication. However, this model does not account for efficient translation-to-replication switching early in infection. The single vRNA from the infecting virion may not produce sufficient 3CD to enable a translation-to-replication switch early in infection. A comparable complication may arise late during infection, when 3CD levels accumulate. In the PCBP2-model, no translation but only replication would be possible late in infection, as 3CD concentrations are higher late in infection compared to early in infection. Therefore, more insight into the dynamics of translation-to-replication switching is required to decipher the efficiency of producing new vRNA and new virions.

Another unknown aspect of replication concerns the ratio between -RNA and +RNAs. Once the replication stage of an infection has started, both -RNAs and +RNAs can be formed and -RNA formation is necessary to generate +RNAs. If all Vpg and 3Dpol are occupied with -RNA synthesis, no +RNAs can be formed. Contrary, if only a single -RNA molecule is produced to then form +RNAs, the +RNA production rate may be suboptimal. Therefore, it is likely that mechanisms control the -RNA versus +RNA synthesis. The +:- RNA ratio during poliovirus infection is ~20:1 (Dave et al., 2019; Schulte and Andino, 2014), which may reflect the optimal ratio to generate maximal +RNAs. However, it is unknown whether this ratio remains equal during early or late infection and how it impacts infection efficiency.

6+7. Assembly & release

Newly synthesized vRNAs and capsid proteins are assembled into provirions. In the presence of vRNA, the final proteolytic processing step (i.e. VP4-2 cleavage into VP4 and VP2) takes place, resulting in mature virions in the cytoplasm of the host cell (Jiang et al., 2014). Release of virions from the infected host cells depends mainly on a lytic pathway. The virion-producing cell bursts and the newly synthesized virions are released to infect other cells, in which the viral life cycle will start with the first step again. Release can also take place in a non-lytic manner, for example through autophagosomal exit (Lai et al., 2016). The manner of release may have implications for the next infection cycle. If a cell is infected by a single virion or a phagosome containing multiple virions, the viral translation, proteolytic processing, and replication during early infection may be different.

MONONEGAVIRALES

Mononegavirales are membrane-enveloped negative-sensed RNA viruses. This order includes well-known pathogens like mumps virus, measles virus, Ebola virus, and RSV. As indicated by the name, the *Mononegavirales* genome is located on a single negative-sensed molecule. The exemplifying *Mononegavirales* RSV has a dynamics shape, ranging from spherical (~100nm) to filamentous shapes (up to 1μm) (Liljeroos et al., 2013). The RSV genome is 15.5 kb and contains 10 genes, encoding 11 proteins including RNA dependent RNA polymerase (L) (Fig. 7A-B). In contrast to the positive-sensed picornaviruses, *Mononegavirales'* genome cannot be directly translated into viral proteins. Instead, viral protein synthesis depends on transcription of viral genes into viral mRNAs that are then translated into viral proteins. Viral gene expression is directed by regulatory sequences in the genome (Gene Start [GS] and Gene End [GE]; Fig. 7C). In contrast to the picornaviruses, *Mononegavirales'* genomes are not naked, but tightly wrapped in nucleoproteins (Ns) (Tawar et al., 2009), required for genome replication and packaging.
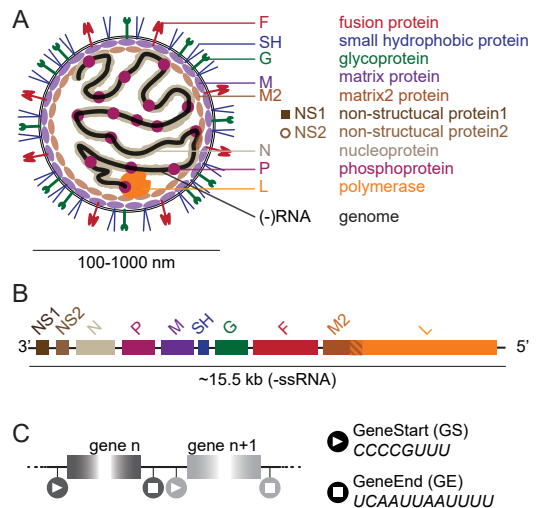
Figure 7. Schematic RSV virion and genome composition.

(A) A single RSV virion (left) in which individual viral proteins and their full name are indicated (right). Note that NS1 and NS2 are also show, even though these proteins are not part of the viral particle. (B) RSV genome showing the order and relative size of the individual viral genes. Note that overlap between M2 and L gene is not depicted at scale; the actual overlap is only 60 nucleotides. Color and name of genes corresponds to the name of the viral proteins encoded by the genes as indicated in A. (C) Each RSV gene is flanked by a GS and a GE. Sequences (italics) are depicted in genomic 3'-5' orientation.

RSV infection can cause severe respiratory distress, including pneumonia and bronchiolitis and -prior to the COVID-19 pandemic- RSV was the leading viral cause of lower respiratory infections requiring hospitalization particularly for infants, elderly, and immunocompromised patients (Htar et al., 2020; Shi et al., 2017a). Although several promising antiviral prophylactic or curative treatments are in design, treatment options for severe RSV infections are currently limited. Deciphering early events during an RSV infection may contribute to design or improvement of successful therapies (Hu et al., 2020).

RSV life cycle

RSV infection is initiated by entry into a host cell, followed by transcription and replication of the viral genome. Ultimately, new virions are assembled and secreted to infect other cells. The individual steps in the RSV life cycle are explained below (also see Fig. 8), particularly focused on the early events in infection.

1. Entry

The exact mode of RSV attachment, a main tropism-determining factor, is not fully understood. Defining the factors required for attachment has been complicated by the discrepancy between *in vitro* and *in vivo* studies. RSV readily infects cultured cells, whereas infection in the respiratory tract is mostly limited to ciliated epithelial cells (Mastrangelo and Hegele, 2013). Heparan sulfate is a possible explanation for this difference; it is absent on airway epithelial cells *in vivo,* whereas it is present in high levels on the surface of cultured cells, where it binds well with glycoprotein (G). The high abundance of heperan sulfate during cell culture may explain why cells in culture are readily infection, whereas the RSV infection efficiency *in vivo* is low (Hu et al., 2020).

RSV virions contain three types of transmembrane proteins (small hydrofobic protein [SH], G, and fusion protein [F]) (Fig. 7A). G and SH are mainly involved in bypassing antiviral host cell responses, but may also contribute to RSV entry by increasing the 'stickyness' of an RSV particle to a host cell (Griffiths et al., 2020; Hu et al., 2020). The main factor required for RSV entry into a host cell is a homotrimer F complex and its conformational change from the native, pre-fusion state into a post-fusion state, which results in pore-formation leading to release of RSV genome into the host cell's cytoplasm. F-dependent entry is supported by nuclear protein nucleolin, whose transport to the cell membrane requires F-IGFR binding followed by PKCζ signaling (Griffiths et al., 2020; Tayyari et al., 2011). Cell type specific PKCζ expression may explain the *in vivo* preference of RSV to infect ciliated airway epithelial cells. In this way, RSV-IGFR binding results in entry specifically into cells in which the entry factor nucleolin can be recruited to the cell membrane (Griffiths et al., 2020). Collectively, RSV entry depends on multiple factors and whilst deciphering the entry process, the natural infection environment should be considered.

2. Replication & Transcription

The RSV genome can both be transcribed and replicated. As the genome itself is non-coding, the polymerase machinery must be brought in from the infecting virion. The

Figure 8. Mononegavirales lifecycle.

Schematic illustrating the key steps during a Mononegavirales infection, using RSV as an example. Note that transmembrane proteins SH, G, and F are translated and processed using the ER-Golgi secretory pathway, which is not included in the schematic for simplicity. See main text for details on each step.
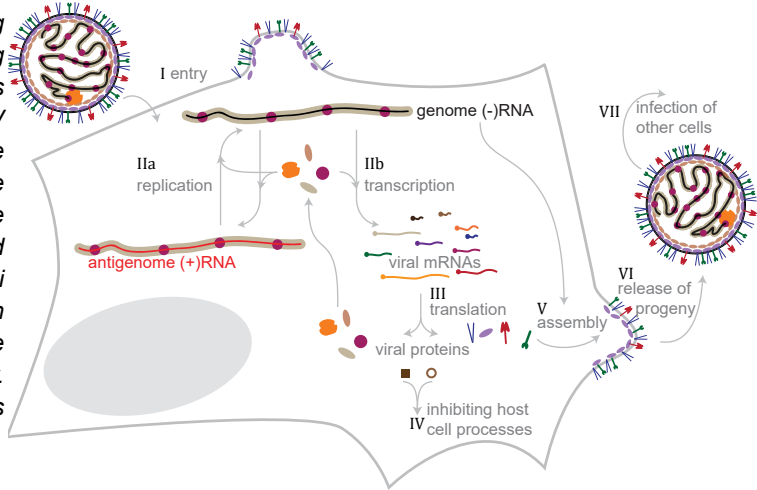
naked genome cannot facilitate any process but relies on proteins already produced and transported along. Both processes initiate on the -RNA genome template, rely on the polymerase machinery (polymerase and accessory proteins) and result in +RNA molecule products. The paradoxical involvement of the same elements in both processes, illustrates the central dogma challenge of RSV: to optimally generate new virions, both replication and transcription need to take place and both processes support each other by generating more templates or proteins required for the processes. The dynamics and regulation of transcription, replication, and switching between the two processes are likely of paramount importance for the success-rate of an RSV infection.

The transcription products (see section 2b) are viral mRNAs encoding the viral proteins. Each viral mRNA molecule is positive sensed, capped, polyadenylated, and represents one viral gene. In contrast, the replication products (see section 2a) are either positive or negative sensed, contain the entire (complementary) viral genome, and cannot be translated. Furthermore, viral mRNAs are naked, whereas the genome and antigenome are encapsidated. During transcription, L utilizes gene starts and stops, whereas these regulatory sequences are not recognized during replication. The encapsidation state of the L product affects L' processivity and thereby L's ability to ignore the gene stop sequences (during replication) or to abortively read the genome (during replication) and several factors may be responsible for coordinating transcription versus replication (Cao et al., 2021; Noton and Fearns, 2015; Noton et al., 2019). First, the mode of L loading onto the 3' end of the genome determines whether transcription or replication is preferred (Cowton et al., 2006; McGivern et al., 2005; Tremaglio et al., 2013). Second, (anti) genome production through replication requires N and P, so high concentrations of these proteins are likely required for successful replication, as is suggested for other *Mononegavirales* (Horikami et al., 1992; Noton and Fearns, 2015; Patton et al., 1984). Third, the relative abundance of the different protein products of the M2 gene may facilitate transcription or replications. M2.1 and M2.2 support transcription or replication respectively (Bermingham and Collins, 1999; Fearns and Collins, 1999; Hardy and Wertz, 1998).

Many questions regarding the dynamics of transcription or replication remain unanswered. It is unknown how RSV balances replication and transcription and whether the balance changes over the course of an infection. Furthermore, it is unclear whether replication and transcription can take place simultaneously on the same genome or whether the processes are mutually exclusive. In summary, both gene expression and replication are essential for virion production and insight into coordination between transcription and replication may help to understand RSV infections.

2a. Replication

To generate more viral genomes, which can serve as additional transcription templates and can be packed into new virions, the genome needs to be replicated. First, the negative-sensed genome is used to produce a complementary, positive-sensed antigenome, which can then serve as a template for the synthesis of new genomes. Antigenome production depends on the 3' Leader promoter in the genome. Specifically, the RSV RNA dependent RNA polymerase (L) must be loaded on the first nucleotide for the replication product to contain the encapsidation signal. Similarly, genome synthesis using the antigenome relies on Trailer promoter at the 3' end of the antigenome (correspondingly encrypted at the 5' end of the genome) (McGivern et al., 2005). N is brought to the (anti)genome by binding to P. This N-P interaction prevents aspecific N-binding to host cell's RNA (Cao et al., 2021; Castagné et al., 2004). The replication products form a tight helical RNP (RNA and protein) complex, consisting of the RSV (anti)genome and decameric nucleocapsids. Each N can interact with seven nucleotides and consequently, each (anti)genomic RNP comprises up to 2000 Ns (Cao et al., 2021; Tawar et al., 2009).

The ratio between genomes and antigenomes probably affects RSV infection efficiency. A high rate of antigenome synthesis may benefit further genome synthesis, but antigenome synthesis may also inhibit genome synthesis by occupying L, N, and P. This ratio may be controlled by M2-2, as M2-2 binds to and inhibits L, specifically M2-2 binding to L repressing +RNA synthesis (Bermingham and Collins, 1999; Kitagawa et al., 2009). Thus, multiple proteins (L, N, P, M2-2) are involved in RSV replication and their presence may affect replication efficiency.

The (anti)genomic RNP is a tight complex and accessibility for RNA-protein interactions may be limited (Tawar et al., 2009). As L moves along an (anti)genome molecule to produce a new RNA molecule, the RNA sequence must be exposed to L, raising questions related to the dynamics of the RNP structure. For example, does the (anti)genome release nucleocapsids, or do the nucleocapsids remain associated during polymerization? Whether or not an (anti)genome RNP may lose its proteins may have implications for replication. Does replication depend on *de novo* viral protein synthesis or can the replication product 'steal' nucleocapsids from its template?

2b. Transcription

RSV mRNAs are transcribed by L. This process depends on the gene regulatory sequences surrounding each gene (GS and GE, see Fig. 8C). Upon recognition of a GS, L initiates transcription on the first nucleotide of the GS and produces a cap

with its capping and cap-methylation domains. Although the process of capping is different from eukaryotic cap synthesis, a cap produced by L is chemically similar to the caps of host cell mRNAs (Cao et al., 2021; Ogino and Green, 2019). Next, L elongates and converts the genomic sequence into a transcript. M2-1 binds to L through P and this interaction is essential for transcription elongation (Blondot et al., 2012). In the absence of M2-1, L is prone to terminate prematurely, producing only short mRNA fragments (Fearns and Collins, 1999). Finally, L reaches the U-rich GE which is utilized to generate a polyA tail on the mRNA, by slipping on the U-rich tract, and to release the mRNA (Noton and Fearns, 2015). Single mutations in GE result in reduced release efficiency (Moudy et al., 2003), illustrating that accurate transcription strongly depends on the conserved GEs and GSs.

L remains associated on the genome after release of an mRNA and scans through the intergenic region to reinitiate transcription on the next GS. During polyadenylation on the upstream gene, M2-1 adopts a different conformation that supports L retaining on the genome. Additionally, transcription of downstream genes is impaired in the absence of M2-1 (Fearns and Collins, 1999; Hardy and Wertz, 1998; Leyrat et al., 2014). Retaining of L during crossing of the intergenic region can be inefficient, leading to dissociation of L after transcription of a gene. Consequently, RSV gene expression is thought to occur in a gradient in which the gene is always transcribed less efficiently than an upstream (3') gene (Noton and Fearns, 2015).

The gradient model assumes that (1) L is loaded on the 3' end of the genome, (2) L moves along the genome unidirectionally, and (3) transcription termination results in either release or continued scanning to the next GS. For several reasons, the gradient model may be too simplified and lacks details about RSV gene expression. First, gene M2 and L are partially overlapping: the GS of L is located upstream of M2's GE. According to the gradient model, the L gene would never be transcribed, as the L protein is unable to initiate transcription on the GS of the L gene. Based on a similar discrepancy of the gradient model of another *Mononegavirales* VSV (vesicular stomatitis virus), L may be able to move along the genome in the opposite direction to identify GSs (Barr et al., 2008). The efficiency of transcription reinitiation on an upstream GS depends on the distance between the GE and the GS. The GS-GE distance should be short (<200 nucleotides). Second, according to the gradient model, transcription termination on a GE and reinitiation on the next GS occur very accurately. However, single mutations in a GE can result in the production of polycistronic mRNAs, illustrating that transcription is very sensitive to divergences (Moudy et al., 2003). Accordingly, it is conceivable that alternative transcription may occur, for example by leaky GS reading resulting in skipping of a gene. Third, analysis of the RSV transcript abundance by RNA-seq or qPCR analysis has revealed that not all RSV genes follow the gradient model (Aljabr et al., 2016; Piedra et al., 2020). Especially G mRNAs seem too abundant to be in accordance with the gradient model. Although these studies mostly focus on time-points relatively late after the start of infection (>24hr), RSV mRNAs have long half-lives and the discrepancy between the gradient model and data is thus not explained by RSV mRNA decay (Piedra et al., 2020). Fourth, loading of L onto the genome, transcription, and transcription reinitiation strongly depends on a collection of RSV proteins. During an infection, inclusion bodies are formed, in which core elements of the viral gene expression machinery are concentrated to facilitate viral gene expreission and replication (Galloux et al., 2020; Rincheval et al., 2017). In contrast, at the onset of an infection

prior to any viral gene expression, the abundance of the viral proteins is low and no inclusion bodies are formed. Viral gene expression efficiency may change as inclusion bodies are formed. However, the gradient model does not account for changes in viral gene expression over the course of an infection. In sum, the gradient model may thus not completely explain how viral gene expression is coordinated and how the dynamics of viral gene expression may change during an infection.

Finally, various other aspects about the dynamics of RSV gene expression remain unclear. The recruitment of L to a genome, the polymerization rate, and the dissociation rate of L from a genome may greatly impact the viral gene expression efficiency. Especially at the onset of an infection, levels of L may be low and the manner of recruiting or retaining L could impact the absolute transcription output. Moreover, it is unclear how many L proteins could transcribe a single genome at the same time and whether simultaneously transcribing Ls function independently, synergistically, or antagonistically. Comparably, it is unclear whether transcription on different genome molecules is coordinated. Prior to the formation of inclusion bodies, each genome may be expressed autonomously, and its expression may only depend on the presence of viral proteins on a genome. Alternatively, transcription could be synchronized on multiple genomes, for example due to the presence of a host cell factor. Collectively, detailed analysis of RSV transcription is required to decipher the dynamics and regulation, especially during early infection.

3. Translation

Synthesis of RSV proteins requires translation of viral mRNAs by host cell's ribosomes. The RSV mRNAs are very similar to host cell's mRNAs, containing a cap at the 5' end and a polyA tail at the 3' end. Consequently, RSV mRNAs can be translated according to the canonical, cap-dependent translation model (see section 'Cellular gene expression'). Notably, some RSV mRNAs contain very short 5'UTRs (7 out of 10 genes have a 5'UTR <20 nt), which may affect translation initiation efficiency according to the canonical translation initiation model (Kozak, 1991).

Each RSV mRNA encodes a single viral protein, except for M2 mRNAs that contain two partially overlapping ORFs encoding M2-1 and M2-2. Translation of both ORFs on the polycistronic mRNAs depends on a rare translation mechanism: coupled translation-termination reinitiation which requires a combination of an RNA structure in the first ORF and termination on the first ORF (Ahmadian et al., 2000; Gould and Easton, 2005). After translation termination of the first ORF, the ribosome scans backwards until it reaches the RNA structure that stabilizes the scanning ribosome on the start codon of the second ORF, leading to translation initiation of the second ORF (Gould et al., 2014; Gunišová et al., 2018). Thus, RSV protein production depends on cellular translation machinery and RSV mRNAs can use the host cell's ribosomes to synthesize proteins in canonical and non-canonical manners.

4. Inhibiting host cell's antiviral response

To prevent antiviral interference during an infection, RSV produces several proteins, including non-structural proteins 1 and 2 (NS1 and NS2). These proteins potently prevent numerous aspects of antiviral responses, including sensing of viral RNAs and avoiding apoptosis (Sedeyn et al., 2019). Additional RSV proteins, like SH and G, also

contribute to host cell inhibition by circumventing TNFα signaling and dampening chemokine production required for T-helper cells (Anderson et al., 2021; Fuentes et al., 2007). Moreover, G proteins are synthesized as transmembrane or soluble proteins and the soluble isoforms affect the adaptive immune response to an RSV infection (Anderson et al., 2021). Furthermore, matrix protein (M) can also contribute to evasion of an antiviral response. During RSV infection, M can be imported into the nucleus and then M can interfere with host cell's gene expression (Ghildyal et al., 2003, 2006). Notably, the subcellular location of M changes over the course of an infection. Especially early in infection, M is enriched in the nucleus, whereas M accumulates on the cytoplasm during budding, implying that manipulation of the host cell's response to an RSV infection can change during an infection. Finally, inclusion body formation during an RSV infection may also facilitate escaping host cell detection. Accumulation of viral components in inclusion bodies results in high concentrations of transcription and replication machinery locally, and it may also serve to shield RSV from antiviral sensors.

The RSV life cycle consists of many steps and mistakes during the production of new virions can result in the formation of defective particles (DV), which may lack essential RSV genes (Sun and López, 2016; Vignuzzi and López, 2019). Consequently, DVs are incapable of establishing a successful infection independently. However, if the same cell is also infected by a full-length RSV or by other DVs defective for different genes, the DVs can participate in the RSV infection. Specifically, DVs can compete with the full-length RSV as they typically replicate faster. As DVs often are deficient for genes responsible for host cell evasion, the presence of DVs in an infection interferes with RSV's ability to evade host cell's antiviral response and infection with DVs leads to potent activation of an antiviral response (Felt et al., 2021; Vignuzzi and López, 2019). Collectively, the various mechanisms employed by RSV to circumvent the host cell response lead to inefficient launching of an antiviral response, which likely contributes to disease progression. If DVs are present or formed during an infection, the antiviral response may be more potent. The dynamics of host cell evasion and DV formation are thus important parameters in the success-rate of an infection.

5-7. Assembly and release

Secretion of new virions occurs through the assembly of particles on the plasma membrane and budding into single particles. The transmembrane surface proteins SH, G, and F are produced and transported to the plasma membrane through the ER-Golgi pathway, whereas all other components of the virion are produced in the cytoplasm and some accumulate in inclusion bodies. The exact mode of translation of an RNP from the cytoplasm or inclusion bodies to the site of budding is unknown, but probably involves cytoskeletal transport (Kallewaard et al., 2005). Assembly of RNP and glycoprotein-containing membranes requires interactions between M and M2-1 with RNP and the transmembrane proteins (Hu et al., 2020). Finally, RVS virions are released from the host cell, although the exact mechanism of budding is unknown (Hu et al., 2020; Shaikh et al., 2013). The released virions are heterogeneously shaped and have different sizes and multiple genomes may be packed into a single large virion (Liljeroos et al., 2013). The nature of the released particles can have implications for the infection of the next cell. If a large virion contains multiple genomes, the onset of infection in the next cell may differ from an infection that was initiated by a virion

with a single genome.

The presence of F on the plasma membranes is required to produce virions that can infect by fusion with the membrane of the next host cell (see step 1). Likewise, the Fs on the membrane of a virion-producing cell can induce fusion with neighboring cells, resulting in syncytia, which is responsible for the name of RSV. Whether syncytia formation is simply a consequence of the F presence on virion producing cells, or whether an RSV infection benefits from syncytia formation is unclear. It is conceivable that bypassing extracellular secretion to enter a next host cell may benefit the virus as it limits detections by the immune system. Alternatively, fusion with unaffected cells may supply more host cell's resources (such as ribosomes, plasma membranes, etc.) that could boost the efficiency of an infection. Examining the dynamics of an infection in a single cell versus a syncytium may reveal how RSV could benefit from syncytia formation.

In summary, infection by an +ssRNA or -ssRNA virus involves cascades of events, started by entry of a virion into a host cell and completed with the release of newly synthesized particles to initiate another infection cycle in the next cell. To accurately convert the viral genome into viral proteins and more viral genomes, the virus needs to balance viral protein production and replication. Especially at the onset of an infection, when the viral machinery required for replication and viral protein production may be scarce, proper coordination between viral protein production and replication may be of utmost importance for establishing a successful infection. Moreover, the regulation of viral gene expression versus replication on individual genomes may depend on the stage in infection. Early in infection, when only a few viral molecules may be present in an infected cell, the dynamics of replication and viral protein synthesis may differ from later in infection, when many viral RNAs and proteins may have been formed in the cell. To further complicate the understanding of the dynamics during an infection, multiple processes may occur simultaneously on a single viral RNA molecule. For example, replication of a viral genome may not take place one molecule at the time. Instead, multiple polymerases may use the same RNA molecule at the same time to produce multiple new genomes. Therefore, to understand the outcome of an infection, sensitive assays are required with the ability to follow an infection over time and to detect multiple single molecules at the same time.

III.   INNATE ANTIVIRAL RESPONSE TO INHIBIT RNA VIRUSES

Whilst a virus is establishing an infection, the host cell attempts to interfere with the production of new virions. The innate antiviral response to inhibit an RNA virus depends on the sensing of the virus and the activation of the antiviral response in the infected cell. This involves, launching an antiviral state in neighboring cells, in which the virus would be readily detected and suppressed if the virus infects the neighboring cell (Fig. 9A). Here, I briefly introduce the key steps of an antiviral response against RNA viruses, such as *Picornavirales* and *Mononegavirales*.

SENSING OF VIRUS INFECTION AND ACTIVATION OF ANTIVIRAL RESPONSE

Detection of an RNA virus infection by the host cell generally depends on the recognition of viral RNA by host cell sensors that recognize 'unusual' RNA states, most importantly dsRNA or RNAs lacking host cell specific post-transcriptional modifications (Hur, 2019; Wu and Chen, 2014) (Fig. 9B).

During replication of an RNA virus's genome, an RNA molecule complementary to the genomic RNA is generated, which then serves as a template for synthesis of new genomic RNAs. Consequently, double-stranded replication intermediates are formed during replication and are abundantly present in infected cells. Moreover, subcellularly-localized viral accumulates, such as replication organelles or inclusion bodies, are formed during replication of many RNA viruses (Dolnik et al., 2021; Paul and Bartenschlager, 2015; Rincheval et al., 2017; van der Schaar et al., 2016; Du Toit, 2020), which facilitates concentration of viral molecules, but may also enable the formation of dsRNA structures between complementary viral RNAs. The presence of dsRNA in a cell is thus an indication that the cell is infected by an RNA virus and serves as an initiator of an antiviral response.

Detection of viral dsRNA in the cytoplasm of an infected cells depends on proteins from the RLR (RIG-I [Retinoic acid-Inducible Gene I] like receptor) family. The two main RLRs, namely RIG-I and melanoma differentiation-associated protein 5 (MDA5), recognize different aspects of viral RNAs (Hur, 2019). Viral RNA recognition by MDA5 depends on long (>1kb) dsRNAs, which accumulate during replication of +ssRNA (Feng et al., 2012; Weber et al., 2006). In contrast, RIG-I recognizes shorter dsRNA molecules with 5′-triphosphate, which are typically removed from cellular mRNAs, but present on viral mRNAs of most -ssRNA viruses (Schlee et al., 2009). In addition to cytoplasmic dsRNA sensors, a host cell can also sense dsRNA in endoplasmic compartments, where viral dsRNA may reside upon endoplasmic entry into a host cell (Alexopoulou et al., 2001). For example, RSV enters a cell through endocytosis and its RNA can be detected by Toll-like receptor 3 (TLR3) (Botos et al., 2009; Groskreutz et al., 2006). Collectively, the viral sensors can detect viral RNA from a broad range of viruses, but are typically not activated by cellular RNAs, as cellular RNA molecules lack long dsRNA structures and are devoid of 5′-triphosphate.

In addition to dsRNA detection, some RNA virus infections can also be sensed by host cells in an dsRNA-independent manner. For example, nucleotide-binding oligomerization domain 2 (NOD2) recognizes ssRNA and can mount an antiviral response during RSV infection, although the exact mechanism of NOD2-based RSV recognition is unclear (Sabbah et al., 2009; Vissers et al., 2012).

If viral RNA in a host cell is detected by one of the sensors, a cascade of events can start, leading to expression of antiviral genes, such as interferons (IFNs) (Fig. 9B) (Sun et al., 2010; Wu and Chen, 2014). In brief, sensing of viral RNA results in activation of the kinases TANK-binding kinase-1 (TBK1) and IκB kinase homologs (IKKs) and these kinases phosphorylate transcription factors IFN regulatory factor 3 and 7 (IRF3/7) (Fitzgerald et al., 2003; Sharma et al., 2003). Upon phosphorylation, IRF3/7 translocates into the nucleus, resulting in transcription of type I IFN genes, including IFN-α, IFN-β, IFN-ε, IFN-κ, and IFN-ω (referred to as IFN). Thus, accurate sensing of a virus infection and signaling of the antiviral pathways results in IFN synthesis and secretion from the infected cell.
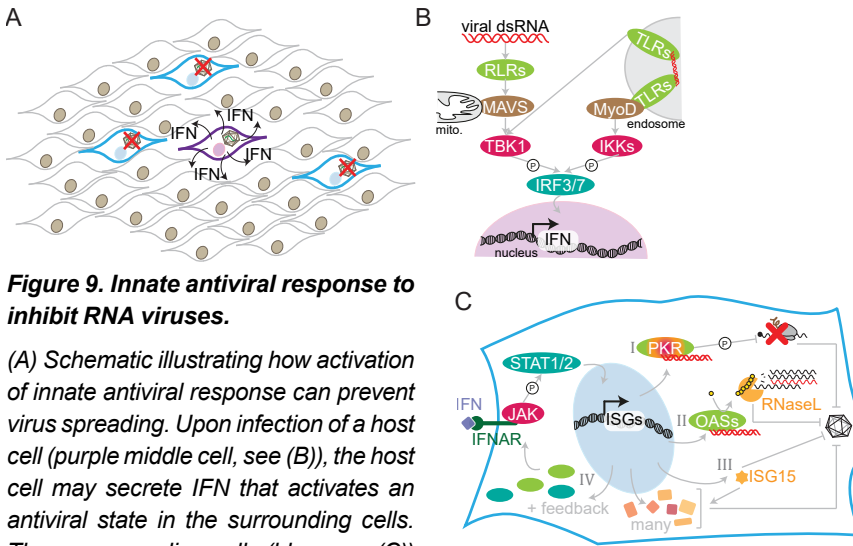
Figure 9. Innate antiviral response to inhibit RNA viruses.

(A) Schematic illustrating how activation of innate antiviral response can prevent virus spreading. Upon infection of a host cell (purple middle cell, see (B)), the host cell may secrete IFN that activates an antiviral state in the surrounding cells. Those surrounding cells (blue, see (C)) are then able to inhibit a virus, if they get infected. (B) Simplified illustration of two of the main pathways responsible for sensing the presence of an RNA virus, leading to the activation of IRF3/7 kinases. Upon phosphorylation, IRF3/7 translocate into the nucleus and induce transcription of IFNs. (C) Simplified illustration of activation of antiviral state by IFN and main antiviral mechanisms to inhibit an RNA virus. Binding of IFN to its receptor causes activation of the JAK/STAT pathway, resulting in phosphorylation and nuclear translation of transcription factors STATs (left). IFN induces expression of many ISGs that can interfere with RNA viruses in various ways. Upon binding to dsRNA, ISG PKR (1) is active and induces global translation inhibition and ISG OAS (2) produces an ATP oligomer (yellow; 2-5A) that activates RNaseL. Upon activation, RNaseL cleaves both host cell and viral RNAs. ISG15 (3) inhibits viral infection by directly interfering with various viral processes and by supporting other ISGs. Finally, many components (4) of its signaling pathways are also ISGs, resulting in a positive feedback. Color legend: light-green: viral sensing protein; pink: kinase that phosphorylates (P) upon activation; blue/green: transcription factor; orange: ISGs that inhibit virus. See main text for further details and abbreviations.

However, there is massive cell-to-cell variation in IFN expression upon an RNA virus infection; the majority of cells fails to induce IFN expression, whereas a fraction of cells potently produces IFN (Chen et al., 2010; Doğanay et al., 2017; Russell et al., 2018; Zawatzky et al., 1985; Zhao et al., 2012). The heterogeneity in IFN induction can been attributed to various factors, including variable expression of viral sensors and signaling pathway components, differences in expression of viral proteins evading sensing, the microenvironment of the infected cell, or a combination of different factors.

To study activation of the antiviral response, cells are often stimulated with synthetic dsRNA (i.e. Polyinosinic-polycytidylic acid [polyIC]). However, a dsRNA response upon polyIC transfection may not accurately reflect how an dsRNA response is induced during an infection. Upon polyIC transfection, the cell is instantly exposed to many dsRNA molecules, whereas the number of dsRNAs during an infection is low during early infection and may accumulate over time. Moreover, viruses express various

proteins to evade detection, and the ability to prevent launching of the antiviral response may therefore change over the course of an infection as the concentration of viral proteins can change. Virus-based evasion of an antiviral response or its temporal changes are typically not included in polyIC-based experiments. PolyIC-based examination of launching an antiviral response may therefore not mimic the dynamics of the antiviral activation during a virus infection. Instead, the activation of the antiviral response should be studied in the context of an infection and these studies should consider the moment of measuring the antiviral response relative to the onset of an infection in each single cell.

ANTIVIRAL RESPONSE AND INHIBITION OF VIRUS

Production and release of IFN from an infected cell causes an activated antiviral state in the infected cell and in surrounding cells that are then equipped to inhibit an infection (Fig. 9A, C) (Hoffmann et al., 2015; Schneider et al., 2014). In brief, binding of IFN to its receptor (IFNAR [IFN-α receptor]) induces the JAK/STAT (Janus kinase signal transducer and activator of transcription) pathway. Upon phosphorylation of STAT by JAK, the transcription factor translocates into the nucleus and induces expression of IFN stimulated genes (ISGs): a plethora of genes that encode various proteins to interfere during viral entry, uncoating, genome transcription or translation, and assembly of new virions (Der et al., 1998; Schneider et al., 2014). Here, I discuss some example ISGs that illustrate how activation of an antiviral state can result in inhibition of an RNA virus and I highlight some open questions related to these ISGs.

Protein kinase R (PKR)

One of the main ISGs, EIF2AK2, encodes the translation initiation inhibitor PKR. The potency of PKR in inhibiting an infection is illustrated by the wide range of viral proteins preventing PKR activation; viral strains deficient in those PKR antagonists or inhibitors are very susceptible to inhibition of infections by PKR (García et al., 2006).

Activation of an PKR requires phosphorylation by another PKR monomer, which can occur if two or more PKR monomers assemble on the same dsRNA molecule (Dey et al., 2005). Multimerization of PKR monomers strongly depends on the concentration of dsRNA molecules with > 30bp (Lemaire et al., 2008). PKR is therefore particularly potent in cells infected with an RNA virus that form such dsRNA structures. Active PKR phosphorylates eIF2α, which causes inhibition of translation of both viral and cellular RNA molecules (see section 'eIF2 and integrated stress response'). Activation of PKR thereby interferes with an infection by preventing synthesis of viral proteins. Furthermore, PKR activation can cause apoptosis of an infected cell (García et al., 2006), which also contributes to limiting the spread of an infection by clearance of infected cells.

The exact dynamics of PKR activation and PKR-induced inhibition of an infection are unclear. In polyIC-transfected cells, there is a rapid increase in phosphorylated eIF2α levels, suggesting rapid and potent activation of PKR (Burke et al., 2019; Rath et al., 2019). However, global translation inhibition due to PKR activation may not

occur until 5 hours after polyIC transfection, implying that PKR only has a minor role during early stages of an infection (Rath et al., 2019). Moreover, the concentration of dsRNA in a cell influences PKR activation. If the dsRNA concentration is too high, PKR activation is inefficient, since multimerization of PKR monomers is impaired in the presence of many dsRNA molecules (Lemaire et al., 2008). As the number of dsRNA molecules in an infected cell may change over the course of an infection, the ability to activate PKR is likely non-uniform throughout an infection. To understand the contribution of PKR in inhibiting an infection, more insight is required into the dynamics of PKR activation relative to the dynamics of the infection.

Oligoadenylate synthetases (OASs)

Another important group of ISGs are the OAS genes that encode for dsRNA-dependent 2′,5′-oligoadenylate (2-5A) producers. The main OAS isoform is OAS3 (Li et al., 2016), although inhibition of some viruses depends on one of the other OAS isoforms. For example, inhibition of the recently emerged coronavirus SARS-CoV2 relies on OAS1 (Wickenhagen et al., 2021). Like PKR, activation of OASs relies on oligomerization on dsRNA molecules and is therefore tailored towards infected cells (Li et al., 2016). 2-5A facilitates dimerization and activation of endonuclease RNaseL, which leads to cleavage of ssRNA molecules at an UN^N sites (Han et al., 2014). Consequently, activation of RNAseL causes degradation of any type of RNA species in a cell, including tRNAs, rRNAs, mRNAs, and vRNAs (Chakrabarti et al., 2011). OAS-dependent RNAseL activation results in inhibition of an infection by clearing vRNAs and by preventing viral protein synthesis (Donovan et al., 2017).

The mechanism of this RNAseL-mediated translation inhibition is not fully understood, as RNAseL-mediated translation inhibition precedes rRNA, tRNA, and mRNA cleavage (Donovan et al., 2017). Remarkably, ribosomes remain able to translate, even after RNAseL-mediated rRNA cleavage (Rath et al., 2019). Moreover, some mRNAs are partially resistant to RNAse, further complicating the understanding of RNAseL-mediated repression of protein synthesis. In particular, IFN and ISG mRNAs can escape or outcompete RNAseL. These transcripts can therefore still be translated despite potent cleavage of RNAs by RNAseL (Chitrakar et al., 2019; Rath et al., 2019). Thus, to understand how OAS-mediated activation of RNAseL contributes to inhibition of an infection, more insight into the dynamics and specificity of RNAseL is needed.

ISG15

ISG15 is one of the most strongly induced genes upon exposure to IFN and encodes a ubiquitin-like protein that can be conjugated to target proteins by ISGylation (Der et al., 1998; Perng and Lenschow, 2018). Many host cell proteins as well as viral proteins can be ISGylated and the result of ISGylation differs per target protein (Perng and Lenschow, 2018). For example, ISGylation of PKR stimulates PKR and leads to potent translation inhibition (Okumura et al., 2013). In contrast, ISGylation of viral proteins, such as capsid proteins, impairs the formation of new virions (Durfee et al., 2010). ISGylation has also been suggested to reduce RSV, but the exact mode of RSV repression by ISG15 is unknown (González-Sanz et al., 2016). The exact contribution of ISGylation to an antiviral response during different moments of

an infection is hard to define, as there is only limited understanding of ISG15 and ISGylation dynamics. For example, the specificity of ISGylation (i.e. identification of ISG15 target proteins) is poorly defined. In fact, analysis of ISGylation during protein synthesis indicated that ISGylation occurs aspecifically during translation: any newly synthesized protein can get ISGylated (Durfee et al., 2010). Based on this model of co-translational ISGylation, ISG15 can only suppress an infection at very high levels of ISG15. If ISG15 levels are low, ISGylation to suppress an infection is unlikely, as there would not be sufficient ISG15 to ISGylate all viral proteins. Similarly, co-translational ISGylation to inhibit viral proteins would be most efficient if a large fraction of ribosomes produces viral proteins. Therefore, the contribution of ISG15 to inhibit a virus during early infection may be minor. More experiments are required to access the role of ISG15 during different moments of an infection.

Positive feedback

Among the many ISGs, there are some genes that encode for proteins involved in sensing an infection and launching an antiviral response, such as RIG-I and STAT1 (Der et al., 1998). As a result, there is positive feedback in cells exposed to IFN to stimulate an antiviral response.

COMPETITION BETWEEN VIRUS AND HOST

As described above, viruses balance the synthesis of viral proteins and genome replication to produce new virions (see 'Viral gene expression & replication'). To complicate the challenges of a virus even more, host cells have sophisticated mechanisms in place to detect and neutralize foreign structures to prevent viral reproduction (see 'Innate antiviral response to inhibit RNA viruses'). In turn, viruses try to evade detection and prevent activation of cellular antiviral strategies (Hoffmann et al., 2015). For example, CVB3's proteases 2A and 3C potently degrade many host cell proteins (Saeed et al., 2020), interfering with the host cell's ability to launch an antiviral response. Likewise, RSV's NS1 and NS2 suppress an antiviral response by antagonizing IFN production and signaling (Sedeyn et al., 2019). Therefore, to multiply efficiently, a virus needs to hijack gene expression components of the host cell, ensure replication of its genome, and circumvent an antiviral response. The outcome of an infection thus depends on a competition between the various viral processes and the host cell: if the virus manages to generate new virions, the virus wins, whereas the host cell is the winner, if the infection is successfully suppressed (Fig. 10A).

A virus infection typically starts with a single virion entering one cell and releasing one copy of the viral genome. As the infection progresses, more and more viral proteins and RNAs accumulate in the infected cell. The number of potential triggers for launching an antiviral response thus increases during an infection, however, the number of viral molecules that need be neutralized by the antiviral response increases as well. The competition between viral gene expression, viral genome replication, and evading antiviral signaling may therefore change over the course of infection (Fig. 10B). Along the infection timeline, there may be moments that

either the virus or the host cell can tip the virus-host balance in their favor. Antiviral inhibition of the virus by the host cell is probably most successful early in infection, as there are only few viral elements to repress. However, early in infection there are only few triggers for the antiviral response and a potent antiviral response is therefore unlikely. Thus, more knowledge on the onset, outcome, and temporal dynamics of the virus-host competition is required.

Most assays currently used to investigate viruses and antiviral responses are suboptimal to explore the complex host-virus competition from the start of an infection onwards. First, many assays lack sensitivity to detect single molecules and are therefore incapable to zoom-in on the single molecules involved in an infection. Highly sensitive assays are required to study the earliest moments of an infections, when only a single or a few viral molecules may be present. Second, creating a reliable read-out often requires a large sample consisting of multiple cells. Not all cells may be infected, and the exact stage of the infection may differ between each individual cell. A single read-out based on all cells at once is thus not inadequate to study details of the virus-host competition. Third, many assays require lysis or fixation of the sample, impairing temporally resolved analysis of the virus-host dynamics. Finally, the virus behavior and the host cell's response are often studied separately, whereas analysis of the virus and antiviral response simultaneously is required to understand the competition. To overcome these limitations, live-cell single-molecule imaging of RNA virus infection is needed to understand the onset and outcome of the virus-host competition, which may expose vulnerabilities, either in the viral life cycle or in the antiviral response, that may serve as targets for antiviral therapies.

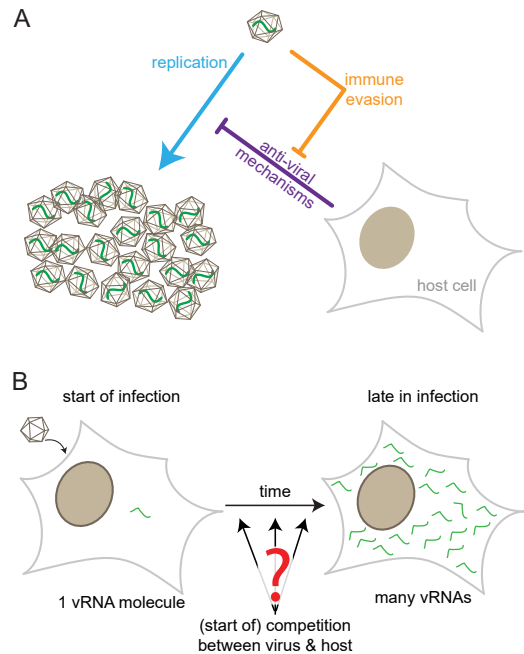**Figure 10. Virus-host competition.**

(A) The outcome of a virus infection depends on the virus replication, the host-cell dependent virus inhibition, and the virus's ability to evade or repress the host cell's response. (B) The virus-host competition may change over the course of an infection and high-resolution, temporally resolved assays are required to interrogate the virus-host competition.

OUTLINE OF THESIS

This thesis describes various projects focusing on the development of live-cell single-molecule imaging assays to grasp dynamic processes.

Chapter 2 and 3 focus on heterogeneity of translation on single mRNAs over time. Inspired by the SunTag translation imaging system, an orthogonal system called MoonTag is developed in **Chapter 2**. The SunTag and MoonTag system can be combined to simultaneously analyze translation of multiple mRNA species in the same cell and to facilitate investigations into translation of multiple ORFs on the same mRNA. Additionally, yet another tag is developed - the MashTag - to interrogate canonical and non-canonical translation simultaneously on single mRNAs. Using the expanded single-molecule translation imaging toolbox, we observe widespread non-canonical translation due to alternative translation start site selection. Canonical and non-canonical translation start site selection is heterogenous between mRNA molecules of the same gene and can even be variable over time on the same mRNA molecule.

On single mRNAs, we frequently observed fluctuations in translation over time. To examine the contribution of ribosome recruitment to variations in translation, we analyze the dynamics of eIFs in **Chapter 3**. We uncover that binding of a single eIF4E molecule to an mRNA molecule can be responsible for translation initiation of multiple consecutive ribosomes. Furthermore, we speculate that the dynamics of eIF4E binding to a cap contribute to the temporal fluctuations in translation of single mRNAs. Collectively, Chapter 2 and 3 indicate that regulation of mRNA translation contributes to the proteome composition.

Chapter 4 and 5 illustrate how single-molecule imaging can be leveraged to investigate virus infection, replication, transcription, translation, and the interplay between a virus and a host cell. In **Chapter 4**, we establish a live-cell single-molecule imaging assay, called VIRIM, to examine translation and replication dynamics of +RNA viruses. VIRIM enables us to trace an infection from the first RNA molecule of the infecting viral particle and measure the timing, efficiency, and effect on the host cell of an infection. Specifically, we use VIRIM to identify the bottleneck of successful infection in the viral life cycle and study how, when, and how well the host cell's antiviral response can try to attack this bottleneck.

Like +RNA viruses, the outcome of a -RNA virus infection depends on viral and host cell dynamics. In **Chapter 5**, we therefore expand the single-molecule virus imaging toolbox to also investigate -RNA viruses, particularly the *Mononegavirales* RSV. A combination of live-cell and fixed cell experiments reveals coordination and heterogeneity in expression of RSV genes. Furthermore, we observe that activation of the antiviral response anticorrelates with viral transcription, indicating that the heterogeneity in viral expression may have profound implications for the outcome of an infection.

Collectively, the new assays developed throughout this thesis have uncovered variations in dynamic processes and serve as a great starting point for many follow-up studies. In **Chapter 6**, I provide a summarizing discussion focusing on the limitations of the current tools and on potential future improvements and applications of the new assays.
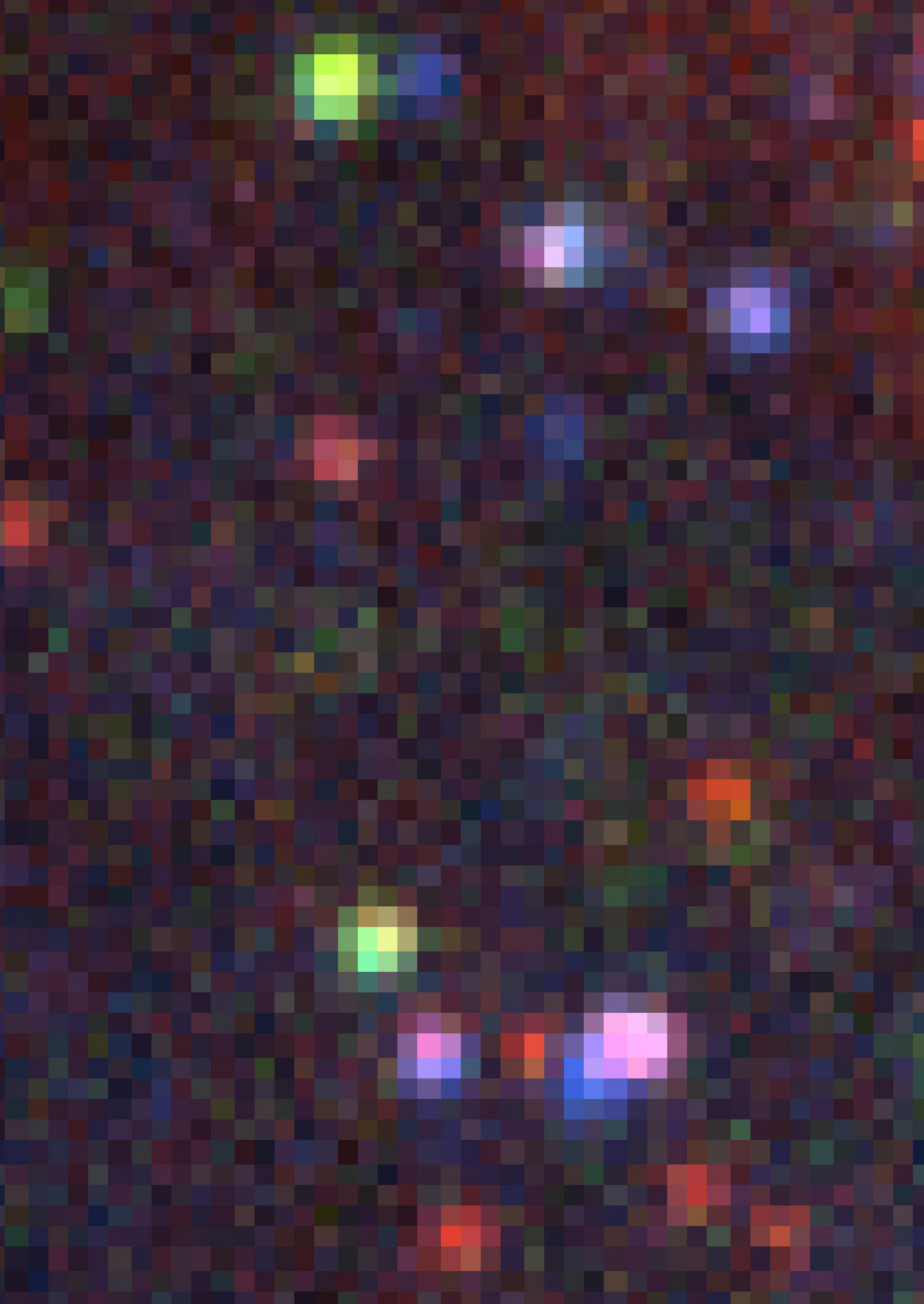
Chapter 2

Multi-Color Single-Molecule Imaging Uncovers Extensive Heterogeneity in mRNA Decoding

Sanne Boersma[1,*], Deepak Khuperkar[1,*], Bram M. P. Verhagen[1],
Stijn Sonneveld[1], Jonathan B. Grimm[2], Luke D. Lavis[2],
and Marvin E. Tanenbaum[1]

[1] Oncode Institute, Hubrecht Institute–KNAW and University Medical Center
Utrecht, Utrecht, The Netherlands.

[2] Janelia Research Campus, Howard Hughes Medical Institute, Ashburn,
Virginia, USA.

[*] Co-first authors.

*Background illustration: example image of uORF-MashTag reporter mRNAs
Related to: Chapter 2 - Fig. 6C*

SUMMARY

mRNA translation is a key step in decoding genetic information. Genetic decoding is surprisingly heterogeneous, as multiple distinct polypeptides can be synthesized from a single mRNA sequence. To study translational heterogeneity, we developed the MoonTag, a fluorescence labeling system to visualize translation of single mRNAs. When combined with the orthogonal SunTag system, the MoonTag enables dual readouts of translation, greatly expanding the possibilities to interrogate complex translational heterogeneity. By placing MoonTag and SunTag sequences in different translation reading frames, each driven by distinct translation start sites, start site selection of individual ribosomes can be visualized in real-time. We find that start site selection is largely stochastic, but that the probability of using a particular start site differs among mRNA molecules, and can be dynamically regulated over time. Together, this study provides key insights into translation start site selection heterogeneity, and provides a powerful toolbox to visualize complex translation dynamics.

HIGHLIGHTS & GRAPHICAL ABSTRACT

- *Development of MoonTag, a fluorescence labeling system to visualize translation*

- *Combining MoonTag and SunTag enables visualization of translational heterogeneity*

- *mRNAs from a single gene vary in initiation frequency at different start sites*

- *Ribosomes take many different "paths" along the 5' UTR of a single mRNA molecule*
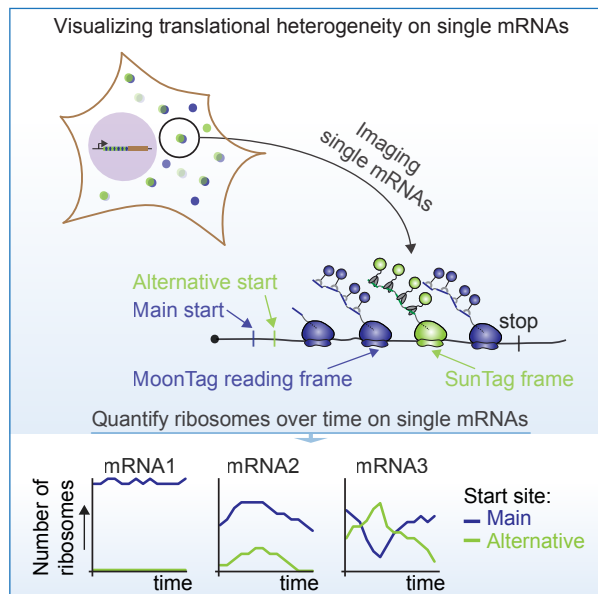
INTRODUCTION

Translation of mRNAs by ribosomes is a key step in decoding the genetic information stored in DNA and mRNA, and regulation of translation plays an important role in shaping the proteome (Hinnebusch et al., 2016; Schwanhäusser et al., 2011). Typically, translation initiates at the most upstream (i.e. most 5') translation start codon, usually an AUG codon, and then continues in the same reading frame until it encounters the first in-frame stop codon (here referred to as canonical translation). However, more recent work has shown that translation of many, if not most mRNAs, is far more complex, and that different regions of an mRNA can be translated. For example, many mRNAs contain multiple open reading frames, including upstream ORFs (uORFs), which are short ORFs upstream of the 'main' ORF that generally repress translation of the main ORF (Calvo et al., 2009; Johnstone et al., 2016). Moreover, ribosomes can translate each nucleotide sequence in 3 different reading frames, resulting in 3 completely unrelated polypeptides (Atkins et al., 2016). Ribosomes translating some eukaryotic or viral RNAs can also undergo frameshifting, changing the reading frame during translation elongation (Dinman, 2012; Dunkle and Dunham, 2015). Finally, ribosomes can bypass stop codons under certain conditions to generate C-terminally extended proteins (Dunn et al., 2013; Schueren and Thoms, 2016). While many examples are known where non-canonical translation occurs productively to generate functional proteome diversity (Barbosa et al., 2013; Dinman, 2012), it is important to note that non-canonical translation may also occur inappropriately, due to errors in translation (Barbosa et al., 2013; Gao et al., 2017). Such errors likely result in synthesis of misfolded and/or dysfunctional polypeptides, which may inhibit the function of the natively folded protein, and can cause proteotoxic stress to the cell.

Critical for determining the translated region of the mRNA, is the selection of the correct translation start site. In eukaryotes, the translation start site is selected during a process in which the 43S translation pre-initiation complex, including the small ribosomal subunit, scans along the mRNA in a 5'-to-3' direction in search of a start codon (Aitken and Lorsch, 2012; Hinnebusch et al., 2016). Identification of the correct start site by a scanning ribosome is complex, as; 1) many genes contain one or more AUG sequences in their 5' untranslated region (UTR) (Iacono et al., 2005); 2) translation can also initiate, albeit generally less efficiently, on near-cognate start codons (e.g. GUG, CUG) (Ingolia et al., 2011; Lee et al., 2012); 3) the canonical start site may not be recognized with 100% efficiency (Kozak, 1986; Lind and Aqvist, 2016); 4) after translating of a short ORF (e.g. a uORF), a ribosome can reinitiate translation at a downstream start site, thus initiating at multiple start sites on a single mRNA molecule (Calvo et al., 2009; Hinnebusch et al., 2016; Johnstone et al., 2016). An additional layer of complexity in selection of a start site is the existence of multiple different transcript isoforms for many genes. For example, alternative transcription start site (TSS) usage or alternative splicing could create different transcript isoforms, and some isoforms may contain translation start sites or translation regulatory elements that are not present in all isoforms (Wang et al., 2016b).

While ensemble measurements have identified multiple translation start sites for many genes, it is currently unclear whether all start sites are used on each individual mRNA molecule, and if so, how their relative usage is regulated. In the simplest model, ribosomes initiate translation on each possible start site with a pre-defined

I

II

III

IV

V

VI

&

probability, which depends on the sequence of the start codon and its local sequence context (i.e. Kozak consensus sequence). In this model, all possible start sites are used on each mRNA molecule and translation start site selection by the scanning pre-initiation complex is purely stochastic. Alternatively, relative start site usage could vary among different mRNA molecules originating from the same gene, for example due to differences in the transcript isoforms, RNA structure or due to regulatory factors, such as RNA binding proteins (RBPs) or RNA modifications. Regulation of start site usage would provide an intriguing possibility to tune protein levels as well as protein sequence in space and time.

While the mechanisms of canonical translation have been extensively studied, the prevalence and underlying causes of heterogeneity in mRNA translation have remained largely unexplored. Currently used technologies, like ribosome profiling and fluorescence reporters, are not ideally suited to detect variability in mRNA translation, including variability in translation start site selection, because 1) they cannot distinguish which translation start sites are used on which mRNA molecules, or whether multiple start sites are used on individual mRNA molecules; 2) they cannot track translation start site usage in space and time for individual mRNA molecules; 3) it is challenging to detect infrequently used start sites above the experimental noise, especially if many different infrequently used start sites exist in an mRNA; 4) static measurements may not readily detect start sites that trigger mRNA degradation. Start sites that result in out-of-frame translation, which likely represent the majority of non-canonical translation initiation events, may trigger nonsense mediated-mRNA decay (Lykke-Andersen and Jensen, 2015), resulting in rapid decay of the mRNAs that preferentially use such alternative start sites. Therefore, new tools are required to uncover the dynamics and heterogeneity in translation start site selection.

RESULTS

Development of the MoonTag, a fluorescence labeling system to visualize translation of single mRNAs

We have recently developed a fluorescence labeling strategy, called SunTag, consisting of a genetically-encoded fluorescently-labeled intracellular antibody and a peptide epitope (Tanenbaum et al., 2014). We and others have shown that the SunTag system (Wang et al., 2016a; Wu et al., 2016; Yan et al., 2016), or a labeling system with a purified antibody (Morisaki et al., 2016), can be applied to fluorescently label nascent polypeptides, enabling visualization of translation of individual mRNA molecules over time. However, the SunTag system only provides a single read-out of translation and is therefore not suited to study more complex translation events. To obtain multiple readouts of translation of single mRNA molecules in real-time, we aimed to establish a second, orthogonal, genetically-encoded antibody-epitope pair for nascent chain labeling (Fig. 1A). An extensive literature search identified seven single chain antibodies (e.g. nanobodies) that bound a linear epitope with high affinity *in vitro* (See STAR Methods). We found that one of these antibody-peptide pairs retained robust binding in cells (gp41; Fig. 1A-D). The gp41 peptide is a 15 amino acid peptide from the HIV envelope protein complex subunit gp41. The gp41 antibody is a 123 amino acid Llama VHH nanobody (clone 2H10), which binds the

peptide *in vitro* with an affinity of ~30nM (Lutje Hulsik et al., 2013). As this antibody-peptide system is orthogonal to our 'SunTag' system, we refer to it as the 'MoonTag' system.

To determine the binding stoichiometry of the MoonTag nanobody to its peptide array, we created peptide arrays containing 4, 12 or 24 MoonTag peptides, which were fused to Mito-mCherry. The binding stoichiometry of the MoonTag nanobody and peptide array was then determined by quantitatively comparing mCherry and GFP fluorescence near mitochondria (See STAR Methods), which revealed that up to ~10-12 MoonTag nanobodies bind to an array of 24 MoonTag peptides (Fig. 1D), slightly less than was observed for the SunTag (Tanenbaum et al., 2014). Similar labeling efficiency was observed when MoonTag peptides were separated by shorter (5 amino acid) linkers, which were used for all subsequent experiments. Fusion of MoonTag peptides to either a histone or a membrane protein resulted in recruitment of the MoonTag nanobody to DNA and the plasma membrane, respectively (Fig. S1A-B), indicating that MoonTag-fused proteins localize correctly to different cellular compartments. The MoonTag nanobody could also be labeled with the far-red dye, JF646 using the HaloTag (Grimm et al., 2015) (Fig. S1C), providing the possibility to label the SunTag and MoonTag in different colors and combine both systems in a single cell.

Next, we introduced a sequence encoding the MoonTag peptide array in our previously developed translation imaging reporter (Fig. 1E) (Yan et al., 2016). In brief, the MoonTag is inserted upstream of a gene of interest (the kinesin Kif18b). During translation, the MoonTag peptides are synthesized before the protein of interest and rapidly bound by the MoonTag nanobody co-translationally. This results in bright fluorescence labeling of the nascent polypeptide, providing a direct readout of translation of single mRNA molecules. Additionally, the reporter mRNA contains 24 binding sites for the PP7 coat protein (PCP) (Chao et al., 2008) in the 3'UTR. Co-expression of PCP-2xmCherry enables fluorescence labeling of the mRNA independently of translation. The PP7 system was also used to tether the mRNAs to the plasma membrane, which substantially increases signal-to-noise during imaging and facilitates long-term tracking of individual mRNAs that undergo cytoplasmic translation, without detectably altering translation dynamics (Fig. 1E) (Yan et al., 2016). When transfected into human U2OS cells stably expressing MoonTag-Nb-GFP and PCP-2xmCherry-CAAX, MoonTag foci could be observed that co-localized with single mRNAs (Fig. 1E, F and Movie S1), indicating active translation of those mRNA molecules. These results demonstrate that the MoonTag system can be applied to label nascent polypeptides and visualize translation of single mRNA molecules in real-time, similar to the SunTag system.

For simultaneous analysis of two different types of mRNAs in single cells, we generated SunTag and MoonTag translation reporters containing different genes (Kif18b and GAPDH) and co-expressed these reporter mRNAs in U2OS cells stably expressing SunTag-scFv-GFP, MoonTag-Nb-Halo[JF646] and PCP-2xmCherry-CAAX (referred to as Moon/Sun cells). Red mRNA foci were observed that co-localized with either SunTag or MoonTag signal, but no mRNAs were observed that contained both signals (Fig. 1G and Movie S2, n = 253 mRNAs; 2 repeats), demonstrating that the SunTag and MoonTag systems are fully orthogonal and can be used together in the same cell to visualize translation of two different mRNAs.
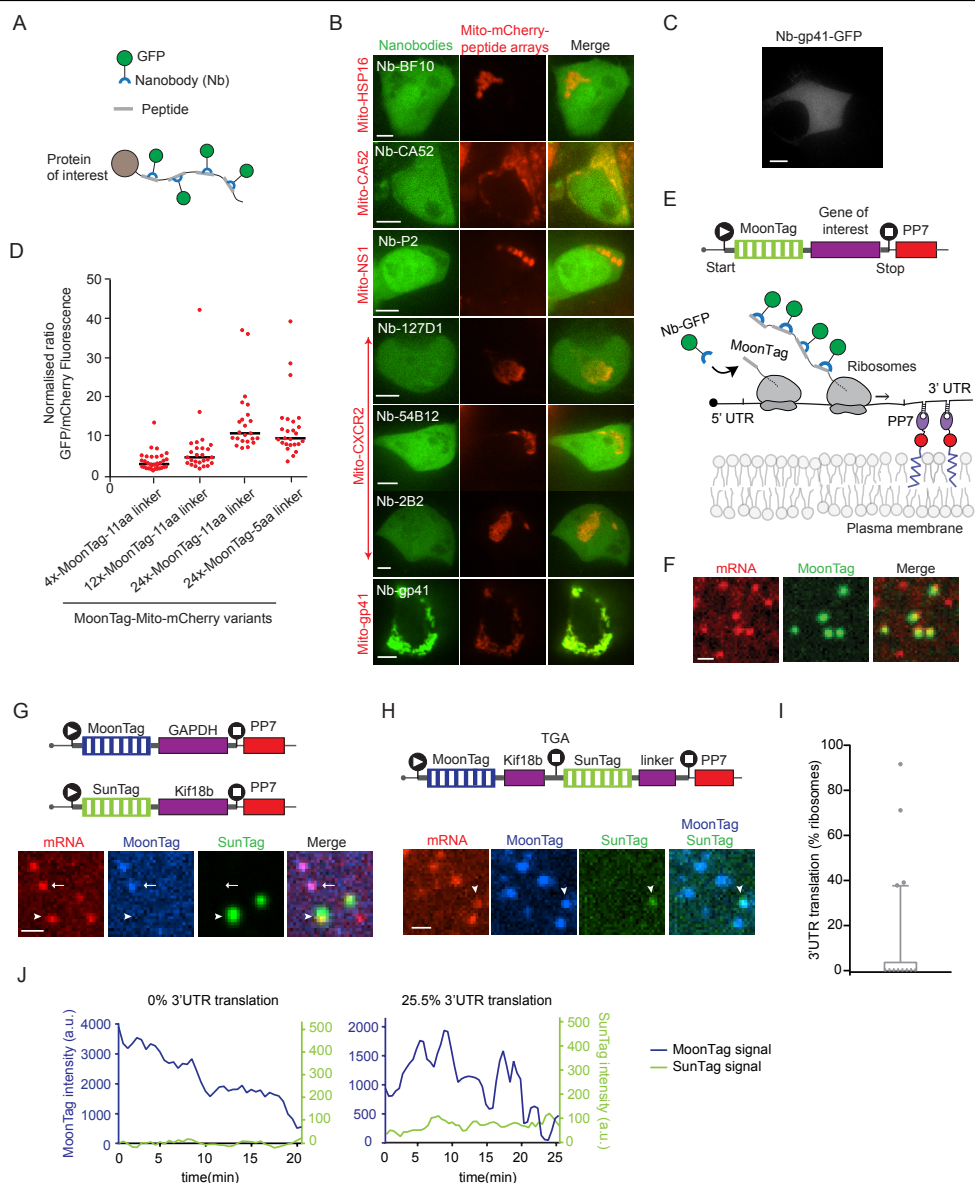
Figure 1. Development of the MoonTag, a fluorescence labeling system to visualize translation of single mRNAs

A) Schematic representation of the nanobody-peptide labeling system. B, C) Representative images of HEK293T cells transfected with indicated constructs. D) Indicated constructs were transfected in U2OS cells stably expressing GFP-tagged MoonTag nanobody. The GFP:mCherry fluorescence intensity ratio on mitochondria was quantified. Each dot represents a single cell, and lines indicates the average. E) Schematic of translation reporter (top panel) and nascent polypeptide labelling strategy using the MoonTag system (bottom panel). F) Representative image of U2OS cells expressing MoonTag-Nb-GFP and PCP-mCherry-CAAX and the MoonTag translation reporter shown in (E). G-H). Schematic of reporters (top) and representative images of Moon/Sun cells expressing indicated reporters (below). G) Arrow head and arrow indicate SunTag and MoonTag

Figure 1 continued:

translation, respectively. H) Arrow head indicates mRNA with 3'UTR translation. I-J) Moon/Sun cells were transfected with the reporter indicated in (H) and MoonTag and SunTag intensities on single mRNAs were tracked over time. I) Boxplot indicates frequency of 3'UTR translation (percentage of ribosomes) calculated for each mRNA. Dashed line represents median value, box indicates 25-75% range, and whiskers indicated 5-95% range. J) Example dual-color intensity traces of two mRNAs with MoonTag (blue) and SunTag signal (green). Number of experimental repeats and mRNAs analyzed per experiment are listed in Table S1. See also Movie S1-3. Scale bars, 5µm (B), 10µm (C) or 1µm (F-H).

To test whether the SunTag and MoonTag systems could also be combined in a single mRNA, we generated a translation reporter with the MoonTag encoded in the ORF, followed by a stop codon and the SunTag (the SunTag sequence was placed in frame with the stop codon) to visualize translation of the 3'UTR (Fig. 1H, schematic). Most mRNAs showed strong MoonTag translation signal, but a small subset of MoonTag-positive mRNAs showed SunTag signal as well (Fig. 1H). The SunTag signal intensity was generally low, and most individual SunTag translation events only lasted a few minutes (Fig. 1H and Movie S3). 3'UTR translation was likely caused by occasional stop codon read-through by individual ribosomes. While translation reinitiation after termination at the MoonTag ORF stop codon is also possible, it is less likely because no AUG start codons were present in the SunTag reading frame downstream of the stop codon. Surprisingly, large variations in the frequencies of 3'UTR translation were observed between different mRNA molecules (Fig. 1I, J). The majority of mRNAs (68.9 %) did not show any 3'UTR translation over the time period of imaging (mean track length 16.9 ± 5.2 min (mean ± SD)), while other mRNAs showed continuous 3'UTR translation, indicative of translation by multiple ribosomes (Fig. 1I, J). The differences in the frequency of 3'UTR translation between different mRNAs were not caused by corresponding differences in the translation initiation rate of those mRNAs (Fig. S1D), suggesting that different mRNA molecules may have a distinct susceptibility for stop codon read-through even though these mRNAs were derived from the same gene. Thus, the SunTag and MoonTag systems can be combined in single cells, and even in single mRNAs to visualize complex aspects of mRNA translation with single ribosome sensitivity.

Development of a translation reading frame reporter that reports on translation start site selection

Alternative translation start site selection is an important form of translational heterogeneity, as the majority of mRNAs contain multiple translation start sites, and translation start site selection can determine both protein sequence and expression levels. Since the translation start site determines the reading frame of a ribosome, we reasoned that a reporter of translation reading frame could be leveraged to report on translation start site selection. To develop a translation reading frame reporter, we designed a tag in which MoonTag and SunTag peptides were "mashed" together: they were fused in alternating fashion and positioned in different reading frames. All SunTag peptides were located in the -1 reading frame with respect to the MoonTag peptides (Fig. 2A). The +1 frame does not contain any SunTag or MoonTag sequences, and is referred to as the 'blank' frame. We named this ribosome reading

frame reporter the M̲oon̲Tag a̲nd S̲un̲Tag h̲ybrid (Mash)Tag (Fig. 2A). To enable the MashTag to report on translation start site selection, we designed two versions of the MashTag reporter: both versions contained 36 copies of the MashTag (devoid of stop codons in all frames), a downstream gene of interest, followed by stop codons in all three frames. As gene of interest, we designed a BFP sequence lacking stop codons in all frames to ensure that the coding sequence length of the MoonTag and the SunTag frame is equal. Finally, 24 PCP binding sites were introduced in the 3'UTR of the MashTag reporter to visualize and tether mRNAs. One version of the MashTag reporter contained an AUG translation start codon in frame with the MoonTag peptides ('MoonStart' reporter), while the other contained an AUG in frame with the SunTag peptides ('SunStart' reporter) (Fig. 2B, C, schematics). Both AUGs were placed in strong initiation sequence context (Kozak consensus sequence) and no other AUG codons were present in the 5'UTR or MashTag sequence. During initial attempts to image cells expressing the MashTag reporters, we noticed that at high expression level, the 'mature' (i.e. ribosome released) protein encoded in the SunTag frame of the MashTag tended to form protein aggregates. The observed protein aggregation was likely caused by an aggregation prone amino acid sequence that is produced when the MoonTag peptides are translated in the -1 frame (i.e. when translating the MashTag in the SunTag frame). Therefore, the MashTag reporter was expressed from an inducible promoter (a CMV-based Tet-On promoter) and induced briefly (15-20 min) before imaging to reduce protein synthesis before the onset of imaging, and all cells that showed protein aggregates were excluded from further analyses.

MoonStart and SunStart reporters showed predominantly MoonTag and SunTag translation signals, respectively (Fig. 2B, C), indicating that they accurately report on the dominant translation start site. Upon addition of the translation inhibitor puromycin, the MoonTag and SunTag fluorescence signals disappeared from the MashTag mRNAs, confirming that the MoonTag and SunTag signals on MashTag mRNAs reflects translation (Fig. 2B, C). Surprisingly, when analyzing the MoonStart reporter, we observed frequent brief pulses of SunTag signal on mRNA molecules that also showed MoonTag signal (Fig. 2D and Movie S3). These pulses of SunTag signal could not be explained by background fluorescence or bleed-through from the MoonTag signal, as similar fluorescence signals were not observed on mRNAs containing only the MoonTag (Fig. S2A). Furthermore, to exclude that dual labeling of mRNAs in both MoonTag and SunTag channels is due to coincidental colocalization of two or more mRNAs, each translating only a single reading frame, we compared the mCherry (i.e. mRNA) fluorescence intensity of mRNAs with only MoonTag signal to mRNAs with both MoonTag and SunTag signal. This analysis revealed that mCherry fluorescence in both categories of mRNAs is similar, arguing against mRNA

Figure 2 continued:

a single ribosome translating a MashTag reporter mRNA. (G-H) Fluorescence intensities of single ribosomes translating the reporter mRNA in the SunTag frame, either when the SunTag is in the main frame (SunStart reporter, G) or is OOF (MoonStart reporter, H). Intensity traces are aligned at the last time point that contains SunTag signal (i.e. just before translation termination). Solid lines indicate experimentally derived values; shaded areas surrounding solid lines indicate SEM. Dashed lines (F-H) indicate expected single ribosome intensity trace of SunTag reading frame. Number of experimental repeats and mRNAs analyzed per experiment are listed in Table S1. See also Movie S4. Scale bars, 1μm.
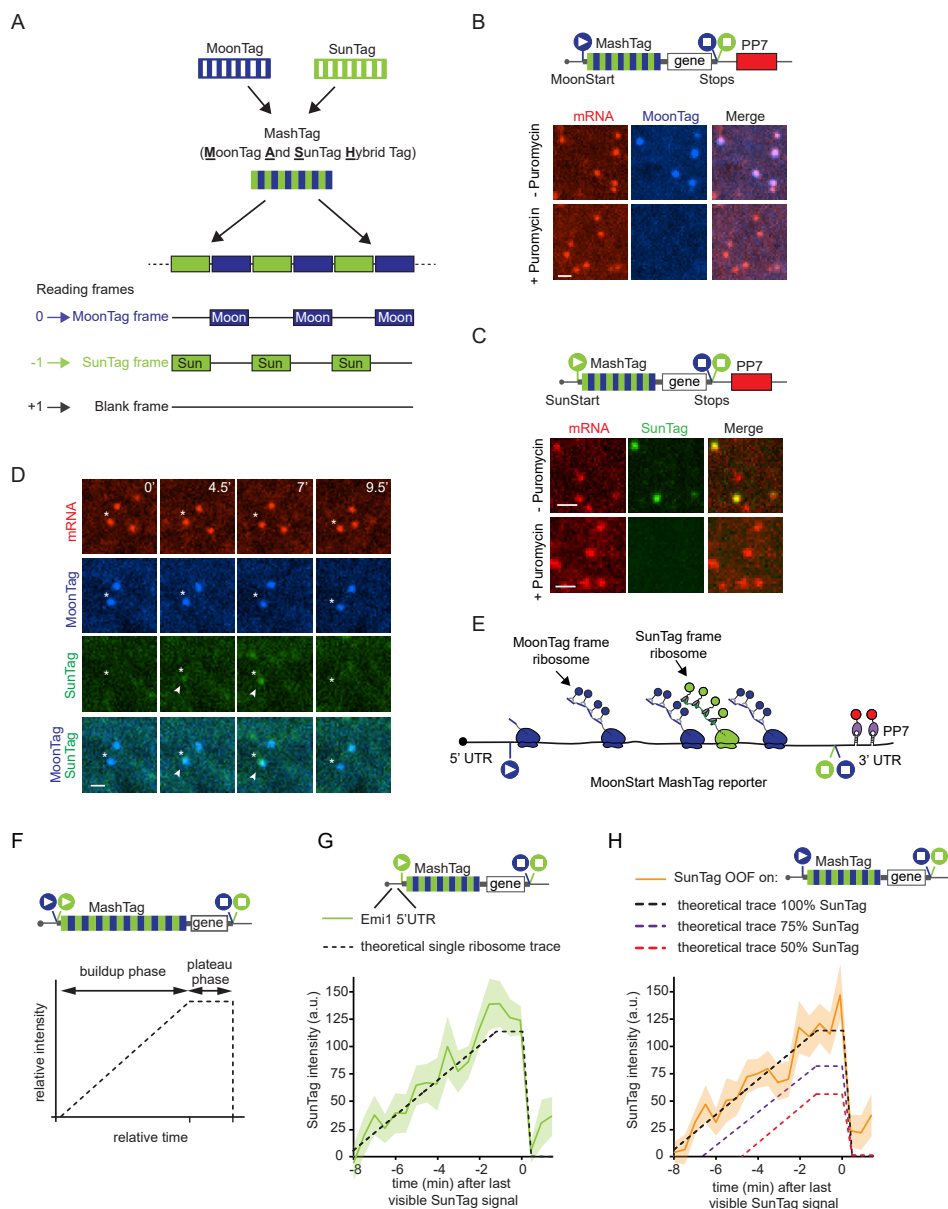
Figure 2. The MashTag: a reading frame sensor to visualize translation start site selection

A) Schematic of the MashTag design. (B, C, F-H) Schematic of MashTag translation reporters (top panels). Circles with triangles and squares represent start and stop sites, respectively. Colors of the filled circles indicate reading frame (blue: MoonTag reading frame; green: SunTag reading frame). For simplicity, 24xPP7 sites in the 3'UTR are not depicted in the schematics. B, C) Representative images of mRNAs in Moon/Sun cells expressing the indicated translation reporters prior to (top images) and after (bottom images) 5 min incubation with the translation inhibitor puromycin. D) Moon/Sun cells expressing the MashTag translation reporter shown in (B). Asterisk in (D) indicates an mRNA with OOF translation. Arrow heads indicate the OOF SunTag signal. Time is indicated in min. E) Schematic of OOF translation on the MashTag reporter. F) Theoretical intensity trace of

multimers as the cause of dual SunTag and MoonTag positivity of a subset of mRNAs (Fig. S2C). Instead, the SunTag pulses on the MoonStart reporter mRNAs represent a subset of ribosomes on the same mRNA that are translating the MoonStart reporter in the SunTag reading frame, which we will refer to as out-of-frame (OOF) translation (Fig. 2E). Together, these results show that the MashTag reporter can accurately report on the dominant translation start site of an mRNA, and can simultaneously reveal non-canonical, OOF translation events on individual mRNA molecules.

OOF translation is mainly due to alternative translation start site selection

OOF translation in the MoonStart reporter could either be due to alternative translation start site selection or ribosome frameshifting. Alternative translation start site selection presumably occurs near the 5'end of the MashTag and thus is expected to include most if not all SunTag peptides. In contrast, if OOF translation is caused by ribosome frameshifting on the MashTag reporter, the OOF translation event would contain only a subset of SunTag peptides, reducing both the SunTag fluorescence intensity and duration of the fluorescence signal of the OOF translation event. To differentiate between these scenarios, we wished to compare the fluorescence of OOF translation events to the expected fluorescence signal of a single ribosome translating the entire array of 36 SunTag peptides (referred to as the 'theoretical single ribosome intensity trace'). The theoretical single ribosome intensity trace contains three distinct phases: 1) a fluorescence intensity buildup phase. During the buildup phase the SunTag peptides are sequentially synthesized and fluorescently labeled by antibodies; 2) a plateau phase when the gene downstream of the MashTag (i.e. the BFP sequence) is translated. During the plateau phase no new SunTag peptides are synthesized, and the fluorescence remains constant; 3) a sudden drop in fluorescence when translation is terminated and the nascent chain is released and diffuses away from the mRNA (Fig. 2F). To determine the duration of the buildup and plateau phases, we calculated the ribosome elongation speed using harringtonine run-off experiments (See STAR Methods), which revealed an elongation speed of 2.9 ± 2.0 codons/s (mean ± SD) (Fig. S2D), similar to our previously determined translation elongation rate in U2OS cells (Yan et al., 2016). Using the nucleotide length of the MashTag and BFP sequences, combined with the experimentally-derived translation elongation rate, the duration of the buildup and plateau phases could be calculated (429 sec and 74 sec, respectively). Next, we determined the fluorescence intensity during the plateau phase. The plateau intensity represents the fluorescence intensity of a single, fully synthesized array of 36 SunTag peptides encoded by the MashTag, and was determined to be 110 ± 53 a.u. (mean ± SD) (Fig. S2E; see STAR Methods).

To validate the values for the theoretical single ribosome intensity trace, we directly determined the fluorescence intensity over time of a single ribosome translating the entire 36 repeats of the MashTag reporter in the SunTag frame. To image single translating ribosomes, we introduced the highly repressive 5'UTR of Emi1 into the SunStart reporter, which reduces translation initiation rates by ~50-fold (Tanenbaum et al., 2015). As a result, mRNA molecules are translated by no more than one ribosome at the time (Yan et al., 2016). Comparison of the theoretical and observed single ribosome intensity traces revealed highly similar traces (Fig. 2G), demonstrating that the theoretical intensity trace accurately represents the fluorescence associated with a single ribosome translating the entire 36 repeats of

the MashTag.

We also generated two additional theoretical intensity traces, which represent translation of either 18 or 27 SunTag peptides by a single ribosome, the approximate average number of SunTag peptides that would be translated if SunTag OOF signals were caused by frameshifting at random positions within the MoonStart mRNA sequence (Fig. 2H; see STAR Methods). We then analyzed the SunTag fluorescence intensity traces of OOF translation events on the MoonStart reporter and compared them to either the trace containing all 36 SunTag peptides, or the traces containing 18 or 27 peptides. This comparison revealed that the intensity profile of single OOF translation events was very similar to the theoretical intensity trace of 36 SunTag peptides (Fig. 2H), indicating that OOF translation is predominantly caused by alternative start site selection near the 5' end of the ORF. Comparison of the SunTag fluorescence intensity of mature polypeptides synthesized from either the SunStart reporter or through OOF translation of the MoonStart reporter also revealed similar intensities (Fig. S2F), confirming that frameshifting is not a major cause of OOF translation signal. Note that OOF fluorescence could, in theory, also be explained by frameshifting that occurs exclusively at a unique sequence near the 5' end of the MashTag. However, this is unlikely, as the nucleotide sequence of the MashTag is quite repetitive, so any frameshifting sequence in one of the first MashTag repeats is likely present multiple times in the MashTag, and thus not unique to the 5' end. Together, these analyses indicate that most of the OOF SunTag translation events are caused by alternative start site selection. Therefore, our MashTag reporter can be used to study translation start site selection kinetics and variability.

A computational pipeline to quantitatively interpret fluorescence signals on translating mRNAs

To understand the heterogeneity and dynamics of translation initiation at both canonical and alternative start sites, it is essential to extract quantitative information from microscopy images on the frequency and timing of both types of initiation events. To facilitate fluorescence intensity measurements, we developed an automated analysis package in Matlab with a graphical user interface (GUI) ("TransTrack", freely available through Github). TransTrack enables simultaneous mRNA tracking and fluorescence intensity measurements in multiple colors, and generates fluorescence intensity traces for both SunTag and MoonTag frames for each mRNA as output.

Next, we wished to convert SunTag and MoonTag fluorescence intensities to the number of ribosomes on the mRNA at each time point. We made use of the theoretical fluorescence intensity profile of a single ribosome in both SunTag and MoonTag frames (Fig. 2F, see STAR Methods). By positioning one or more theoretical single ribosome intensity traces along the time-line of an experimentally observed translation site intensity trace, and summing up their intensity profiles at each time point, the experimentally observed intensity trace of an mRNA translated by multiple ribosomes can be reconstructed *in silico* (Fig. 3A). We developed an iterative stochastic modeling approach to determine the number and temporal position of translation initiation events that generated the best fit with the experimental data (RiboFitter) (Fig. 3B , S3D-F; see STAR Methods). To validate our tracking and RiboFitter, we generated three control reporters: one reporter containing only

I

II

III

IV

V

VI

&
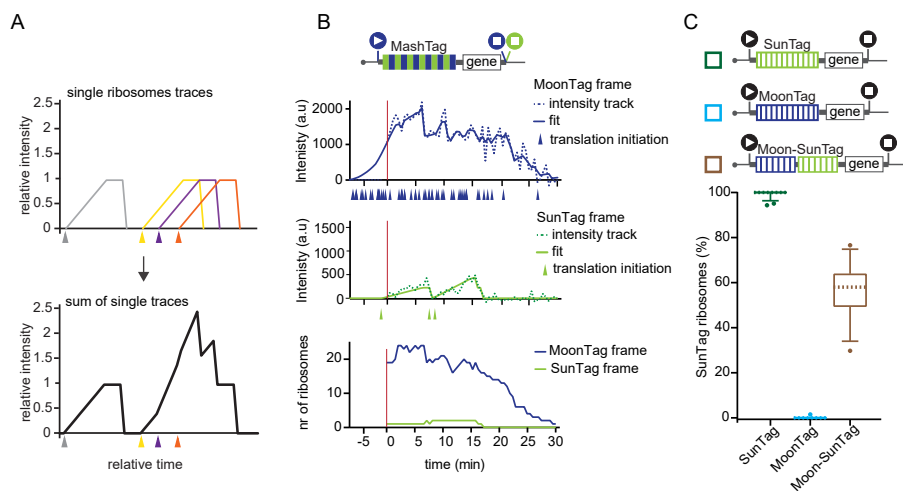
Figure 3. A computational pipeline to quantitatively interpret fluorescence signals

A) Schematic illustrating how the fluorescence originating from multiple ribosomes translating an mRNA together generates a fluorescence intensity trace. Each color represents a single translating ribosome. Triangles indicate the moment of translation initiation. B-C) Schematics of translation reporters (top). For simplicity, 24xPP7 sites in the 3'UTR are not depicted in the schematics. Black start/stop sites in (C) indicate that only a single reading frame contains MoonTag or SunTag peptides. B) An example dual-color intensity trace of a single MoonStart-MashTag mRNA with MoonTag (top panel) and SunTag (middle panel) signal. Dashed lines indicate experimentally observed intensities, solid lines display the optimal fit. Colored triangles below the x-axes of top and middle graphs represent translation initiation events. Bottom panel shows ribosome occupancy per reading frame over time as determined by RiboFitter. C) Boxplots indicate the relative percentage of ribosomes translating the SunTag frame on single mRNAs of the reporter mRNAs indicated above. Dashed line represents median value, box indicates 25-75% range, and whiskers indicated 5-95% range. Number of experimental repeats and mRNAs analyzed per experiment are listed in Table S1.

SunTag peptides, one containing only MoonTag peptides, and one reporter containing both SunTag and MoonTag peptides that were placed in the same reading frame ('Moon-SunTag' reporter). As expected, when SunTag- or MoonTag-only reporters were analyzed, ribosomes were detected almost exclusively in the SunTag and MoonTag frames, respectively (Fig. 3C). Furthermore, the Moon-SunTag reporter showed a narrow distribution in the ratio of SunTag and MoonTag signals, centering close to 50% (Fig. 3C), confirming the accuracy of our analysis pipeline.

Analysis of translation start site selection dynamics and heterogeneity

To determine the frequency of OOF translation on the MoonStart reporter, intensity traces were generated for 85 mRNA molecules that contained detectable translation in either reading frame, and the number of ribosomes translating either reading frame was determined for each mRNA. Traces had a duration of 26 ± 6 min (mean ± SD) and contained 38 ± 30 (mean ± SD) translation initiation events. Most mRNAs were strongly translated in the MoonTag frame; 87% mRNAs had an initiation rate of > 0.5 ribosomes/min in the MoonTag frame. The majority of mRNA molecules (66%,

56 out of 85) showed both SunTag and MoonTag translation events, indicating that multiple translation start sites are used intermittently on most mRNA molecules originating from this reporter gene. Surprisingly, we observed widespread variability in the frequency of OOF translation, ranging from 0% to 100% of the ribosomes (median = 7%) (Fig. 4A, B; see STAR Methods). To rule out that the variability in the OOF translation frequency observed among mRNAs is due to transient transfection of the plasmid encoding the reporter gene, we generated a knock-in of the MoonStart-MashTag reporter in a single genomic locus (the AAVS1 safe harbor locus, a site in the PPP1R12C gene) in Moon/Sun cells. Integration in the correct genomic site was confirmed by Northern Blot (Fig. S4B). MoonStart-MashTag mRNAs expressed from a single genomic locus displayed similar levels of OOF translation (median = 6.7%; p = 0.61, Mann-Whitney test) and variability in OOF translation among mRNAs as mRNAs expressed from transiently transfected plasmids (Fig. S4C).

Two possible explanations could account for the observed variability in OOF translation frequency on different mRNA molecules. First, it is possible that translation start site selection is stochastic, and some mRNAs have more OOF translation than others by chance. In this model, the probability of initiating translation at each potential start site is identical for every mRNA molecule. Alternatively, the probability of alternative start site selection may be distinct for different mRNA molecules. To distinguish between these possibilities, we performed statistical analyses, which revealed that for 25% (21 of 85) of mRNAs start site usage frequency deviated significantly from the population (Fig. 4C, S4D) (p < 0.01, binomial test; see STAR Methods). These results indicate that different mRNA molecules originating from a single gene can be heterogeneous with respect to translation start site usage.

To test whether alternative start site selection frequency depends on the overall translation initiation rate of an mRNA (i.e. the sum of MoonTag and SunTag initiation rates), we compared the frequency of OOF translation with the overall translation efficiency for each mRNA molecule. The OOF translation frequency was similar over a range of translation initiation rates (Fig. 4D), demonstrating that OOF translation does not depend on the overall translation efficiency. Next, we asked whether translation initiation rates in the MoonTag and SunTag frames were correlated over time. We performed linear regression analysis on the intensities of SunTag and MoonTag translation signals for all time points of an mRNA. As a positive control, the level of correlation between SunTag and MoonTag fluorescence over time was determined in the Moon-SunTag reporter, which showed a strong positive correlation, as expected (Fig. 4E). Of note, this linear regression analysis likely underestimates the correlation between MoonTag and SunTag signal, as a strong R^2 value is only expected when substantial changes in the fluorescence intensity occur. In parts of the intensity traces without strong changes intensity fluctuations are dominated by noise, which is not expected to correlate in different fluorescence channels (Fig. S4E; see STAR methods). Analysis of SunTag and MoonTag fluorescence on the MoonStart reporter also revealed a positive correlation between translation in both reading frames for many mRNAs, albeit not as strong as the Moon-SunTag reporter; 56% of MoonStart mRNAs (18/32; note that only 32 of 85 mRNAs could be included in this analysis; see STAR Methods) showed a positive correlation ($R^2 > 0.2$) (Fig. 4E, S4F). The positive correlation between MoonTag and SunTag translation initiation events over time may be explained by temporal fluctuations (i.e. bursting) in the rate of ribosome recruitment to the mRNA, which could affect the initiation rate at all start
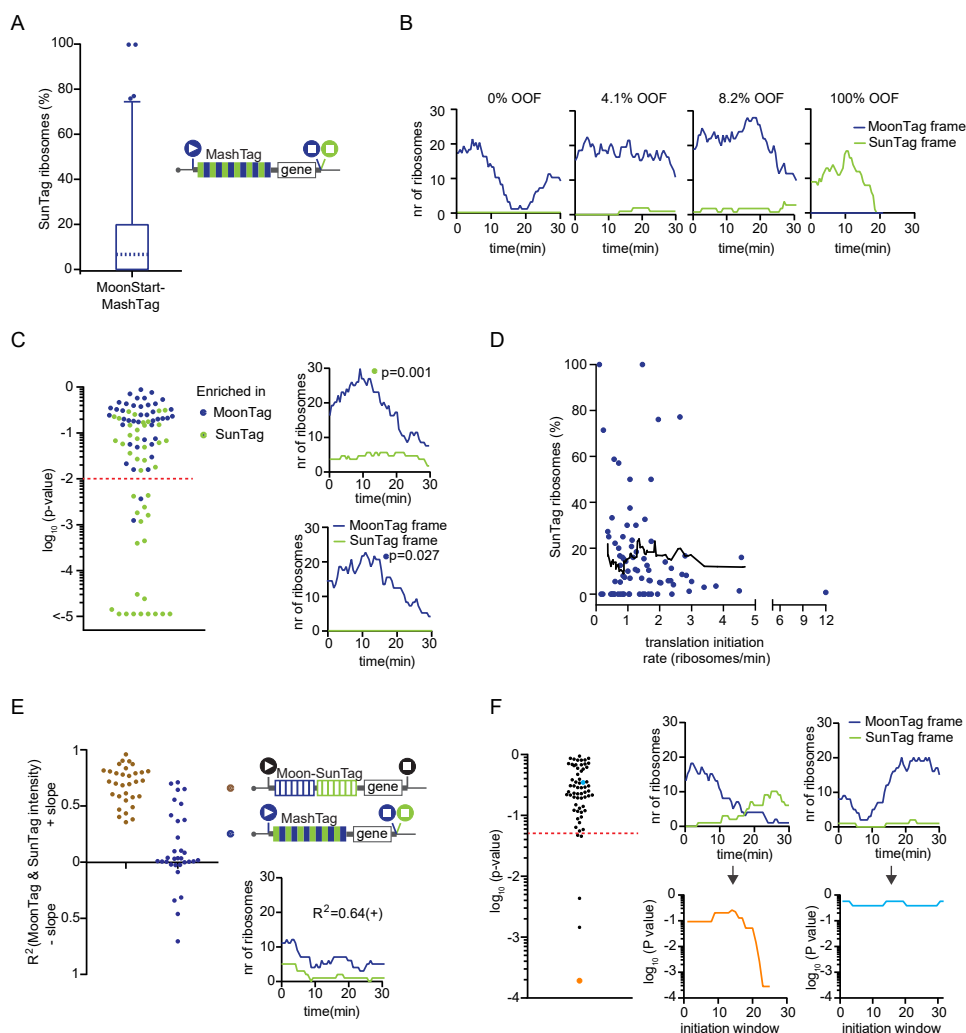
Figure 4. Heterogeneity in translation start site selection among different mRNA molecules

A,F) Indicated reporters were transfected into Moon/Sun cells and MoonTag and SunTag intensities on single mRNAs were tracked over time. For simplicity, 24xPP7 sites in the 3'UTR are not depicted in the schematics. (A) Boxplot indicates the relative percentage of ribosomes translating the SunTag frame on single mRNAs. Dashed line represents median value, box indicates 25-75% range, and whiskers indicated 5-95% range. B) Example graphs of four representative mRNAs in which the number of ribosomes in each reading frame is plotted over time of the reporter indicated in (A). The percentages of SunTag ribosomes on each mRNA is shown (% OOF). C) P-values for enrichment of ribosomes translating either the SunTag or MoonTag frame on individual mRNAs. Every dot represents a single mRNA (left graph). The color of the dot indicates the reading frame that is enriched. Example traces of single mRNAs that show enrichment of ribosomes translating either the SunTag or MoonTag frame (right graphs). D) Correlation between overall translation initiation rate and relative SunTag frame translation frequency for individual mRNAs of the reporter indicated in (A). Every dot represents a single mRNA and the line depicts moving average over 15 mRNAs. E) Linear regression analysis of MoonTag and SunTag intensities for indicated reporter mRNAs

Figure 4 continued:

(left graph). Example trace of one mRNA is shown (right bottom graphs) with indicated R² value. F) Sliding window analysis (see Fig. S4H, I for details) of initiation events in MoonTag and SunTag reading frames on mRNAs of the reporter indicated in (A). Every dot depicts the strongest p-value of a single mRNA (left graph). Example traces show the number of ribosomes in each reading frame over time (top right graphs) and corresponding sliding window p-values (bottom right graphs). Number of experimental repeats and mRNAs analyzed per experiment are listed in Table S1.

sites. Observed changes in fluorescence intensities were not due to imaging noise or fluctuations in nanobody occupancy on the peptide array, as fluorescence intensities remained mostly constant in the presence of the translation inhibitor cycloheximide, which locks ribosomes on the mRNA and prevents translation-dependent changes in fluorescence (SD = ~15% of the mean intensity; Fig. S4G). Moreover, changes in fluorescence due to altered translation occur over multiple consecutive time points (i.e. minutes time-scale), while the observed 'technical' noise acts over milliseconds to seconds.

To investigate whether initiation at different start sites could also be controlled independently, the relative frequency of SunTag frame and MoonTag frame initiation was determined over shorter periods of time to detect 'bursts' in the usage of particular translation start sites. The SunTag and MoonTag translation initiation frequencies were determined in a sliding window of 10 sequential translation initiation events, and the relative initiation frequencies for each window were compared to the average translation frequencies of both frames of the entire trace (Fig S4H, I; see STAR Methods). We then calculated the probability of observing the relative SunTag and MoonTag initiation frequency of each window and determined the lowest window p-value of each mRNA. This sliding window analysis revealed that the majority of temporal fluctuations in the relative frequency of SunTag and MoonTag translation can be explained by chance, indicating that on individual mRNAs start site selection is largely stochastic. However, on a small number of mRNAs (8%; 5/63) a statistically significant change in translation start site selection was observed during the time of observation ($p < 0.05$, binomial test; Fig. 4F, S4J), suggestive of bursts in initiation in a single translation reading frame. While the observed frequency of such bursts in translation start site usage in our dataset was relatively low, our average observation time of individual mRNAs was only 26 min. On a transcriptome-wide level, the fraction of mRNAs that undergoes changes in translation start site usage during their lifetime may be higher.

Alternative translation start site selection can occur on near-cognate start sites both upstream and downstream of the AUG start codon

Since the MoonStart reporter does not contain any AUG start sites in the SunTag frame, SunTag frame translation must initiate on near-cognate start codons, which could be located upstream or downstream of the MoonTag AUG start site (Fig. 5A). Downstream start sites could be encountered by ribosomes after scanning over the MoonTag AUG start site without initiating ('leaky scanning'). To test whether leaky scanning of the MoonTag AUG start codon results in OOF translation on the MoonStart reporter mRNAs, a second AUG start codon was inserted into the mRNA downstream of the MoonTag AUG. Introduction of additional start sites in the MoonTag or blank
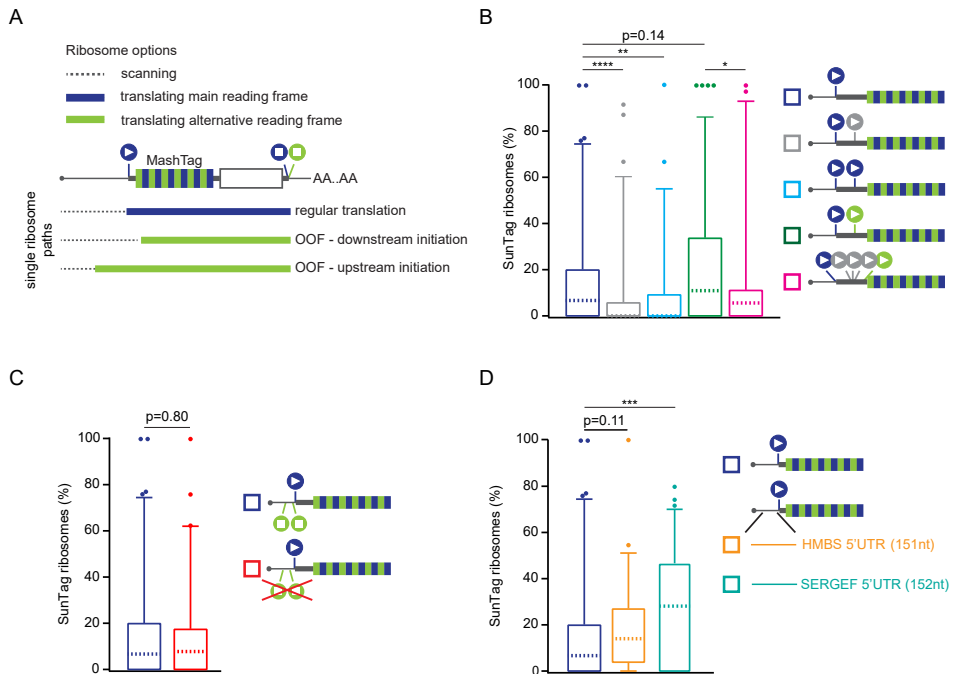
Figure 5. Alternative start site selection contributes to OOF translation

A) Schematic of different possible translation paths of individual ribosomes on a MashTag mRNA. B-D) For simplicity, reporter schematics only indicate 5' region of the mRNA. Indicated reporters were transfected into Moon/Sun cells and MoonTag and SunTag intensities were tracked over time on single mRNAs. Boxplots indicate the relative percentage of ribosomes translating the SunTag frame on single mRNAs. P-values are based on two-tailed Mann-Whitney test; * p < 0.05; ** p < 0.01; *** p < 0.001; **** p < 0.0001. For comparison, data indicated with dark blue (B-D) is re-plotted from Fig. 4A. Dashed line represents median value, box indicates 25-75% range, and whiskers indicated 5-95% range. Number of experimental repeats and mRNAs analyzed per experiment are listed in Table S1.

frame significantly reduced the number of translation initiation events in the SunTag frame (p < 0.01 and p < 0.001, respectively, Mann Whitney test; Fig. 5B). Addition of a start site in the SunTag frame downstream of the MoonTag start site slightly increased the SunTag translation signal, although this effect was not significant (p = 0.14, Mann Whitney test; Fig. 5B). However, introduction of additional start sites in the blank frame between the MoonTag and the newly introduced SunTag start site did significantly decrease initiation in the SunTag frame (p < 0.05, Mann Whitney test; Fig. 5B). Together, these results show that leaky scanning of the MoonTag start site followed by downstream initiation on a near-cognate start codon in the SunTag frame contributes to OOF translation on the MoonStart reporter.

Next, we wished to examine the role of upstream start site selection in OOF translation of the MoonStart reporter. The MoonStart reporter contained two stop codons in the SunTag frame upstream of the MoonTag AUG, which would prevent upstream translation initiation from generation SunTag signal. However, removal of these stop codons (MoonStart ΔSunStops) did not significantly increase the level

of SunTag translation (Fig. 5C), suggesting that upstream initiation in the SunTag reading frame does not strongly contribute to OOF translation on this reporter mRNA.

Rocaglamide A (RocA), an inhibitor of the translation initiation factor eIF4A, was recently shown to stimulate upstream translation initiation (Iwasaki et al., 2016). Treatment of cells expressing the MoonStart ΔSunStops reporter with 0.5 µM RocA resulted in a 37% reduction in overall translation, consistent with inhibition of a key translation initiation factor (Fig. S5A). However, the relative fraction of ribosomes initiating translation in the SunTag frame markedly increased from 8.7% to 21.4 % (median values, p < 0.01, Mann Whitney test; Fig. S5B). These analyses show that upstream translation start site selection can also result in OOF translation, and confirm the previous finding that RocA can stimulate upstream translation initiation.

To access upstream translation initiation within endogenous 5'UTR sequences, two additional MashTag reporters were generated that contained the 5'UTRs of two genes, HMBS and SERGEF (151 and 152 nucleotides in length, respectively). Both 5'UTRs lack start and stop codons in the SunTag frame. While the HMBS 5'UTR reporter showed a similar OOF translation frequency as the MoonStart reporter, introduction of the SERGEF 5'UTR into the reporter resulted in a significant increase in the OOF translation frequency (median, 28.6% vs 7.0%; p < 0.01, Mann Whitney test; Fig. 5D). Interestingly, the overall initiation rate of the SERGEF 5'UTR reporter was also reduced by 35% compared to the MoonStart reporter (p < 0.001, Mann Whitney test; Fig. S5C), indicating that the SERGEF 5'UTR contains translation regulatory elements that result in OOF translation. These results demonstrate that extensive upstream translation initiation occurs on endogenous 5'UTR sequences, suggesting that alternative start site selection might be a widespread phenomenon on endogenous mRNAs.

A real-time sensor to visualize translation of uORF-containing mRNAs

uORFs are present in thousands of mRNAs and generally represses translation of the downstream (main) ORF (Calvo et al., 2009; Johnstone et al., 2016). Ribosomes that translate a uORF can dissociate from the mRNA after translation termination at the uORF stop codon, thus preventing translation of the downstream ORF. Translation of the main ORF can occur either through uORF skipping (i.e. leaky scanning of the uORF start site), or through translation reinitiation at the downstream ORF after translation termination at the uORF stop codon. While a previous study has used the SunTag system to visualize translation of a protein coding ORF downstream of a uORF (Wang et al., 2016a), real-time visualization of schmultiple translation paths (e.g. uORF translation vs. uORF skipping) of a uORF-containing mRNA was not feasible, and therefore the frequency and heterogeneity in path selection by different ribosomes could not be assessed. To determine uORF translation, uORF skipping, and translation reinitiation in real-time on single mRNAs, we generated a single-molecule uORF sensor using the MashTag (Fig. 6A). The uORF sensor is based on the MoonStart reporter and contains an AUG start codon in the MoonTag frame. Upstream of the MoonTag AUG, the reporter contains a short uORF (48 nucleotides; similar to the median human uORF length (Calvo et al., 2009)). The uORF start codon was placed in the blank frame, so initiation at the uORF start site could not result in MoonTag or SunTag fluorescence. A third AUG codon was inserted into the coding
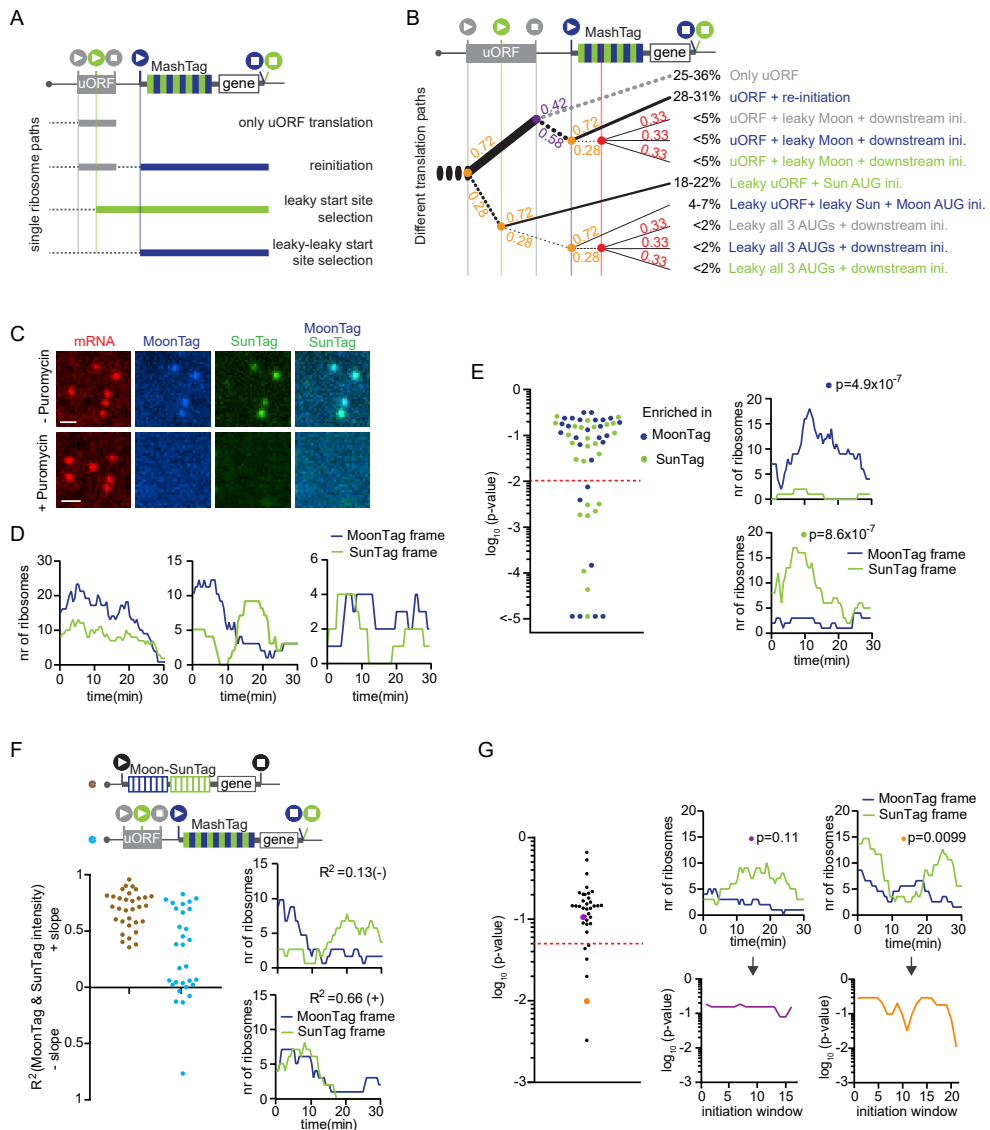
Figure 6. A single molecule uORF sensor based on the MashTag

A, B and F) Schematics of translation reporters (top). For simplicity, 24xPP7 sites in the 3'UTR are not depicted in the schematics. A) Schematic of different possible ORFs that can be translated on a uORF-MashTag mRNA. B) Fraction of ribosomes undergoing each translation path. Thickness of lines reflects the relative usage frequency. Solid lines indicate translation; dashed black lines indicate ribosome scanning; dashed grey line indicates ribosome dissociation from the mRNA. Colored numbers at branch points indicate relative fraction of ribosomes that follow each path. Red vertical line indicates non-canonical start sites in either of the three frames in the MashTag. C-G) MashTag reporters were transfected into Moon/Sun cells and MoonTag and SunTag intensities on single mRNAs were tracked over time. C) Cells were either untreated (top) or treated with puromycin for 5 min (bottom) and representative images are shown. Scale bar, 1μm. (D) Example graphs of single mRNAs of the number of ribosomes translating either SunTag or MoonTag frame over time

Figure 6 continued:

in cells expressing the reporter indicated in (A). E) P-values for enrichment of ribosomes translating either the SunTag or MoonTag frame on individual mRNAs. Every dot represents a single mRNA (left graph). The color of the dot indicates the reading frame that is enriched. Example traces of single mRNAs that show enrichment of ribosomes translating either the SunTag or MoonTag (right graphs). P-values are indicated for example traces. F) Linear regression analysis of MoonTag and SunTag intensities for indicated reporter mRNAs (left graph). For comparison, data indicated in brown is replotted from Fig. 4E. Example graphs of two mRNAs are shown with indicated R^2 values (right graphs). G) Sliding window analysis of initiation events in MoonTag and SunTag reading frames on mRNAs of the reporter indicated in (A). Every dot depicts the strongest p-value of a single mRNA (left graph). Example traces are shown of the number of ribosomes in each reading frame over time (top right graphs) and corresponding sliding window p-values (bottom right graphs). Number of experimental repeats and mRNAs analyzed per experiment are listed in Table S1.

sequence of the uORF, and was placed in frame with the SunTag (Fig. 6A). In this reporter, SunTag signal reports on leaky scanning of the uORF start codon, while a MoonTag signal mainly reflects translation reinitiation after uORF translation. Ribosomes that dissociate from the mRNA after uORF translation are not directly observed, but can be inferred from the decrease in MashTag translation (i.e. MoonTag + SunTag translation) upon introduction of the uORF into the reporter.

Based on the translation rates in both SunTag and MoonTag frames, and based on the overall reduction of translation of the MashTag upon insertion of the uORF, we could estimate the frequency of usage of all translation paths along the uORF reporter (See STAR Methods); 25-36% of ribosomes translate the uORF and do not reinitiate, 28-31% of ribosomes translate the uORF and reinitiate on the downstream MoonTag start site, 18-22% of ribosomes skip the uORF AUG through leaky scanning and initiate on the SunTag AUG, and the remaining ribosomes follow more complex paths (Fig. 6B). We also swapped SunTag and MoonTag start sites, such that SunTag signal reports on translation reinitiation and MoonTag signal reports on uORF AUG leaky scanning, which resulted in similar values for uORF translation and reinitiation (Fig. S6A). To experimentally confirm our calculations on the usage of different translation paths, we removed the uORF stop codon, extending the uORF coding sequence beyond the MoonTag AUG start site (Fig. S6B-D). In this reporter, MoonTag signal can no longer be caused by translation reinitiation. Based on our calculations, we predict that this would result in 80-88% reduction in MoonTag signal, close to the observed 79% reduction in MoonTag signal (Fig. S6B; see STAR Methods). The SunTag translation rate was unaffected (p = 0.75, Mann Whitney test), as predicted (Fig. S6C). This result quantitatively confirms our calculations of the different translation paths, and also confirms that the large majority of MoonTag translation is due to translation reinitiation (Fig. 6B). Together, these results reveal that the MashTag-based uORF sensor can provide a quantitative readout of all possible paths that ribosomes can take along a uORF-containing mRNA.

Next, translation of individual uORF-containing mRNA molecules was examined in more detail. The large majority of mRNA molecules (44/53) contained both SunTag and MoonTag signal (Fig. 6C, D), demonstrating that ribosomes following different paths along the mRNA (e.g. uORF skipping and translation reinitiation) co-exist on most mRNA molecules. However, the relative frequency of the different translation paths varied between different mRNA molecules. A subset of mRNAs (15/53) showed

a significantly greater fraction of translation in either the MoonTag or the SunTag frame than expected based on the total population of mRNAs ($p > 0.01$, binomial test; Fig. 6E), demonstrating that the probability of uORF skipping and translation reinitiation is variable among different mRNA molecules.

When examining the precise moment of translation initiation of ribosomes translating either the SunTag or MoonTag reading frames, a temporal correlation between SunTag and MoonTag translation signals was observed on many mRNAs (Fig. 6F). As discussed before, this correlation is likely caused by temporal fluctuations in the rate of ribosome recruitment to the mRNA. Detailed analysis of translation initiation timing using the sliding window approach (Fig. S4H, I; see STAR Methods) revealed that a subset of mRNA molecules (6/37, $p < 0.05$, binomial test) showed statistically significant bursts of either translation reinitiation or uORF skipping (Fig. 6G), suggesting that uORF translation may be dynamically regulated over time on individual mRNAs. Bursts in translation start site selection did not take place simultaneously on all mRNAs in the same cell, suggesting that the regulation of uORF translation does not occur in a cell-wide manner, but rather at the level of individual mRNA molecules. Together, these results provide the first real-time observations of uORF translation, uORF skipping and translation reinitiation, and offer a quantitative assessment of all the paths that ribosomes take along the 5'UTR of a uORF-containing mRNA and provide a powerful assay to study mechanisms of translation regulation by uORFs.

DISCUSSION

Applications of dual-color single-molecule translation imaging

Expression of SunTag and MoonTag mRNAs in the same cell enables the direct comparison between two different types of mRNAs, for examples between different genes or different mRNA isoforms. Adding a third orthogonal nascent chain labeling system, for example the recently developed 'Frankenbody' (Zhao et al., 2018), would further increase the possible number of mRNA species that can simultaneously be analyzed. The SunTag and MoonTag systems can also be combined in single mRNAs to interrogate complex aspects of translation. In this study, we show that dual color translation imaging can be used to assess translation of the 3'UTR, translation start site selection, and the dynamics of uORFs translation. A parallel study independently developed a multi-color translation reading frame imaging approach to visualize the kinetics of ribosome frameshifting on a viral RNA sequence (Lyon et al., 2018). While the two studies investigate different biological processes, both studies uncover a high degree of translation heterogeneity among different mRNA molecules, and it is possible that widespread translational heterogeneity may be the norm, rather than an exception for most aspects of mRNA translation. The multi-color translation imaging approach will be an important tool to unravel the prevalence, kinetics, and molecular mechanisms of such translational heterogeneity.

Mechanisms of translation start site selection heterogeneity

Using the MoonStart reporter, we found that, overall, ~7% of ribosomes show OOF (i.e. SunTag frame) translation. This value likely represents a lower limit for

endogenous genes, as; 1) our MashTag system only reports on translation of one of the two alternative reading frames; 2) the MoonStart reporter contains a very strong translation start site sequence context, limiting leaky scanning and downstream initiation. In contrast, many endogenous mRNAs have suboptimal start site context (Noderer et al., 2014); 3) The MashTag reporter contains a short, unstructured 5'UTR lacking regulatory elements or additional AUG sequences, limiting upstream start site selection. Endogenous 5'UTRs can be far more complex and therefore could result in a substantially higher upstream initiation rate. Indeed, introducing the endogenous 5'UTR sequence of SERGEF significantly increased OOF translation. Together, these findings suggest that alternative start site selection and OOF translation is likely a widespread phenomenon on many mRNAs.

A subset of mRNA molecules (~25%) showed a significantly altered likelihood of translation initiation on alternative start sites as compared to the bulk of mRNAs in our analysis. There are several possible explanations for the variable frequency of alternative translation start site usage on different mRNAs. First, TSS usage is known to be highly variable in mammalian cells (Forrest et al., 2014) and differences in TSS usage create mRNAs with different 5'UTRs that contain distinct translation start sites. Second, RNA modifications, specific RNA structures or binding of regulatory proteins may alter the probability that translation is initiated on a given start site. Indeed, certain mRNA structures can bias translation initiation site selection in yeast (Guenther et al., 2018). While differences in nucleotide sequence would result in a permanent difference in translation start site usage, RNA modifications, RNA structures and RBP-dependent regulation could be dynamically regulated to alter start site usage over time. For a small number of mRNAs, we indeed observed a change in relative start site usage over time, suggesting that start site selection might indeed be dynamically regulated for single mRNAs. Identifying regulatory mechanisms that shape start site usage is an important future goal, and the MashTag system will be a valuable tool for investigating such mechanisms.

For many mRNA molecules, the average usage of different translation start sites was similar. Nonetheless, the timing of translation initiation and the precise order of initiation events in different reading frames was unique for each mRNA molecule, which likely reflects the inherent stochasticity in start site selection by individual ribosomes. Our results also revealed that the frequency of translation initiation at MoonTag and SunTag start sites was positively correlated over time on many mRNAs. We have previously shown that the translation rate is not constant over time on individual mRNAs, but can show a burst-like behavior (Yan et al., 2016). The fact that multiple translation start sites show correlated bursting, suggests that the burst-like behavior of translation originates upstream of translation start site selection, likely at the step of 43S pre-initiation complex recruitment to the mRNA. In summary, while translation start site selection by individual ribosomes appears mostly stochastic, the probability of usage of individual start sites is under tight control, probably both transcriptionally and post-transcriptionally.

Translation of uORF-containing mRNAs showed many similar characteristics as translation of mRNAs with a single AUG start codon: 1) most mRNAs contained multiple, intermittently-used translation start sites; 2) the selection of a translation start site by individual ribosomes appeared stochastic; 3) usage of different start sites tended to correlate over time; 4) a substantial fraction of mRNA molecules

(~28%) showed a distinct translation start site usage pattern compared to the bulk of mRNAs; 5) evidence for temporal burst in uORF translation, uORF skipping and/or translation reinitiation was obtained. These results suggest that the dynamics and heterogeneity of start site selection are inherent properties of translation and are likely valid for many types of mRNAs.

Consequences of widespread alternative translation start site selection

Pervasive variability in start site selection likely has major implications for cellular function. In-frame alternative start site selection results in N-terminally extended or truncated proteins, which would especially affect the function of proteins containing N-terminal localization signals, like mitochondrial targeting sequences. OOF translation initiation results in polypeptides with a completely different amino acid sequence, which are likely misfolded and non-functional; they would not only waste cellular energy, but could cause a considerable proteotoxic stress to the cell as well. The frequency of OOF translation may be somewhat limited by nonsense-mediated mRNA decay, which may degrade mRNAs that have a high probability of OOF translation (Lykke-Andersen and Jensen, 2015). While widespread alternative translation initiation can be costly, a high degree of flexibility in translation start site selection can also be exploited by the cell. For example, it enables different types of post-transcriptional gene regulation (e.g. uORF-dependent translational control), and it may be important for regulated changes in N-terminal protein sequences as well. An important future question is whether extensive OOF translation is generally functionally important for the cell, or rather reflect errors in translation start site selection.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS
SB, DK, and MET conceived the project. JBG and LDL provided reagents. BMPV and SS developed and optimized the software. SB and DK performed all the experiments. SB, DK, BMPV and MET analyzed the data. SB and DK prepared the figures and MET wrote the manuscript with input from SB and DK.

DECLARATION OF INTEREST

The authors declare no competing interests.

STAR METHODS

RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
DMEM	Gibco	Cat# 31966021
Leibovitz's L15 medium	Gibco	Cat# 21083-027
Opti-MEM	Sigma-Aldrich	Cat# 11058-021
Fetal Bovine Serum (FBS)	Sigma-Aldrich	Cat# F7524
TRIsure	Bioline	Cat# 38033
FuGENE 6 Transfection Reagent	Promega	Cat# E231A
DMSO	Sigma-Aldrich	Cat# D8418-1L
Polyethylenimine	Polysciences Inc	Cat# 23966
Penicillin-Streptomycin	Gibco	Cat# 15140-122
Polybrene	Santa Cruz Biotechnology, Inc	Cat# sc-134220
Doxycycline	Sigma-Aldrich	Cat# D9891-1G
Rocaglamide	Sigma-Aldrich	Cat# SML0656-100UG
Puromycin	ThermoFisher Scientific	Cat# 12122530
Cycloheximide	Sigma-Aldrich	Cat# C4859
Anti-Digoxigenin-AP	Sigma-Aldrich	Cat# 11093274910
Experimental Models: Cell Lines		
Human U2OS cells	Tanenbaum lab	Cat# HTB-96
HEK293T cells	Tanenbaum lab	Cat# CRL-3216
Recombinant DNA		
See Document S1 for all plasmids used in the paper	This study	N/A
Software and Algorithms		
ImageJ	NIH	https://imagej.nih.gov/ij/
Micromanager	Micro-Manager 1.4.22	http://micro-manager.org
NIS-Elements Imaging Software	Nikon	HC 5.11.01
Graphpad Prism 7	GraphPad Software Inc	http://www.graphpad.com/scientific-software/prism/
MatLab R2012b	The Mathworks, Inc.	https://nl.mathworks.com/products/matlab.html
R 3.5.1	R Project for Statistical Computing	http://www.r-project.org/
TransTrack (MatLab)	This Study	https://github.com/TanenbaumLab
RiboFitter (R)	This Study	https://github.com/TanenbaumLab
Other		
96-well glass bottom imaging plates-(Matriplates)	Brooks Life Science Systems	Cat# MGB096-1-2-LG-L
NorthernMax-Gly Kit	ThermoFisher Scientific	Cat# AM1946
DIG RNA Labeling Mix	Sigma-Aldrich	Cat# 11277073910

CONTACT FOR REAGENTS AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by Marvin Tanenbaum (m.tanenbaum@hubrecht.eu). Key plasmids will also be deposited on Addgene.

EXPERIMENTAL MODEL AND SUBJECT DETAILS
Cell lines

Human U2OS cells and HEK293T (ATCC) were grown in DMEM (4.5 g/L glucose, Gibco) containing 5% fetal bovine serum (Sigma-Aldrich) and 1% penicillin/ streptomycin (Gibco). Cells were grown at 37°C and with 5% CO_2.

METHOD DETAILS
Plasmids and reporters

Sequences of all plasmids used in this study are provided in Data S1. The following nanobody sequences were obtained and ordered as G-blocks from IDT:

- Nb-BF10;
- Nb-CA52;
- Nb-2B2;
- Nb-127D1;
- Nb-54B12;
- Nb-P2;
- Nb-gp4.

All peptide array sequences were synthesized by Genewiz. To design the MashTag, the following considerations were taken into account: 1) each repeat of the SunTag or MoonTag in the MashTag had to encode the same SunTag or MoonTag amino acid sequence; 2) no AUG start codons or stop codons (TGA, TAA, or TAG) were introduced in any reading frame; 3) different codons were used for the same amino acid sequence in different copies of the SunTag and MoonTag peptides to introduce nucleotide sequence variation between individual repeats; 4) all sites for restriction enzymes were removed.). Of note, all translation start sites in the MashTag reporters contain a strong Kozak sequence (GCCACC<u>AUG</u>G). After generation of a MashTag containing plasmid, the size of the MashTag was checked by enzyme digestion, and the 5' and 3' ends were sequence verified. Because of difficulties in sequencing due to the repetitive nature of the MashTag, the middle part of the MashTag was not sequence verified for all plasmids.

Lentiviral infection and cell line generation

To produce lentiviruses, HEK293T cells were infected with the lentivirus plasmid of interest and lentiviral packaging vectors ps.Pax and p.MD2 using PEI (Polyethylenimine). One day after transfection, the medium was replaced with fresh medium. Virus-containing medium was collected 3 days after transfection. To infect U2OS cells with lentivirus, cells were seeded 24h before infection and grown to ~60% confluency at moment of infection. The supernatant of the HEK293T cells

containing the lentivirus was added to the U2OS cells. U2OS cells were spin-infected for 90-120 minutes at 2000 rpm at 25°C. After spin-infection, the medium was replaced with fresh medium and cells were cultured for at least 2 days before any further analysis or processing. Where applicable, cells were FACS-sorted as single cells in 96-well plates to generate monoclonal cell lines.

To generate a cell line with stable expression of the MoonStart-MashTag reporter from a single genomic locus, a TALEN-based knock-in of the reporter was made into the AAVS1 locus (in the PPP1R12C gene). The Moon/Sun cells were transfected with two TALEN plasmids (to cut both strands of the DNA) and a plasmid encoding the MoonStart-MashTag reporter driven by a tetracycline inducible promoter, two homology arms to direct homologous repair and a P2A-puro cassette followed by a BGH polyadenylation sequence to select for cells with correct integration. To select for cells with a successful knock-in, cells were subjected to puromycin (2 μg/ml) treatment 4 days after transfection. To check whether the knock-in of the MashTag reporter was successful and had occurred in the correct location, a northern blot was performed (NorthernMax-Gly, ThermoFisher). A probe was designed targeting the BGH polyadenylation sequence. Genomic integration into the correct locus should yield an mRNA with a length of ~1.3 kb; 0.3kb of the endogenous PPP1R12C mRNA fused to 0.9kb of P2A-puro-BGH sequence. On the Northern blot, only a single band was visible at the correct size (1.3kb), indicating that the knock-in site was correct and that no off-target integration had occurred.

Single-molecule translation imaging

For translation imaging experiments, all imaging was done using U2OS cell lines stably expressing TetR (for inducible expression), PP7-2xmCherry-CAAX, either MoonTag-Nb-GFP or MoonTag-Nb-Halo[JF646] and scFv-sfGFP. Cells were seeded in glass bottom 96-wells plates (Matriplates, Brooks) at 15-20% confluency 2 days before imaging. DNA plasmids encoding reporter mRNAs were transfected 1 day prior to imaging using Fugene (Promega) and for MashTag imaging experiments, a BFP-encoding plasmid was co-transfected (DNA ratio 1:1), which was used for initial identification of transfected cells. One hour prior to imaging, medium was replaced with CO_2-independent pre-warmed L15/Leibovitz's (Thermo Fisher) containing 50nM Halo[JF6464]. After Halo incubation for 1h at 37°C, the cells were briefly rinsed twice with L15/Leibovitz's medium and washed once with L15/Leibovitz's for 15 minutes. Doxycycline (1 μg/ml) was added 15-20 minutes before start of imaging to induce transcription of the reporter. To select cells for imaging, approximately 50 positions were first selected based on BFP signal (the co-transfection marker). From this selection, approximately 10 positions were chosen for time-lapse imaging based on the presence of translation sites and the absence of protein aggregates. For time-lapse imaging, images were acquired at 30s interval with 500ms exposure times for 30 minutes, unless otherwise noted. A single Z-plane was imaged, which focused on the bottom plasma membrane of the cells. Images were acquired using a Nikon TI inverted microscope with perfect focus system equipped with a Yokagawa CSU-X1 spinning disc, a 100x 1.49 NA objective and an iXon Ultra 897 EM-CCD camera (Andor) using Micro-Manager Software (Edelstein et al., 2010) or NIS Elements Software (Nikon).

QUANTIFICATION AND STATISTICAL ANALYSIS
Screening of antibody-peptide pairs

For screening of antibody-peptide pairs, seven different nanobodies fused to GFP, were cotransfected with their respective mito-mCherry-peptide arrays in HEK293T cells and analyzed for co-localization at mitochondria using a spinning disc confocal microscope.

Stoichiometry of MoonTag nanobody-peptide

The number of MoonTag nanobodies that could bind to a peptide array was determined as described previously (Tanenbaum et al., 2014).

Tracking single mRNAs using TransTrack

Only mRNAs were selected for analysis that contained translation signal in either one of the two channels (i.e. SunTag or MoonTag). Maximally 10 mRNAs were tracked per cell. To measure the intensity of fluorescence on mRNAs over time, we generated a semi-automated translation spot tracking in MatLab called TransTrack. TransTrack software is freely available, including documentation, through Github.

Fluorescence intensity of mRNAs

To analyze whether mRNAs associated with both SunTag and MoonTag signals are mRNA multimers, mCherry fluorescence intensity was measured for mRNAs that were associated with only MoonTag signal or for mRNAs associated with both SunTag and MoonTag signals. For intensity measurements, an ROI of 8x8 pixels was created that was used to measure mRNA mCherry intensity. Local background was subtracted from all measurements.

Normalization of fluorescence

Bleach correction was performed in TransTrack. In brief, to correct for photobleaching during the imaging, the fluorescence intensity of the entire field of view was determined at each time point of the movie. The fluorescence intensity over time was fit with an exponential decay distribution to determine the bleaching rate, and this rate was used to correct all fluorescence images.

We also found that cells with higher expression of the MoonTag-nanobody showed on average higher intensities of MoonTag translation sites (See Fig. S3A). Therefore, MoonTag translation site intensities were normalized to total cell MoonTag intensities. As SunTag translation site intensities poorly correlated with total cell intensities, no further correction was performed for the SunTag signal (See Fig. S3B).

Translation elongation rates of MashTag mRNAs

The elongation rates of ribosomes on individual mRNAs was determined by harringtonine run-off experiments as described previously (Yan et al., 2016). In brief, the translation inhibitor harringtonine (3 ug/ml) is added to cells expressing

the SunStart reporter, which freezes ribosomes on the start codon. Therefore, harringtonine prevents new ribosomes from translating the reporter, and allows ribosomes that are already in the translation elongation phase to continue translating until they reach the stop codon. As ribosomes terminate one-by-one, the SunTag signal of the terminating ribosome dissociates from the mRNA, resulting in a gradual reduction of GFP fluorescence on the mRNA until all ribosomes have terminated translation. The fluorescence decrease was tracked for each mRNA and normalized to the average intensity of the 3 time points before drug administration. Translation elongation rates were then calculated based on the slope of the GFP intensity trace, as described previously (Yan et al., 2016).

Plateau intensity of SunTag-frame ribosomes

The plateau intensity of a single ribosome translating the MashTag reporter (See Fig 2F), is equal to the intensity of a mature MashTag protein. To determine the intensity of mature proteins translated in the SunTag frame, mature proteins were tethered to the plasma membrane by encoding a C-terminal prenylation sequence (CAAX) in the reporter in the SunTag frame. This reporter was transfected into Moon/Sun cells used for MashTag translation imaging. No doxycycline was added in these experiments to keep the amount of mature protein to a minimum and thereby enable single-molecule imaging of mature proteins. Using identical imaging parameters as those used during translation imaging, images were acquired of cells containing mature membrane-tethered SunTag protein. For 15 foci (i.e. mature proteins) per cell the intensity was determined. For local background correction, the intensity within a ROI of the same size was determined in a region next to the foci. The mean intensity of 24xMashTag foci in the SunTag frame was 73.5 ± 35.7 a.u. (mean ± SD). For each reporter, the plateau intensity was normalized to the number of SunTag repeats. For example, this value was corrected to 110.3 a.u. for the MoonStart-MashTag reporter, as the MashTag in this reporter contained 36 instead of 24 repeats (See '**Theoretical single ribosome traces parameters per reporter'** for other reporters).

Plateau intensity of MoonTag-frame ribosomes

The plateau intensity of a ribosome translating the MashTag in the MoonTag frame could not be determined with the approach described above, as single mature proteins encoded in the MoonTag frame could not be reliably detected. As an alternative approach, the MoonTag translation site intensity of MoonStart mRNAs was determined (See Fig. S3C). For this, only mRNAs were included that contained exclusively MoonTag signal. The mean MoonTag intensity on MoonStart mRNAs was 1354.13 ± 739.65 a.u. (mean ± SD). This fluorescence signal originated from multiple ribosomes. Therefore, to determine the approximate MoonTag intensity associated with a single ribosome, the average number of ribosomes translating a MashTag mRNA was determined. Since the MashTag reporter contained the same promoter, 5'UTR and start codon as our previously described SunTag reporter (Yan et al., 2016), and both reporters had similar translation elongation rates, we assumed that both reporters had similar ribosome densities. Based on the previously determined inter-ribosomal distance on the SunTag reporter (Yan et al., 2016), the average number of ribosomes on a MashTag mRNA could be calculated (correcting

for mRNA length), which was 17.8. Based on the MoonTag translation site intensity and the number of ribosomes per mRNA, the average MoonTag intensity associated with a single ribosome on the MashTag reporter could be calculated, which was 76.07 a.u. Since only the ribosomes located downstream of the MashTag (i.e. on the BFP sequence) were associated with the full 36 copies of the MoonTag, a correction needed to be applied to obtain the intensity of a 36xMoonTag protein (i.e. the plateau intensity). After applying this correction (as described previously (Yan et al., 2016)), the plateau intensity of a ribosome translating the 36xMashTag in the MoonTag frame was calculated to be 132.7 ± 72 a.u. (mean ± SD), and the intensity of a single MoonTag frame encoded MashTag repeat was derived (3.69 a.u.). The combination of this single repeat intensity value and the number of repeats in a reporter was used to determine the plateau intensity of a single ribosome on each reporter (See '**Theoretical single ribosome traces parameters per reporter'**).

Theoretical single ribosome traces parameters per reporter

Based on the specific parameters for each reporter, and based on the equations shown below, we calculated the build-up time, plateau time, and plateau intensity of a single ribosome. For each reporter, the values describing the theoretical single ribosome trace were calculated for both the MoonTag and the SunTag signal. To check whether any SunTag ribosomes would be called on the MoonTag only reporter by our analysis pipeline, the theoretical single ribosome trace of MoonStart- 36xMashTag values were used. Similarly, the SunStart-36xMashTag frame values were used to test whether any SunTag ribosomes would be scored on the MoonTag reporter. These theoretical traces were selected, as these reporters did not contain MoonTag and SunTag peptides, respectively, so no true theoretical single ribosome intensity trace could be generated.

Equations:

$$t_{buildup} = (nt_{start\_plateau} - nt_{1st\ peptide})/elongation speed$$
$$t_{plateau} = (nt_{termination} - nt_{start\ plateau})/elongation speed$$
$$I_{plateau} = I_{single\_peptide} \times r$$
$$nt_{1st\ peptide} = |r| + start\_linker + rib.exit$$
$$nt_{start\ plateau} = nt_{1st\ peptide} + |r| \times (r - 1) + linker$$

<u>*Definitions:*</u>

$t_{buildup}$ = time in sec to build-up from no signal to plateau intensity

$t_{plateau}$ = time in sec from reaching plateau intensity to termination

$I_{plateau}$ = plateau intensity in a.u.

$nt_{1st\ peptide}$ = nucleotide position of ribosome where peptide can first be bound by antibody

$nt_{start\_plateau}$ = nucleotide position of ribosome where the plateau phase starts

$nt_{termination}$ = nucleotide position of ribosome at moment of termination. The length of the coding sequence (cds) in nucleotide is used.

Fixed values for each reporter:

elongation speed = ribosomal translocation speed during translation in nt/s. Based

on the elongation speed of 2.88 codons/s, this constant is 8.65 nt/s.

rib. exit = ribosome exit tunnel in nt. Based on previously used ribosome exit tunnel of 30 amino acids.

$I_{single\ peptide}$ = fluorescence intensity as a results of one antibody binding to a peptide epitope in a.u. One MoonTag repeat was 3.69 a.u. and one SunTag repeat was 3.06 a.u. using our imaging settings.

Constants based on reporter:

|r| = length of 1 peptide epitope in nt.

start_linker = distance between AUG and start of 1st peptide epitope in nt.

r = number of repeats.

linker= length of an optional linker in repeat array used for cloning (in nt).

reporter	translation signal	type of translation	length cds (nt)	Start_linker (nt)	|r| (nt)	r	linker (nt)	build-up time (s)	plateau time(s)	plateau intensity (a.u.) MoonTag	plateau intensity (a.u.) SunTag
Mstart-Mash1-BFP	MoonTag	main frame translation	4587	35	105	36	36	429	75	133	
	SunTag	alternative start site selection	4587	39	105	36	36	429	74		110
	SunTag	frameshifting after 25% of MashTag	4587	980	105	27	36	320	75		83
	SunTag	frameshifting after 50% of MashTag	4587	1925	105	18	36	210	75		55
Sstart-Mash1-BFP	SunTag	main frame translation	4587	39	105	36	36	429	74		110
	MoonTag	alternative start site translation	4587	35	105	36	36	429	75	133	
Sstart-SunTag-kif18b	SunTag	main frame translation	4386	69	72	24	0	191	289		74
	MoonTag	background MoonTag signal on SunTag	4587	35	105	36	36	429	75	133	
Mstart-MoonTag-kif18b	MoonTag	main frame translation	4017	12	60	24	0	159	286	88	
	SunTag	background SunTag signal on MoonTag	4587	39	105	36	36	429	74		110

12xMoon-12xSun-kif18b	SunTag	main frame translation	4440	39	81	12	0	103	386	44	
	MoonTag	main frame translation	4440	1002	72	12	0	92	287		37
MT-kif18b-STOP-ST-BFP	MoonTag	no read-through	4017	12	60	24	0	159	286	88	
	MoonTag	read-through	6495	12	60	24	0	159	545	88	
	SunTag	read-through	6495	4056	72	24	0	191	45		74

Noise in MoonTag intensity measurements

To determine the effects of differential nanobody occupancy over time, as well as imaging noise on the fluorescence intensity of the MoonTag signal, cells were treated with cycloheximide (CHX) (200 μg/ml). CHX locks ribosomes on the mRNA and limits changes in MoonTag signal due to changes in ribosome occupancy. Thus, changes in MoonTag fluorescence in CHX-treated cells mostly represent nanobody occupancy changes and imaging noise. The MoonTag fluorescence intensity was measured on individual mRNAs imaged at a time interval of 1s or 60s. A small decrease in MoonTag intensities was observed after prolonged CHX administration (which may reflect recycling of stalled ribosomes), therefore a correction was applied to normalize for this effect. To calculate the overall intensity fluctuations of MoonTag signals in CHX-treated cells, the standard deviation of the fluorescence intensity was determined for 6 consecutive time points and the standard deviation was then divided by the mean fluorescence intensity over the same time interval.

Excluding frameshifting as a source of SunTag signal

Based on the combination of the translation elongation rate, the plateau intensities and the length of the MashTag and downstream BFP, the theoretical single ribosome traces in either the MoonTag or the SunTag frame can be calculated (See **'Theoretical single ribosome traces parameters per reporter'**).

To determine the theoretical single ribosome intensity trace for a ribosome translating a MoonStart mRNA in the SunTag frame, two different scenarios were considered: 1) ribosomes could translate the MashTag reporter in the SunTag frame from the start of the coding sequence until the stop codon, or 2) ribosomes could start translating the coding sequence in the MoonTag frame and frameshift into the SunTag frame at some point during translation elongation. In the first scenario, the SunTag intensity trace on the MoonStart reporter would be similar to the SunTag intensity trace on the SunStart reporter (i.e. containing all 36 SunTag peptides). In the second scenario, fewer SunTag peptides would be translated, resulting in a shorter duration of the intensity trace and a lower plateau intensity. Assuming frameshifting occurs at a random position on the MashTag sequence, frameshifting would occur *on average* at the end of 18th MashTag repeat (halfway through translating the 36xMashTag). Therefore, an intensity trace was calculated that was predicted by the second (frameshifting) model, which contains 18 SunTag repeats. Frameshifting events that occur in the last few repeats of the MashTag may result in a very weak

signal, which could evade detection and bias the set of experimentally-detected frameshifted translation events towards more upstream frameshifting events that contain more SunTag peptides. Therefore, we also calculated the theoretical intensity trace assuming on average 27 SunTag peptides were translated, which corresponds to a detection limit of 18 SunTag peptides, well within our detection range (Yan et al., 2016).

To generate intensity traces of single ribosomes translating the SunTag frame, both for the Emi1-5'UTR-SunStart reporter and the MoonStart reporter, translation events were manually selected that were most likely to represent single ribosome translation events. For this, all translation events were selected that contained SunTag signal in at least 3 consecutive time points (1.5 min) and at most 16 consecutive time-points (8 min). The 8 min threshold was chosen based on the expected duration of a single ribosome translating the MashTag reporters of 8.4 min. All translation events that matched these criteria were aligned at the moment of GFP disappearance (i.e. translation termination) and the GFP intensity of the preceding 8 min was determined. All traces that contained a second translation event within the 8 min preceding time-period, or traces starting in the first 8 min of the movie were removed from the analysis. Average intensities and SEMs of all translation events included in the analysis were then calculated.

Intensity measurements of mature SunTag proteins

The intensity of individual mature SunTag proteins was measured for both the MoonStart-36xMashTag-CAAX and SunStart-36xMashTag-CAAX. Experimental settings and analysis of the MoonStart-36xMashTag-CAAX and SunStart-36xMashTag-CAAX (note that the CAAX motif was in the SunTag frame in both cases) were similar to the experiments described in section **'Plateau intensity of SunTag-frame ribosomes'** to determine the plateau intensity of ribosomes translating the SunTag frame, except that higher laser powers and shorter exposure times were used in the analysis of mature protein intensities to facilitate detection.

Fitting intensity traces using RiboFitter

To determine the number of ribosomes translating an mRNA and the exact time at which single ribosomes initiated translation, we developed a computational model in R, called RiboFitter. RiboFitter is freely available, including documentation, through Github.

RiboFitter reconstitutes a raw intensity trace by positioning one or more theoretical intensity profile of a single translating ribosome along the time axis of the trace. The sum intensity for each time point of all theoretical single ribosome intensity traces is calculated and compared to the raw intensity trace at that time point. By optimizing the number and the position of ribosomes in an iterative fashion RiboFitter achieves the optimal fit to the data.

$$I(t) = \begin{cases} I_{buildup}t & t \in [t, t + t_{buildup}) \\ I_{plateau} & t \in [t + t_{buildup}, \; t + t_{buildup} + t_{plateau}] \end{cases} \qquad [1]$$

Equation [1] describes the intensity profile of a single translating ribosome, which initiates translation at time t (See Fig. 2F). $I_{plateau}$, $t_{buildup}$ and $t_{plateau}$ are constant values for each reporter, and are dependent on the length of the mRNA and number of MashTag repeats, as described in sections '**Translation elongation rates of MashTag mRNAs**', '**Plateau intensity of SunTag-frame ribosomes**', '**Plateau intensity of MoonTag-frame ribosomes**', and '**Theoretical single ribosome traces parameters per reporter'**. The raw intensity traces were reassembled *in silico* by fitting the sum of one or more theoretical intensity traces of individual ribosomes that initiate translation at time points t = t_i [2]. The *in silico* fitting of ribosomes was performed on the SunTag and MoonTag signals independently and the output from both signals per mRNA were combined after the fitting.

$$I(t) = \sum_{t=t_i}^{t=t_i+t_{buildup}} I_{buildup}t + \sum_{t=t_i+t_{buildup}}^{t=t_i+t_{buildup}+t_{plateau}} I_{plateau} \qquad [2]$$

To obtain an initial estimate of the number of ribosomes translating an mRNA based on the raw intensity trace, the area under the curve (AUC) of the entire intensity trace was divided by the AUC of the intensity trace of a single translating ribosome. If the AUC of an mRNA did not exceed half of the expected AUC of a translating ribosome, no further fitting of ribosomes to the intensity trace was performed and the number of ribosomes in the raw intensity trace was determined to be 0. If the raw intensity curve was determined to exceed half of the expected AUC of a translating ribosome, the model estimated the number of translation events in each raw intensity trace based on the AUC values. These translation events were then distributed along the raw intensity trace with a probability that is weighted by the intensity of the trace at each time point. The sum of all the positioned single ribosome intensity traces was then determined, and the sum intensity trace was compared with the raw intensity trace to determine a goodness of fit, which was defined as the root mean square error (RMSE) between the fit and the data.

After initial placement of translation events, the number of ribosomes and their time of initiation (i.e. their relative position along the time axis of the trace) were altered according to the parameters shown below, resulting in a new fit. If the RMSE between the new fit and the data was lower (i.e. the fit improved), the new positions of translation events were used as a starting point for the next iteration. If the RMSE did not improve, the previous positions were used again as a starting point for the next iteration. Of note, if the best fit was achieved with a trace containing no ribosomes, the number of ribosomes was considered 0, even if the AUC of the total intensity trace exceeded the AUC of a single ribosome. The process of re-positioning translation events an accepting or rejecting the new positions was repeated for 1000 iterations to obtain a good fit. A limit of 1000 iterations was selected, because minimal improvements in the fit were achieved with additional iterations (See Fig. S3D-F). Since the fitting process is stochastic, multiple runs of RiboFitter could result in different outcomes. RiboFitter was therefore run 10 independent times for each intensity trace to check for variations in position of translation initiation events and the final RMSE, and the run with the best fit was used to generate the final fit. Note that the 10 runs generally resulted in very similar fits (See Fig. S3D), demonstrating the robustness of this approach.

Parameters used to fit intensity trace to fitting trace:

Parameter	Value
Probability to locate a ribosome at time point t in first iteration	$$I(t)/\sum_{0}^{t_{end}} I(t)$$
Number of iterations	1000
Probability to add or remove a ribosome	0.1
Probability to relocate the time point at which the ribosome initiates	0.3
Relocation distance (sec) (i.e. repositioning a ribosome along the trace)	The distance to move a translation event is randomly drawn from a normal distribution with μ = 0 sec. and σ = 25 sec.

Many mRNAs already contained SunTag and/or MoonTag signal at the first time point of the intensity trace. However, translation initiation of the ribosomes present on the mRNA at the first time point of the movie took place before the start of the trace, preventing proper fitting. To overcome this limitation, hypothetical time points were added before the start of the intensity trace. The number of added time points depended on the $t_{buildup}$ per reporter. Addition of extra time points allowed the model to position initiation events on the trace prior to the start of image acquisition and, hence, to generate a signal that resembles the intensity at t = 0. In the initial placement of ribosomes during the first iteration, the probability of positioning an initiating ribosome before t = 0 was equal to the maximal probability to position an initiation event at any time point of the data. The hypothetical time points were not included in calculating the RMSE, as there was no raw intensity measurement that could be compared to the fit during this period.

As output, the model generated an overview containing the following information for both the MoonTag and the SunTag signal on each mRNA: 1) a graph showing the best fit and the RMSE, 2) the total number of ribosomes per reading frame on the mRNA, 3) the initiation time of each ribosome. The initiation time and the duration of translation of a single ribosome could then be used to calculate for each time point the number of ribosomes per reading frame on the mRNA.

Error range for intensity fitting approach

The intensity profile of a single translating ribosome, which is used in the iterative intensity fitting approach described above (See '**Fitting intensity traces using RiboFitter'**), is based on the experimentally determined *average* ribosomal elongation speed and *average* single ribosome plateau intensities. However, both the translation elongation speed and plateau intensity are highly variable values when comparing different ribosomes, resulting in distinct values for the AUC of intensity traces for different ribosomes. Using the average values for the single ribosome intensity trace does not capture this variation and does not allow for an estimation of the error in RiboFitter that is caused by heterogeneity in single ribosome intensity traces. As the values obtained through RiboFitter are used for subsequent calculations of the frequency of different translation paths (See Fig. 3A, B), no error range can be estimated for the usage frequency of the different translation paths either when using the average single ribosome intensity trace. Therefore, we

developed a second approach (termed the *variable* AUC approach) to estimate the error range in RiboFitter.

Rather than creating a single AUC based on the average $I_{plateau}$, $t_{buildup}$ and $t_{elongation}$, the variable AUC approach creates a unique AUC for each individual ribosome used in the fitting. The variable AUC algorithm simulates an AUC by randomly selecting one value for the plateau intensity and one value for the elongation rate from the respective normal distributions with means and standard deviations equal to the experimentally determined values for these parameters. A lower limit of 0.5 amino acids per second was introduced for the elongation rate, as the normal distribution of the elongation rates that was generated also contained unrealistically small values for the elongation rate (including negative values). After the algorithm generated a single ribosome intensity trace, the AUC of that trace was calculated and compared to the AUC of the translation site intensity trace. If the value for the AUC of the single ribosome intensity traces was smaller than the value of the AUC of the translation site intensity trace, a new unique single ribosome intensity trace was generated, and the value of the AUCs were added together and compared to the value for the translation site intensity again. This process is repeated until the sum of the AUC values of all the individual ribosomes is equal to or greater than the AUC value of the translation site. The number of single ribosome traces needed to match the AUC of the translation site intensity trace is then recorded.

As the intensity traces of the MoonTag and the SunTag signals are separated, the algorithm provides the number of ribosomes translating in each frame. Based on the number of ribosomes and the duration of the intensity trace, the translation initiation rate was calculated in each frame (See '**Translation initiation rates**'). The frequency of SunTag translating ribosomes relative to total number of ribosomes per mRNA was calculated to determine the percentage of OOF translation (See '**SunTag ribosome frequency**'). The process of determining the number of ribosomes translating the mRNA (based on the variable AUC approach), calculating the translation rates and determining the SunTag translation frequency was performed for all mRNA molecules in an experiment and the median value was determined. This entire process was then repeated 1000 times, resulting in 1000 values for the median, allowing us to calculate a range for these median values (i.e. an error range). Comparison of the two different algorithms (fitting intensity traces with fixed or variable single ribosome AUCs) to calculate the OOF frequency indicated that both approaches yielded comparable results (See Fig. S4A). Therefore, both approaches can be used to define the number of ribosomes per reading frame based on the intensity trace. Depending on the downstream analysis, one algorithm can be preferred over the other. The fixed AUC approach provides the possibility to analyze changes in translation over time as individual translation events are positioned along the time axis, whereas the second algorithm with variable single ribosome AUCs enables calculation of error ranges for the fitting process.

SunTag ribosome frequency

As described in the section '**Fitting intensity traces using RiboFitter**', the number of MoonTag and SunTag ribosomes per mRNA was calculated. To determine the fraction of SunTag ribosomes, the number of SunTag ribosomes was divided by the sum of the number of MoonTag and SunTag ribosomes.

Translation initiation rates

To determine translation initiation rates of individual mRNAs, the total number of ribosomes translating both SunTag and MoonTag frames was calculated for each mRNA, as described in the section '**Fitting intensity traces using RiboFitter'**. Initiation rates in each frame were then calculated by dividing the total number of initiating ribosomes by the duration of the trace. Overall initiation rates were based on the sum of all ribosomes translating the MoonTag and the SunTag frames. Ribosomes translating the blank frame were not included in the calculation of initiation rates, as they could not be determined.

3'UTR translation

In the 3'UTR translation reporter, translation of the coding sequence results in MoonTag signal, while translation of the 3'UTR results in SunTag signal. To determine the number of ribosomes translating the 3'UTR for each mRNA, the number of ribosomes translating both MoonTag and SunTag was determined as described in the section '**Fitting intensity traces using RiboFitter'**. For the fitting process of the 3'UTR translation reporter a separate theoretical single ribosome intensity trace had to be developed, because the 3'UTR translation reporter had different numbers of SunTag and MoonTag repeats and the coding sequence was different in length than the MashTag reporter (See '**Theoretical single ribosome traces parameters per reporter'**). The theoretical single ribosome intensity trace of the SunTag was based on a buildup time of 191 s and a plateau time of 45 s, which correspond to the time to translate 24x SunTag repeats and the downstream linker sequence, respectively. The single ribosome trace of the MoonTag was based on a buildup time of 160 s and a plateau time of 286 s, which correspond to the time to translate 24xMoonTag repeats and the downstream sequence (Kif18b). However, if a ribosome reads through the stop codon (which is likely the predominant mechanisms of 3'UTR translation in this reporter), the plateau phase for MoonTag translation is longer (545 s instead of 286 s), as the sequence that is translated downstream of the MoonTag signal now includes the SunTag and BFP as well. A read-through ribosome therefore gives rise to more total fluorescence. To correct for this, a correction factor was calculated (0.78) based on the fold difference in the AUC of a single ribosome translating the MoonTag and Kif18b sequence compared to a ribosome translating the MoonTag, Kif18b, SunTag and BFP. To calculate the frequency of 3'UTR translation, the number of SunTag ribosomes was divided by the number of MoonTag ribosomes, after correction.

Calculating the # of ribosomes per translation path

Translation of a uORF-containing mRNA can result in various different translation paths. The options of single ribosomes were envisioned as a roadmap with a chance to initiate translation or terminate on each start site or stop site encountered by a ribosome, leading to 10 different paths (See Fig. 6B):

1. Initiation on uORF AUG; termination on uORF; dissociation from mRNA.
2. Initiation on uORF AUG; termination on uORF stop codon; reinitiation of scanning; initiation on MoonTag frame AUG; translation of MashTag in MoonTag frame.

3. Initiation on uORF AUG; termination on uORF stop codon; reinitiation of scanning; skipping of MoonTag frame AUG; downstream initiation in blank frame; translation of MashTag in blank frame.

4. Initiation on uORF AUG; termination on uORF stop codon; reinitiation of scanning; skipping of MoonTag frame AUG; downstream initiation in MoonTag frame; translation of MashTag in MoonTag frame.

5. Initiation on uORF AUG; termination on uORF stop codon; reinitiation of scanning; skipping of MoonTag frame AUG; downstream initiation in SunTag frame; translation of MashTag in SunTag frame.

6. Skipping of uORF AUG; initiation on SunTag frame AUG; translation of MashTag in SunTag frame.

7. Skipping of uORF AUG; skipping of SunTag frame AUG; initiation on MoonTag frame AUG; translation of MashTag in MoonTag frame.

8. Skipping of uORF AUG; skipping of SunTag frame AUG; skipping of MoonTag frame AUG; downstream initiation in blank frame; translation of MashTag in blank frame.

9. Skipping of uORF AUG; skipping of SunTag frame AUG; skipping of MoonTag frame AUG; downstream initiation in MoonTag frame; translation of MashTag in MoonTag frame.

10. Skipping of uORF AUG; skipping of SunTag frame AUG; skipping of MoonTag frame AUG; downstream initiation in SunTag frame; translation of MashTag in SunTag frame.

The following equations were used to calculate the fraction of ribosomes in each path:

$path1 = i \times (1 - r)$

$path2 = i \times r \times i$

$path3 = i \times r \times (1 - i) \times b$

$path4 = i \times r \times (1 - i) \times m$

$path5 = i \times r \times (1 - i) \times s$

$path6 = (1 - i) \times i$

$path7 = (1 - i) \times (1 - i) \times i$

$path8 = (1 - i) \times (1 - i) \times (1 - i) \times b$

$path9 = (1 - i) \times (1 - i) \times (1 - i) \times m$

$path10 = (1 - i) \times (1 - i) \times (1 - i) \times s$

i = probability to initiate translation on an AUG start codon in optimal context.

r = probability to reinitiate scanning after uORF translation and termination at the uORF stop codon.

b, m, or s = probability to initiate translation on a near-cognate start site in either the blank (b), the MoonTag (m), or the SunTag (s) frame downstream of the MoonTag AUG start site

Multiple different paths lead to translation signal of one of the three frames. Therefore:

$$fraction\ of\ ribosomes\ translating\ blank\ frame\ =\ path1 + path3 + path8$$

$$fraction\ of\ ribosomes\ translating\ MoonTag\ frame\ =\ path2 + path4 + path7 + path9$$

$$fraction\ of\ ribosomes\ translating\ SunTag\ frame\ =\ path5 + path6 + path10$$

To allow calculations of the values for the different constants (i, r, n) and the fraction of ribosomes in each path, the following assumptions were made:

The probability of a scanning ribosome to initiate translation on an AUG in optimal context (i.e. GCCACC<u>AUG</u>G) that is encountered, is constant irrespective of the position of the AUG in the mRNA.

The probability to initiate on a near-cognate start site downstream of the MoonTag AUG start site is equal for all three frame (MoonTag, SunTag, and blank frame).

The combined translation initiation rate of the MoonTag, SunTag, and blank frame is equal for the MoonStart and uORF reporter.

To determine the probability of a ribosome initiating translation on an AUG start site in optimal context (i), we examined the frequency of translation initiation at the AUG of the MoonStart reporter. The median initiation rate in the MoonTag frame on the MoonStart reporter was 0.63 ± 0.03 ribosomes/min (Note: all calculations in this section are based on the variable AUC approach intensity fitting approach, see '**Error range for intensity fitting approach'**), which represents a combination of translation initiation on the MoonTag AUG start codon, and skipping of the AUG and downstream initiation in the MoonTag frame. While we cannot de-convolve these two paths directly, we could measure the translation initiation rate in the SunTag frame (0.07 ± 0.01 ribosomes/min), which represents leaky scanning of the MoonTag AUG and downstream initiation on a near-cognate start site in the SunTag frame. One of our assumptions is that the frequency of initiation on near-cognate start sites downstream of the MoonTag AUG is equal in all three frames. Based on this assumption b = m = s = 0.33. Thus, if 0.07 ribosomes/min initiate on a near-cognate start site in the SunTag frame, we assume that 0.07 ribosomes/min are also initiating in the MoonTag frame on a near-cognate start site, leaving 0.63 - 0.07 = 0.54 ribosomes/min for initiation on the MoonTag AUG. Addition of a similar blank frame initiation rate (0.07 ribosomes/min) results in an overall initiation rate of 0.07 + 0.07 + 0.07 + 0.54 = 0.75 for all three frames combined. 0.54 of 0.75 ribosomes/min are initiating on the MoonTag AUG, which represents 0.54 / 0.75 = 0.72 of initiating ribosomes. Therefore i = 0.72, which is comparable to a previously determined probability of initiating on an AUG start codon with optimal context (i = 0.86 based on (Ferreira et al., 2013)).

Finally, we determined the remaining constant r, the probability to reinitiate scanning after uORF translation. To calculate r, we used the following equation:

Equation 1

$$fraction\ of\ ribosomes\ translating\ MoonTag\ frame\ =\ path2 + path4 + path7 + path9 =$$

$$i \times r \times i + i \times r \times (1 - i) \times m + (1 - i) \times (1 - i) \times i + (1 - i) \times (1 - i) \times (1 - i) \times m$$

To solve this equation, we first calculated the fraction of ribosomes translating the MoonTag frame. The median MoonTag translation rate on the uORF reporter was 0.28 ± 0.02 ribosomes/min. As we assumed that the combined translation initiation rate is equal for the MoonStart and uORF reporter, and we had already determined the total initiation rate for the MoonStart reporter to be 0.75 ribosomes/min, the total translation rate on the uORF reporter is also 0.75 ribosomes/min and thus the relative fraction of ribosomes translating the MoonTag frame is 0.28/0.75 = 0.37. Using this value, we could solve equation 1, which led to r = 0.51.

In parallel, r was calculated based on the total blank frame translation rate, using the following equation:

Equation 2:

$fraction\ of\ ribosomes\ translation\ blank\ frame = path1 + path3 + path8 =$

$i \times (1 - r) + i \times r \times (1 - i) \times b + (1 - i) \times (1 - i) \times (1 - i) \times b$

To solve this equation, we first calculated the fraction of ribosomes translating the blank frame, based on the assumption that combined translation initiation rate is equal for the MoonStart and uORF reporter (0.75 ribosomes/min). The median MoonTag (0.28 ± 0.02 ribosomes/min) and SunTag translation (0.21 ± 0.02 ribosomes/min) rates were then used to calculate the blank frame translation rate: 0.75 - 0.28 - 0.21 = 0.26 ribosomes/min. The relative fraction of ribosomes translation the blank frame is therefore: 0.26 / 0.75 = 0.35. Using this value, we could solve equation 2, and determine r = 0.64. The mean of the two values for r that were calculated by solving equations 1 and 2 was used as the final value for r (0.58).

To include error ranges in the values for the constants r and i, similar calculations were performed as described above. However, instead of using the *average* value of the 1000 simulated median values that were obtained by the variable AUC simulations (see '**Error range for intensity fitting approach')**, we used values that were ± 1 SD away from the average for subsequent calculations. To generate conservative estimates of the error ranges of r and i, we used a value that was +1 SD for the SunTag initiation rate and -1 SD for the MoonTag initiation rate. By selecting the upper value (+1 SD) for the SunTag and the lower value (-1 SD) for the MoonTag (or vice versa), the largest possible error range is created, since the SunTag/MoonTag ratio is used to calculate the constants i and r. Using this approach, the following error ranges were determined: i = 0.67-0.76 and r = 0.53-0.63. These values were used to determine the error range of the fraction of ribosomes per path (See Fig. 6B).

To confirm these constants, we took advantage of the experiment in which we removed the stop codon of the uORF. Based on the constants, we predicted the effects of removing the uORF stop codon on the translation paths and the translation rates per frame. Removal of the stop codon prevents reinitiation at the MoonTag AUG start codon after uORF translation and thus leads to an r = 0 value. With this new value for r, the fraction of ribosomes translating each translation path was computed. Then, the translation rate of the MoonTag and SunTag combined was calculated (r = 0; 0.18-0.24 ribosomes/min). Similarly, the translation rate of SunTag and MoonTag combined was calculated in the presence of the uORF stop codon (r = 0.58; 0.47-0.52 ribosomes/min). These calculations revealed that the removal of the

uORF stop codon is predicted to induce a 54-62% (from 0.50 to 0.21 ribosomes/min) reduction in the combined SunTag and MoonTag translation rate. When measuring the translation rates on the uORF reporter in the presence or absence of the uORF stop codon, we observed 55% reduction (from 0.56 ± 0.03 to 0.25 ± 0.02 ribosomes/min) in median translation rates, validating our calculated parameters.

Statistics of population differences

To compare different data sets, students' T tests or Mann-Whitney tests were performed, as indicated in the figure legends.

Statistical analysis of heterogeneity in OOF translation

To test whether observed differences in the amount of OOF translation on different MashTag reporter mRNAs could be explained by chance, the OOF translation frequency of individual mRNAs was compared to the OOF translation frequency of the total population of mRNAs. Per mRNA a binomial test was conducted, using the total number of ribosomes on each mRNA, the number of ribosomes translating the SunTag frame on the mRNA and the median SunTag translation frequency on all mRNAs. Note that this analysis had limited statistical power to detect MoonStart mRNAs with increased MoonTag translation; based on analysis of all mRNAs, 93% of ribosomes initiate translation in the MoonTag frame on the MoonStart reporter. Even if all ribosomes initiate in the MoonTag frame during our observation period (~25 min, ~25-50 initiating ribosomes) the bias towards the MoonTag start site is only moderately statistically significant, so much longer traces are needed to achieve strongly significant p-values for enrichment of MoonTag translation. Nonetheless, we do find 2/85 mRNAs with an enrichment in MoonTag translation initiation events with a p-value > 0.01 and 8/85 mRNAs with a p-value > 0.05.

Linear regression analysis of intensities

To perform linear regression analysis on SunTag and MoonTag fluorescence over time, intensity traces of individual mRNAs were first smoothed by generating moving averages of 3 consecutive time points to reduce the technical noise in intensity traces. For every mRNA, MoonTag intensities were then plotted against SunTag intensities for each time point, linear regression analysis was performed on the SunTag and MoonTag intensity scatterplot, and the R^2-value was determined for each mRNA.

Of note, this analysis method may underestimate the correlation between MoonTag and SunTag signals, especially for the Moon-SunTag reporter that was used as control. First, all MoonTag peptides are positioned upstream of the SunTag peptides in the Moon-SunTag reporter. As a result, the MoonTag signal is observed slightly earlier than the SunTag signal during increases in translation of the Moon-SunTag reporter (See Fig. S4E, bottom left graph). The correlation between the MoonTag and SunTag fluorescence is thus slightly lower than expected. This temporal shift only occurs during increases in translation on the Moon-SunTag reporter, not during decreases in translation, as the MoonTag and SunTag signals disappear simultaneously during translation termination. Therefore, this temporal shift cannot be corrected by

shifting the entire SunTag intensity trace in time. This temporal shift does not occur during translation of the MashTag reporter, as the MoonTag and SunTag peptides are interspersed in the MashTag.

A second reason why the correlation between SunTag and MoonTag signals may be underestimated in the linear regression approach is that a strong correlation only occurs if there are substantial changes in the intensities over time and if the intensities are strong. If both signals remain largely constant over time or are low, the regression analysis mainly compares technical noise in the SunTag and MoonTag intensity traces. So, even though both MoonTag and SunTag fluorescence behave similar over time, a very poor correlation is observed (See Fig. S4E, bottom right graph). To somewhat mitigate this issue, only mRNAs with at least two-fold changes in the MoonTag signal and at least 0.25 ribosome/min were included in the linear regression analysis. However, this did not completely eliminate the problem, as even mRNAs with >2-fold changes in the fluorescence over time showed a reduction in the R^2 value due to periods in the time trace that had a constant rate of translation in both frames (See Fig. S4E).

Thus, the linear regression analysis of the MoonTag and SunTag provides a conservative estimate of the temporal correlation between the two signals.

Sliding window p-value analysis

To test whether changes in the usage of translation initiation sites occurred during the life-time of an mRNA, a statistical test was designed, referred to as a sliding window approach (See Fig. S4H). First, ribosomes were fit to raw intensity traces, and the time of each translation initiation event was determined, as described in the section '**Fitting intensity traces using RiboFitter'**. Second, initiation events in both MoonTag and SunTag frames were merged onto a single time-line. Third, a 'window' of 10 consecutive initiation events was defined as the first 10 initiation events of the trace. Additional windows were defined by sliding the window along all initiation events, moving the window one initiation event per iteration. Note that mRNA traces with 10 or less initiation events were excluded from analysis, as we could not generate multiple windows on these mRNAs. In this way, a collection of windows was created along the trace of an mRNA, in which each window represented 10 consecutive initiating ribosomes. For each window, the number of SunTag initiation events was determined. Fourth, the relative SunTag initiation frequency in each window was compared to the SunTag initiation frequency of the entire mRNA to test whether the window contained an increase or decrease in the number of SunTag initiation events as compared to the total trace. To provide a significance value for each window, a binomial test was performed. For each window of an mRNA and the lowest p-value per mRNA was reported.

DATA AND SOFTWARE AVAILABILITY
TransTrack (MatLab) and RiboFitter (R) and documentation are made available through Github.

SUPPLEMENTAL ITEMS

Supplemental Movie S1-6 and Data 1 are available via https://doi.org/10.1016/j.cell.2019.05.001

Table S1 – related to all figures. Number of experimental repeats, cells, and mRNAs analyzed per experiment

Overview of the number of repeats per experiment, and of the number of cells and mRNAs analyzed per experiment. Some datasets are used for multiple analyses, or sometimes the identical analysis is shown again for comparison with a new analysis. In both cases, this is indicated with a colored cell in the column '(re)plotting'. Figure panels based on the same mRNAs are indicated with the same color. Replotting of the same analysis is also indicated in the figure legends.

Fig.	Description	Repeats	Cells	mRNAs or events	(re)plotting data
1D	4xMoonTag-11aalinker	1	32	32	
	12xMoonTag-11aalinker	1	26	26	
	24xMoonTag-11aalinker	1	23	23	
	24xMoonTag-5aalinker	1	25	25	
1I	24xMoonTag-kif18b-STOP-24xSunTag-BFP	4	16	74	(gray)
2G	5'UTR_Emi-Sstart-36xMash1-BFP-3xStop-24xPP7	4	8	17	a
2H	Mstart-36xMash1-BFP-3xStop-24xPP7	3	10	17	a (dark blue)
3C	24xSunTag-kif18b-stop-24xPP7	2	12	54	
	24xMoonTag-kif18b-stop-24xPP7	2	14	61	
	12xMoonTag-12xSunTag-kif18b-stop-24xPP7	3	12	32	(orange)
4A	Mstart-36xMash1-BFP-3xStop-24xPP7	4	17	85	(dark blue)
4C	Mstart-36xMash1-BFP-3xStop-24xPP7	4	17	85	(dark blue)
4D	Mstart-36xMash1-BFP-3xStop-24xPP7	4	17	85	(dark blue)
4E	12xMoonTag-12xSunTag-kif18b-stop-24xPP7	3	12	32	b (orange)
	Mstart-36xMash1-BFP-3xStop-24xPP7	4	17	32	b (dark blue)
4F	Mstart-36xMash1-BFP-3xStop-24xPP7	4	17	63	c (dark blue)
5B	Mstart-36xMash1-BFP-3xStop-24xPP7	4	17	85	(dark blue)
	Mstart-Bstart-36xMash1-BFP-3xStop-24xPP7	3	18	71	
	Mstart-Mstart-36xMash1-BFP-3xStop-24xPP7	2	17	53	
	Mstart-Sstart-36xMash1-BFP-3xStop-24xPP7	4	16	94	
5C	Mstart-36xMash1-BFP-3xStop-24xPP7	4	17	85	(dark blue)
	ΔSunStop-Mstart-36xMash1-BFP-3xStop-24xPP7	4	13	62	(red)
5D	Mstart-36xMash1-BFP-3xStop-24xPP7	4	17	85	(dark blue)
	HMBS_5'UTR-Mstart-36xMash1-BFP-3xStop-24xPP7	3	15	54	(orange)
	SERGEF_5'UTR-Mstart-36xMash1-BFP-3xStop-24xPP7	3	16	64	(cyan)
6E	uORF(Blank)-Sstart(in_uORF)-Mstart(reint)-36xMash-BFP-3xstop-24xPP7	2	10	53	(light blue)
6F	12xMoonTag-12xSunTag-kif18b-stop-24xPP7	3	12	32	b (orange)
	uORF(Blank)-Sstart(in_uORF)-Mstart(reint)-36xMash-BFP-3xstop-24xPP7	2	10	30	b (light blue)
6G	uORF(Blank)-Sstart(in_uORF)-Mstart(reint)-36xMash-BFP-3xstop-24xPP7	2	10	37	c (light blue)
S1D	24xMT-kif18b-STOP-24xST-BFP	4	16	74	(gray)
S2C	Mstart-36xMash1-BFP-3xStop-24xPP7_MoonTag only	3	29	78	(light blue)
	Mstart-36xMash1-BFP-3xStop-24xPP7_MoonTag and SunTag	3	29	59	(light blue)

S2D	Sstart-36xMash1-BFP-3xstop-24xPP7 + Harr	3	8	26	
S2E	Mstart-24xMash1-BFP-CAAX(Sun)	2	59	505	
S2F	Sstart-36xMash1-BFP-CAAX(Sun)	2	34	407	
	Mstart-36xMash1-BFP-CAAX(Sun)	2	33	399	
S3A	Mstart-36xMash1-BFP-3xStop-24xPP7	4	17	85	
S3B	Sstart-36xMash1-BFP-3xstop-24xPP7	4	17	43	
S3D	Mstart-36xMash1-BFP-3xStop-24xPP7	5	36	195	
S4A	Mstart-36xMash1-BFP-3xStop-24xPP7	4	17	85	
S4B	MStart-36xMash1-BFP-3xStop-24xPP7_Northern blot	2			
S4C	Mstart-36xMash1-BFP-3xStop-24xPP7_transient	4	17	85	
	Mstart-36xMash1-BFP-3xStop-24xPP7_knock-in	3	29	165	
S4D	Mstart-36xMash1-BFP-3xStop-24xPP7_transient	4	17	85	
	Mstart-36xMash1-BFP-3xStop-24xPP7_knock-in	3	29	165	
S4F	Mstart-36xMash1-BFP-3xStop-24xPP7_transient	4	17	32	b
	Mstart-36xMash1-BFP-3xStop-24xPP7_knock-in	3	29	45	b
S4G	Mstart-36xMash1-BFP-3xStop-24xPP7_1s interval	2	27	181	
	Mstart-36xMash1-BFP-3xStop-24xPP7_30s and 60s interval	3	36	202	
S4J	Mstart-36xMash1-BFP-3xStop-24xPP7_transient	4	17	63	c
	Mstart-36xMash1-BFP-3xStop-24xPP7_knock-in	3	29	144	c
S5A	ΔSunStop-Mstart-36xMash1-BFP-3xStop-24xPP7	4	13	62	
	ΔSunStop-Mstart-36xMash1-BFP-3xStop-24xPP7 + DMSO	4	13	58	
	ΔSunStop-Mstart-36xMash1-BFP-3xStop-24xPP7 + RocA	3	14	51	
S5B	ΔSunStop-Mstart-36xMash1-BFP-3xStop-24xPP7	4	13	62	
	ΔSunStop-Mstart-36xMash1-BFP-3xStop-24xPP7 + DMSO	4	13	58	
	ΔSunStop-Mstart-36xMash1-BFP-3xStop-24xPP7 + RocA	3	14	51	
S5C	Mstart-36xMash1-BFP-3xStop-24xPP7	4	17	85	
	HMBS_5'UTR-Mstart-36xMash1-BFP-3xStop-24xPP7	3	15	54	
	SERGEF_5'UTR-Mstart-36xMash1-BFP-3xStop-24xPP7	3	16	64	
S6A	uORF(Blank)-Sstart(in_uORF)-Mstart(reint)-36xMash-BFP-3xstop-24xPP7	2	10	53	
	uORF(Blank)-Mstart(in_uORF)-Sstart(reint)-36xMash-BFP-3xstop-24xPP7	3	9	34	
S6B	Mstart-36xMash1-BFP-3xStop-24xPP7	4	17	85	
	uORF(Blank)-Sstart(in_uORF)-Mstart(reint)-36xMash-BFP-3xstop-24xPP7	2	10	53	
	uORF(Blank)_nouORFStop-Sstart(in_uORF)-Mstart(reint)-36xMash-BFP-3xstop-24xPP7	3	17	45	
S6C	Mstart-36xMash1-BFP-3xStop-24xPP7	4	17	85	
	uORF(Blank)-Sstart(in_uORF)-Mstart(reint)-36xMash-BFP-3xstop-24xPP7	2	10	53	
	uORF(Blank)_nouORFStop-Sstart(in_uORF)-Mstart(reint)-36xMash-BFP-3xstop-24xPP7	3	17	45	
S6D	Mstart-36xMash1-BFP-3xStop-24xPP7	4	17	85	
	uORF(Blank)-Sstart(in_uORF)-Mstart(reint)-36xMash-BFP-3xstop-24xPP7	2	10	53	
	uORF(Blank)_nouORFStop-Sstart(in_uORF)-Mstart(reint)-36xMash-BFP-3xstop-24xPP7	3	17	45	

[a] only OOF events that fulfill criteria (see STAR Methods) are included

[b] only mRNAs with initiation events in both reading frames are included

[c] only mRNAs with >10 initiation events during timelapse are included
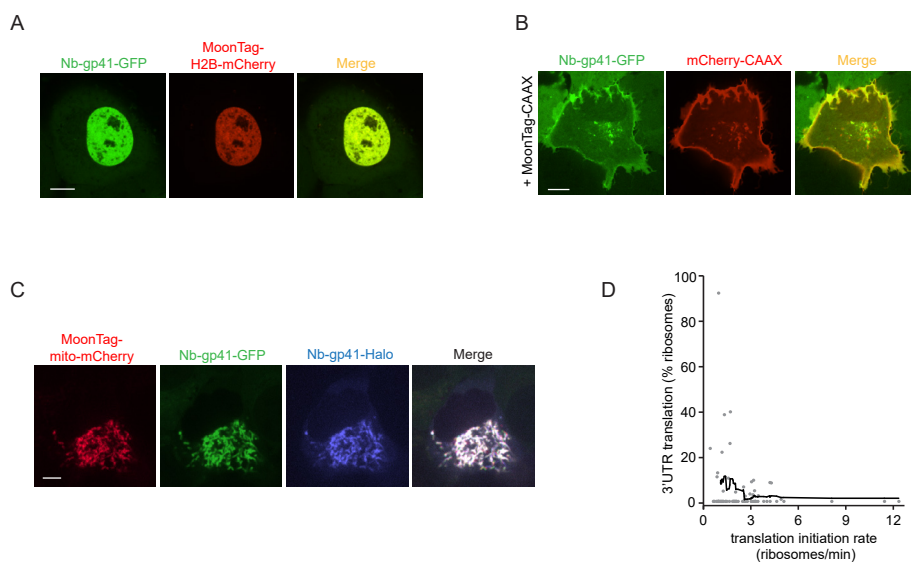
Supplemental figure S1 – related to Fig. 1. Development and characterization of MoonTag system.

A-C) U2OS cells stably expressing MoonTag nanobody-GFP were transfected with 12xMoonTag-H2B-mCherry (A), 12xMoonTag-CAAX and mCherry-CAAX (B), or 12xMoonTag-Mito-mCherry and MoonTag nanobody-HaloJF646 (C). Representative cells are shown. Scale bars, 10µm (A, B), and 5µm (C). D) Moon/Sun cells expressing the reporter indicated in Fig. 1H. Correlation between the main coding sequence translation initiation rate and 3'UTR translation frequency on single mRNAs is shown. Every dot represents a single mRNA and line depicts moving average over 15 mRNAs. Number of experimental repeats and mRNAs analyzed per experiment are listed in Table S1.
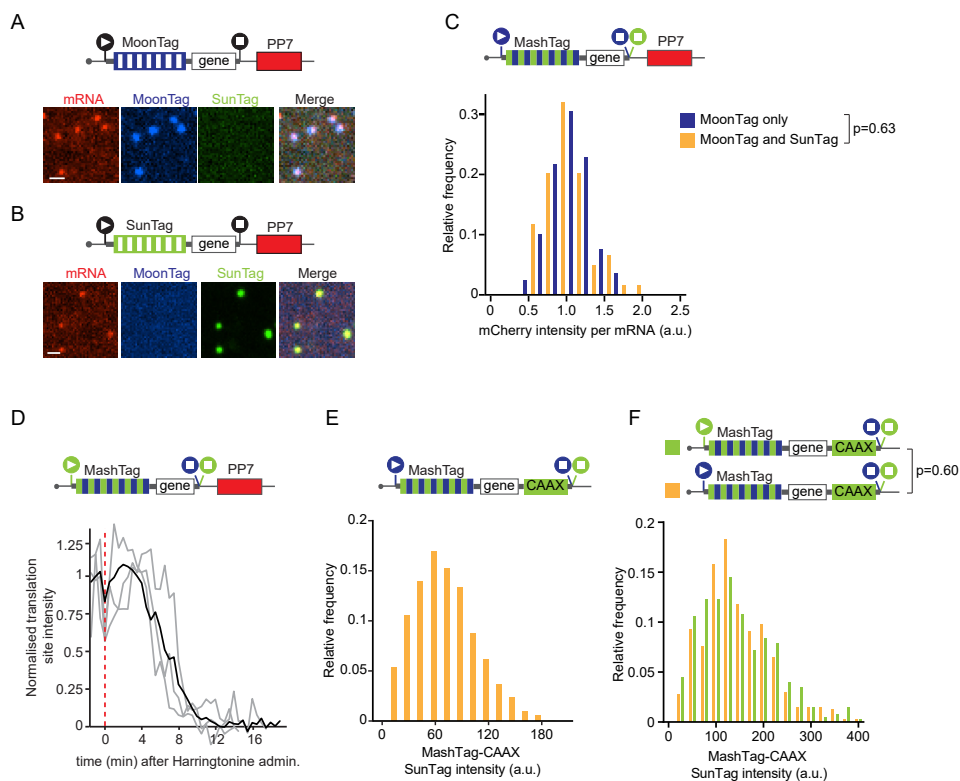
Supplemental figure S2 – related to Fig. 2. Quantifying translation dynamics using the MashTag systems.

A-F) Indicated reporters were expressed in Moon/Sun cells. A-B) Representative images of mRNAs in Moon/Sun cells expressing the indicated translation reporters. C) Distribution of mCherry intensities of mRNAs associated with MoonTag signal only (blue bars) or mRNAs associated with both MoonTag and SunTag signal (orange bars). D) Normalized SunTag intensity on mRNAs after harringtonine treatment. Grey lines depict selected single mRNA intensity traces and the black line shows the average of all mRNAs. Red line indicates harringtonine addition. E) Distribution of the intensity of mature proteins expressed from the SunTag frame. Mature proteins were tethered to the plasma membrane through a CAAX motif. Note that identical imaging settings were used to measure mature protein intensities plotted in (E) and translation site fluorescence intensity traces. F) Distribution of the intensity of mature proteins expressed in the SunTag frame. SunTag proteins were expressed either from the main reading frame (green) or as OOF translation protein products (orange). Mature proteins are tethered to the membrane through a CAAX domain encoded in the SunTag frame. P-values are based on a two-tailed Student's T-Test. Number of experimental repeats and mRNAs analyzed per experiment are listed in Table S1. Scale bars, 1µm.
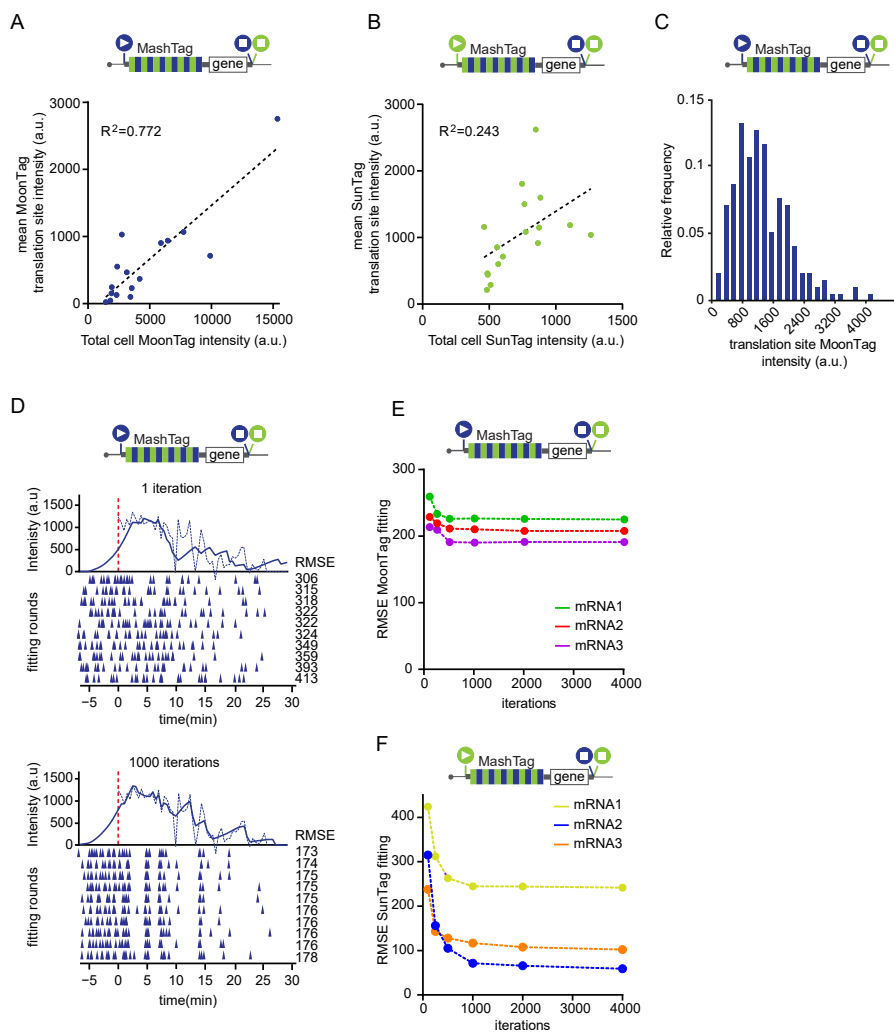
Supplemental figure S3 – related to Fig. 3. Normalizing and quantifying MashTag fluorescence.

A-F) Schematics of translation reporters (top). For simplicity, 24xPP7 sites in the 3'UTR are not depicted. Moon/Sun cells were transfected with indicated reporters. A) Correlation between total cell MoonTag intensity and average MoonTag translation signal on mRNAs. B) Correlation between total cell SunTag intensity and average SunTag translation signal on mRNAs. C) Distribution of the intensity of the MoonTag translation signal of individual mRNAs. D) An example intensity track (dashed line) and fit (solid line) are shown for the MoonTag signal after one iteration (top) or 1000 iterations (bottom) of fit optimization. Colored triangles below the x-axes represent translation initiation events. Each row of triangles illustrates an independent round of fitting. Corresponding root mean squared error (RMSE) values for each round of fitting are shown. (E-F) RMSEs after indicated number of iterations of fit optimization is shown for three representative mRNAs. Number of experimental repeats and mRNAs analyzed per experiment are listed in Table S1.
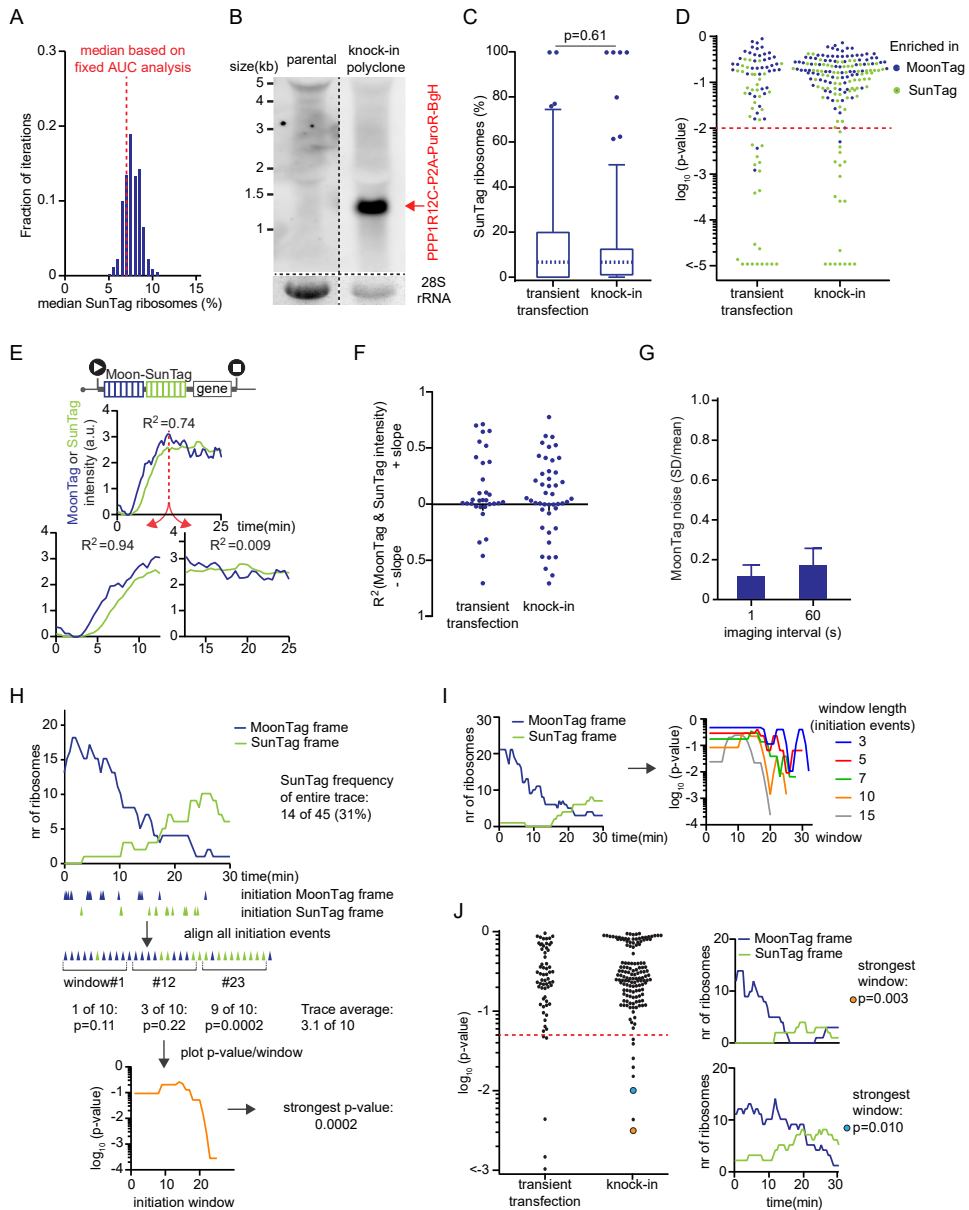
Supplemental figure S4 – related to Fig. 4. Heterogeneity in translation start site selection among different mRNA molecules expressed from a single genomic locus.

The reporter indicated in Fig. 4A was expressed in Moon/Sun cells (A-D, F-J). A) The median frequency of ribosomes translating the SunTag reading frame as determined by either the variable (histogram) or constant (red dashed line) AUC fitting approach (See STAR methods). B) Northern blot with probes against the BGH polyadenylation sequence. RNA was extracted from either parental Moon/Sun cells or from a polyclonal Moon/Sun cell line in which the targeting construct [P2A-Puro-BGH]-[TetOn-MoonStart-MashTag reporter] was integrated into the AAVS1 safe harbor locus (in the PPP1R12C gene). The indicated band represents an mRNA encoding the 5' part of

Figure S4 continued:

PPP1R12C fused to P2A-puro-BGH. 28S rRNA is shown as a loading control (bottom). C, D, F, J) Comparison between MoonStart-MashTag mRNAs either expressed from transiently transfected plasmids (replotted from Fig. 4A, C, E, F) or expressed from the AAVS1 genomic locus. C) Boxplot indicates the relative percentage of ribosomes translating the SunTag frame on single mRNAs. Dashed line represents median value, box indicates 25-75% range, and whiskers indicated 5-95% range. P-value is based on two-tailed Mann-Whitney test. D) P-values for enrichment of ribosomes translating either the SunTag or MoonTag frame on individual mRNAs. Every dot represents a single mRNA. The color of the dot indicates the reading frame that is enriched. Dashed red line indicates p = 0.01. E) Example intensity traces of MoonTag and SunTag on an mRNA of the Moon-SunTag reporter, which illustrates why the linear regression analysis might underrepresent fluorescence intensity correlation over time between MoonTag and SunTag signals; 1) the MoonTag signal appears slightly earlier than the SunTag signal on the Moon-SunTag reporter (left bottom graph) due to its upstream position in the reporter; 2) a poor correlation is obtained if both signals remain largely constant over time (right bottom graph). R^2 values are shown for each graph, as determined by linear regression analysis. F) Linear regression analysis of MoonTag and SunTag intensities of many mRNAs. Each dot represents a single mRNA molecule. G) Cells were treated with 200 µg/ml cycloheximide for 1 min and imaged at indicated time-intervals. MoonTag fluorescence intensity was measured over time. Mean and SD of the intensity was calculated for 6 consecutive time points and the mean/SD is plotted. H) Schematic of sliding window analysis approach. First, ribosomes were fit to raw intensity traces, and the time of each translation initiation event was determined (indicated by triangles under the x-axis of top graph), as described in Fig. 3B (top). Initiation events in both MoonTag and SunTag frames were then merged onto a single time-line. For each consecutive set of initiation events (window length of 10 initiation events is shown), a p-value was calculated using a binomial test, which represents the probability of observing the ratio between MoonTag and SunTag initiation events within that window, based on the MoonTag and SunTag translation initiation rates of the entire mRNA trace (middle). The p-value for each window of 10 consecutive initiation events of an mRNA was plotted and the strongest p-value per mRNA was determined (bottom). The mRNA shown here was also used as an example in Fig. 4F. I) Example graphs showing the number of ribosomes in each reading frame over time for an example mRNAs (left panel) and corresponding sliding window p-value graphs (right) for sliding windows with indicated number of initiation events per window. J) Sliding window analysis of initiation events in MoonTag and SunTag reading frames. Every dot depicts the strongest p-value of a single mRNA (left graph). Example traces of the number of ribosomes in each reading frame over time (right graphs) with corresponding p-values (colored dots). Dashed red line indicates p = 0.05. Number of experimental repeats and mRNAs analyzed per experiment are listed in Table S1.
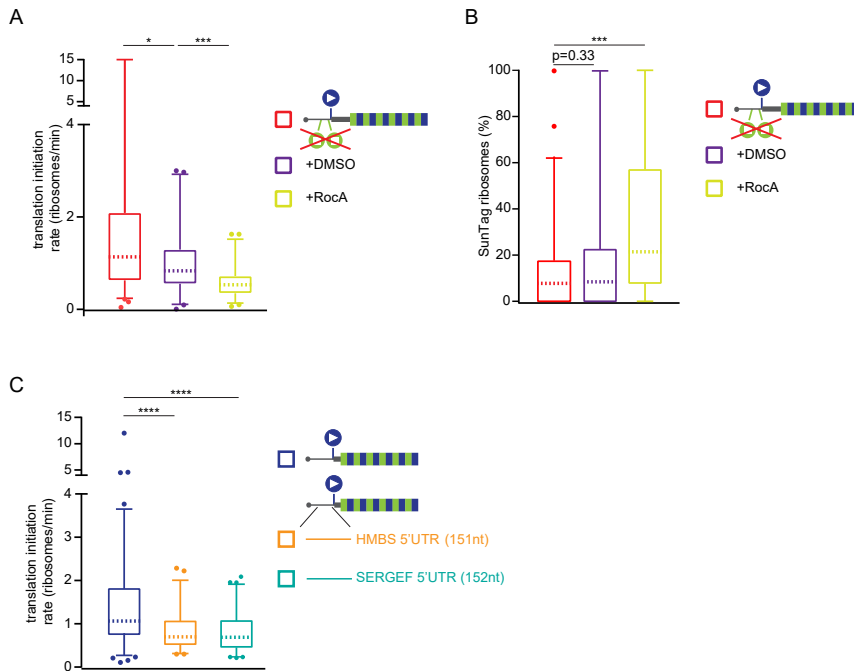
Supplemental figure S5 –related to Fig. 5. Translation initiation dynamics on 5'UTRs.

A-C) Schematics of translation reporters (right). For simplicity, only the 5' region of the mRNA is shown. Indicated reporter mRNAs were expressed in Moon/Sun cells. Boxplots represent the overall translation initiation rates (i.e. initiation rates of MoonTag and SunTag frames combined) (A, C) or fraction of ribosomes translating the SunTag frame (B) for single mRNAs. P-values are based on two-tailed Mann-Whitney tests; * p < 0.05; ** p < 0.01; *** p < 0.001; **** p < 0.0001. Dashed line represents median value, box indicates 25-75% range, and whiskers indicated 5-95% range. Number of experimental repeats and mRNAs analyzed per experiment are listed in Table S1.
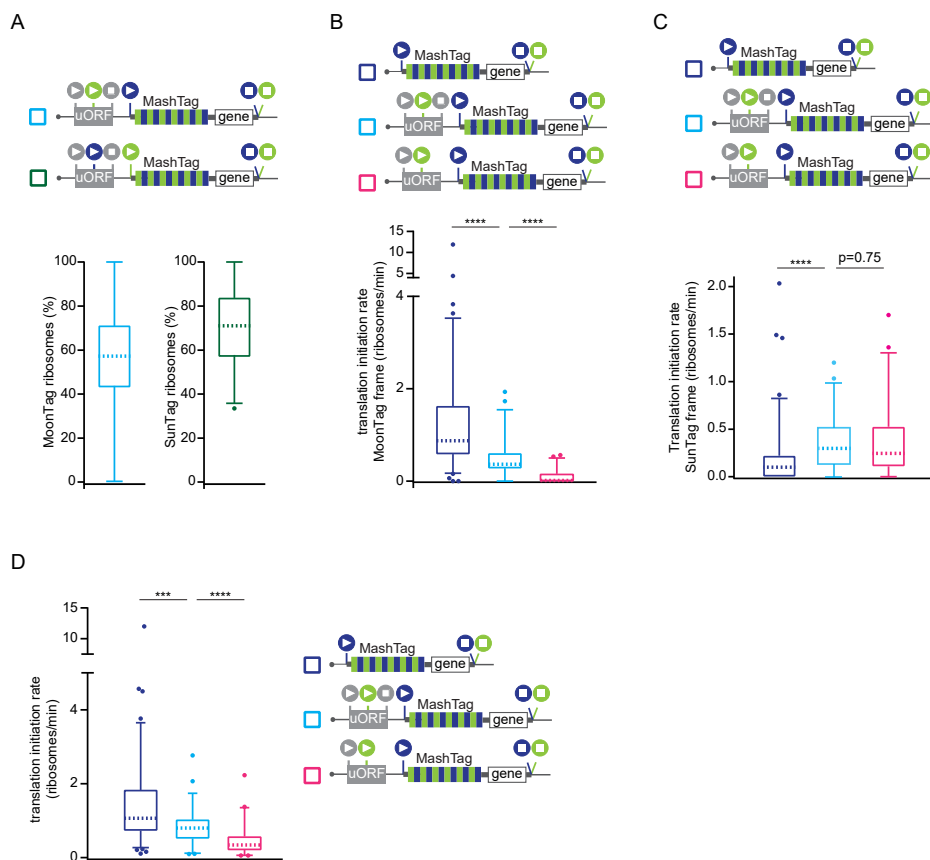
Supplemental figure S6 - related to Fig. 6. Translation initiation dynamics on a uORF-containing mRNA.

A-D) Schematics of translation reporters (top). For simplicity, 24xPP7 sites in the 3'UTR are not depicted. Indicated reporter mRNAs were expressed in Moon/Sun cells. A) Boxplots of relative initiation frequency in the MoonTag frame (left) or SwunTag frame (right) (relative to the sum of the MoonTag and SunTag frame) on single mRNAs. B-D) Boxplots of translation initiation rates in the MoonTag frame (B), SunTag frame (C), or MoonTag and SunTag frame combined (D) for single mRNAs. For comparison, data indicated in dark blue in D is replotted from Fig. S5C. P-values are based on two-tailed Mann-Whitney tests: *** p < 0.001; **** p < 0.0001. Dashed line represents median value, box indicates 25-75% range, and whiskers indicated 5-95% range. Number of experimental repeats and mRNAs analyzed per experiment are listed in Table S1.
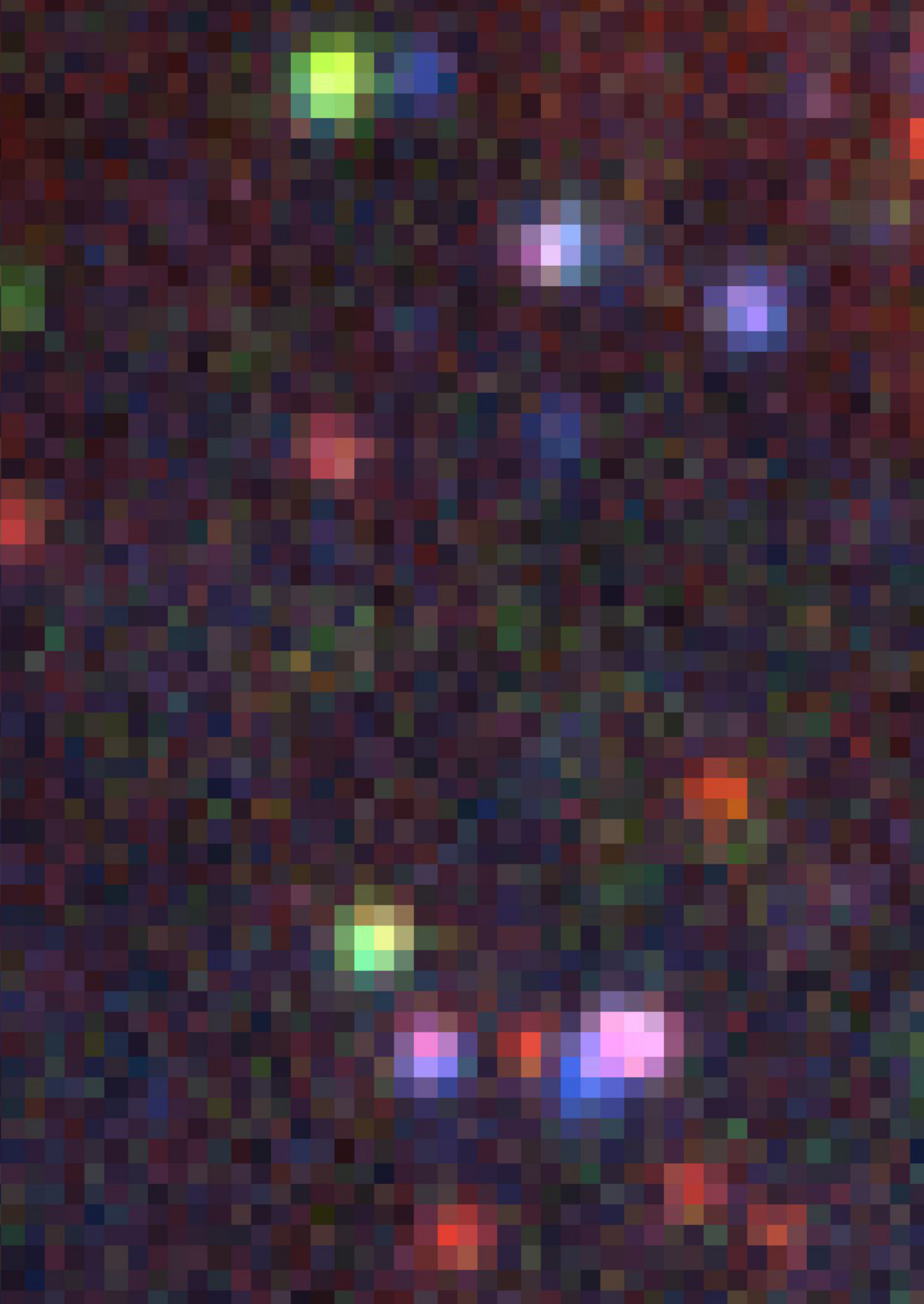
Chapter 3

Effect of Eukaryotic Initiation Factors 4E and 4G Dynamics on Translation Initiation

Sanne Boersma[1], Tim A. Hoek[1], and Marvin E. Tanenbaum[1]

[1] Oncode Institute, Hubrecht Institute–KNAW and University Medical Center Utrecht, Utrecht, The Netherlands.

Background illustration: example image of uORF-MashTag reporter mRNAs
Related to: Chapter 2 - Fig. 6C

SUMMARY

mRNA translation by ribosomes is an essential step during gene expression and depends on the recruitment of ribosomes by eukaryotic initiation factors (eIFs). Live-cell single-molecule imaging studies have uncovered that mRNA translation is heterogeneous both between single mRNA of the same gene as well as over time on a single mRNA. Here, we analyze the dynamics translation initiation, specifically the contribution of eIF4E and eIF4G dynamics to translation initiation. We reveal that multiple ribosomes can be recruited to an mRNA by a single eIF4E and/or eIF4E-eIF4G binding to the cap. Furthermore, we uncover that the durations of the interactions between a cap and an eIF4E molecule and eIF4E and eIF4G molecules has a broad distribution, which may contribute to the temporal heterogeneity of single mRNA translation.

INTRODUCTION

Ribosomes translate mRNAs to produce proteins, and regulation of translation is vital for controlling protein levels in cells (Hinnebusch et al., 2016; Schwanhäusser et al., 2011). Most regulation is exerted at the level of translation initiation, which is further subdivided into several steps (Hinnebusch et al., 2016; Sonenberg and Hinnebusch, 2009): 1) recruitment of a small ribosomal subunit (40S) to the 5' end cap of an mRNA, 2) scanning of the 5'UTR by the small ribosomal subunit, 3) identifying the start of the coding sequence, and 4) assembly of a full (80S) ribosome to then translate the coding sequence (See Introduction of this thesis for more details). (Note that, for simplicity, all compositions of the small ribosomal subunit (40S, 43S, 48S) are referred to as small ribosomal subunit throughout this manuscript). These steps depend on a large collection of proteins called eukaryotic initiation factors (eIFs). The factors eIF4E and eIF4G are of particular interest for the first step (i.e. recruiting the small ribosomal subunit). Eukaryotic mRNA molecules contain a cap structure at the 5' end, which can be bound by the cap-binding protein eIF4E. eIF4G serves as a scaffold between eIF4E and eIF3-interacting small ribosomal subunit, thereby establishing the key link between an mRNA and a ribosome (via cap-eIF4E-eIF4G) necessary to ensure deposition of a small ribosomal subunit on the mRNA.

The link that enables recruitment of the small ribosomal subunit to the cap of an mRNA is subjected to regulation through the mTOR pathway (Gingras et al., 1999). If mTOR is inactive, eIF4E-binding proteins (4E-BPs) bind to eIF4E in the eIF4G binding pocket, thereby preventing eIF4E-eIF4G binding and thus translation initiation. Activated mTOR can phosphorylate 4E-BP, thereby preventing 4E-BP binding to eIF4E and thus supporting translation initiation (Peter et al., 2015; Thoreen et al., 2012). Taken together, eIF4E, eIF4G, and 4E-BPs are key proteins in the regulation of mRNA translation initiation.

We and others have recently developed live-cell single-molecule imaging techniques using SunTag and orthogonal systems to analyze translation of individual mRNA molecules by visualizing nascent polypeptides during translation (Boersma et al., 2019; Lyon et al., 2019; Morisaki et al., 2016; Pichon et al., 2016; Wang et al., 2016; Wu et al., 2016; Yan et al., 2016). The SunTag system consists of a stably- expressed GFP-tagged SunTag antibody (scFv-sfGFP; STAb) that can bind to a short peptide (SunTag peptide) (Tanenbaum et al., 2014). During translation of a reporter mRNA

encoding a SunTag peptide array, STAbs binding to the nascent polypeptide results in accumulation of GFP fluorescence on the mRNA. The GFP intensity on the mRNA depends on the number of ribosomes translating the mRNA and can be used to extract quantitative information on the number of ribosomes translating each reporter mRNA molecule at each time-point (Boersma et al., 2019; Khuperkar et al., 2020).

Using SunTag and orthogonal systems, we often observed variations over time in the number of ribosomes translating an mRNA and between single mRNAs from the same gene over time (reviewed in Sonneveld et al., 2020). Additionally, using the reading frame reporter system MashTag we uncovered extensive heterogeneity in the selection of translation initiation sites between mRNAs of the same reporter gene (Boersma et al., 2019). Similar to previous reports, experiments with the MashTag reporter revealed fluctuations in the number of ribosomes translating an mRNA molecule over time. On most mRNA molecules, we observed a strong correlation between the translational activity from the canonical and non-canonical start site. Even though the frequency of alternative start site selection differed widely between single mRNAs, the ratio in translation from the canonical and non-canonical start site over time on single mRNAs was comparable on many mRNAs, suggesting that a process prior to start site selection is responsible for the observed fluctuations in the number of ribosomes translating an mRNA over time.

Although the interactions between the individual eIFs and interaction partners are well established (Hinnebusch et al., 2016; Jackson et al., 2010), less is known about the dynamics of these interactions in live cells. Furthermore it is unclear whether and how ribosome recruitment dynamics may contribute to the variation in the number of ribosomes translating an mRNA over time. To investigate the dynamics of ribosome recruitment to an mRNA, we performed live-cell single-molecule translation analysis of translation initiation. Our results uncover that binding of a single eIF4E to a cap can facilitate translation initiation of multiple ribosomes. Furthermore, we find that binding and unbinding of eIFs to the cap may contribute to the previously observed temporal heterogeneity of translation on single mRNAs.

RESULTS

A single eIF4E-cap binding event can drive translation initiation of multiple ribosomes

To study the dynamics of eIF4E-cap binding, we sought to analyze the effect of inhibiting the interaction between eIF4E and the cap on the translation of single mRNAs. To date, there are no drugs inhibiting the eIF4E-cap binding that are suitable for experiments with live cells. However, non-cell permeable compounds, such as ARCA (anti-reverse cap analog) have high affinity for eIF4E (Cai et al., 1999; Chen et al., 2012; Jemiely et al., 2003; Kocmik et al., 2018) and may thus be promising reagents to test the inhibition of the eIF4E-cap binding if introduced into live cells (Fig. 1A). We therefore designed a set-up to introduce ARCA into cells by micro-injection (See Methods) and then follow the effect of eIF4E inhibition on the translation of single mRNA using the SunTag translation imaging system (Yan et al., 2016). Specifically, we used STAb-U2OS cells that can express SunTag reporter mRNAs which are immobilized on the plasma membrane, enabling long-term analysis of single translating mRNA at microscopy settings compatible with micro-

injections (see Methods). The fluorescence intensity of the individual SunTag foci was used to compare the translation of an mRNA before and after micro-injections.

Prior to ARCA administration, we observed relatively steady fluorescence at most SunTag foci, in agreement with our previous observations that translation of single mRNA molecules is often stable. After ARCA administration, no new SunTag foci appeared and we observed a decline in fluorescence intensity of foci that were already present (Fig. 1B-D), confirming that ARCA inhibits translation initiation. Such declines in fluorescence intensity of translation foci were not detected upon control micro-injections, excluding micro-injection-induced stress as a cause for translation inhibition (Fig. 1D). Decline in fluorescence intensity of translation foci, due to ribosome run-off, was also observed upon treatment with harringtonine (Fig. 1D), a drug that leads to inhibition of initiating ribosomes. The run-off is the consequence of a gradual decrease in the number of ribosomes translating an mRNA; the already initiated ribosomes continue translating and terminating, but no new ribosomes can initiate (Yan et al., 2016).

The average run-off time after harringtonine administration was shorter compared to the run-off after ARCA micro-injection (Fig. 1D). Close inspection of the intensity traces revealed temporal variations in fluorescence decline; although the reduction in fluorescence was observed for almost all mRNAs, the exact timing of fluorescence decline varied between treatments and between single mRNAs (Fig. 1B, C). However, this temporal variation between treatments and between individual traces was diminished when individual intensity traces were aligned at the moment of complete run-off (i.e. the time-point that the SunTag fluorescence signal has completely disappeared) (Fig. 1E, F). Thus, the onset of run-off (i.e. the time-point that the stable fluorescence signal starts to decline) is different between treatments and between individual traces, whereas the duration between the onset of run-off and complete run-off is similar for all traces.

To examine the run-off dynamics of single mRNAs in more detail, we extracted quantitative information on the timing between drug treatment, onset of run-off, and complete run-off for all traces. Each experimentally observed fluorescence intensity trace was fitted to a model run-off trace consisting of two parts: 1) the constant part, and 2) the declining part (Fig. 1G, H; see Methods, (Hoek et al., 2019)). The constant part reflects continuous translation of the mRNA molecule, during which initiation and termination take place at similar rates. As a result, approximately the same number of SunTag nascent peptides is associated with the mRNA molecule at any time-point of the constant part, and the fluorescence intensity does therefore not change over time (Khuperkar et al., 2020; Ruijtenberg et al., 2020; Yan et al., 2016). The declining part represents loss of translation initiation of new ribosomes, causing run-off of the fluorescence signal. The transition moment between the two parts

Figure 1 continued:

G) Model run-off trace consisting of constant and declining part. H) Example of experimentally observed fluorescence intensity trace fitted to the model run-off trace. I) Kaplan-Meier distribution curve of the duration of the constant part of individual traces. Dashed lines illustrate the time-point when 50% of the traces has started to decline. J) Violin plots of the timing of the decline part. Statistic based on two-tailed unpaired Students' t-test. Solid lines and shaded areas represent mean and SEM respectively (D, F, I). The number of experimental repeats, cells, and mRNAs are listed in Table S1.
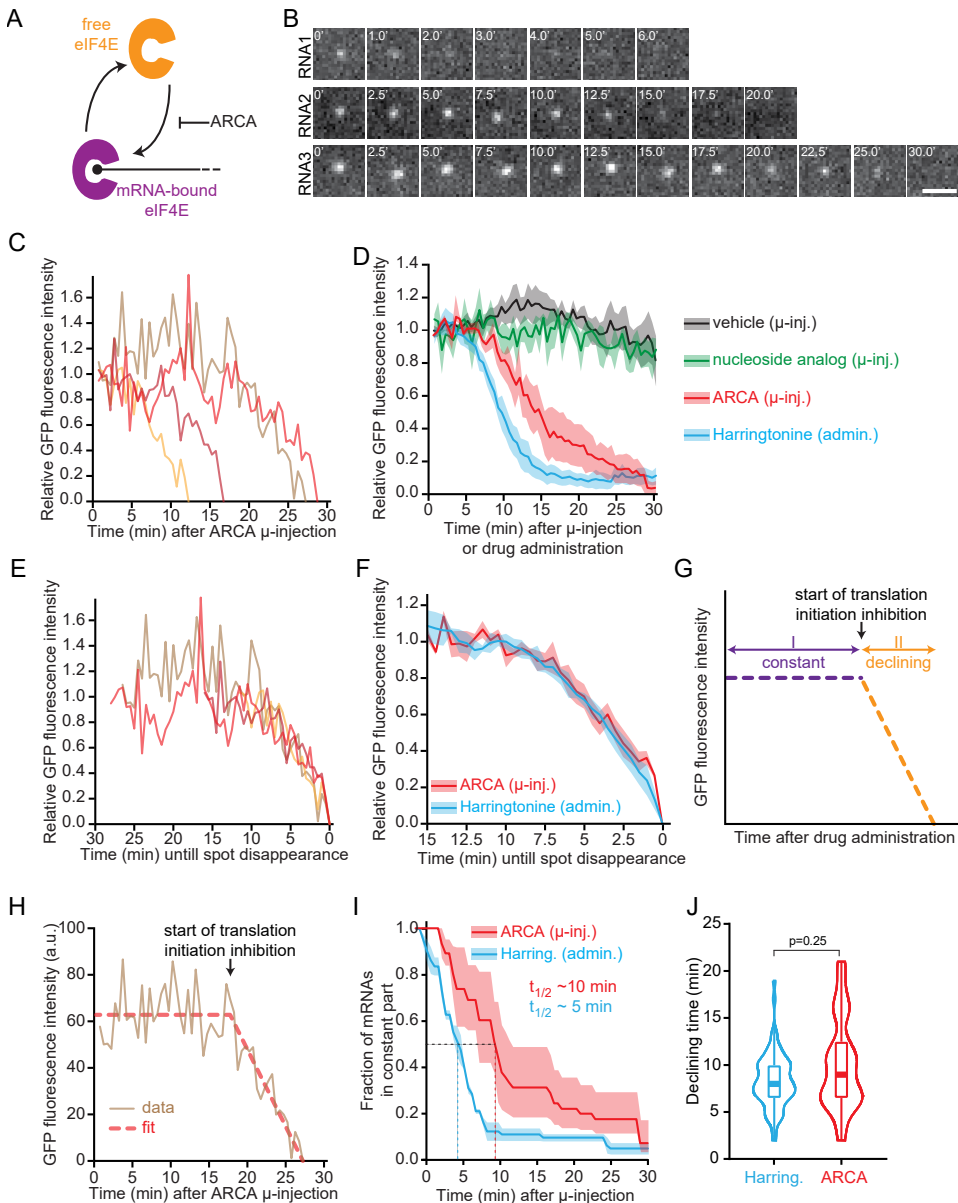
Figure 1. Single-molecule heterogeneity in eIF4E-cap binding.

A) Schematic illustrating ARCA-mediated inhibition of eIF4E-cap binding. B) Example images of 3 SunTag reporter mRNAs expressed in a STAb cell. Time is indicated in minutes after ARCA micro-injection. Scale bar, 3 µm. C) Fluorescence intensity traces of 4 representative mRNAs after ARCA micro-injection. D) Average fluorescence intensity traces of mRNAs after indicated treatment. E) Fluorescence intensity traces of 4 representative mRNAs (same as C) after ARCA micro-injection, aligned to the moment of complete run-off. F) Average fluorescence intensity traces after ARCA micro-injection or after harringtonine administration (same as D), aligned to the moment of complete run-off. Only traces in which moment of complete run-off was observed were included in analysis.

represents the time-point from which onwards no new ribosomes can initiate on the mRNA molecule, due to stalling of the initiating ribosome (due to harringtonine treatment) or due to loss of the cap-bound ribosome recruitment complex (due to ARCA treatment).

The time between drug administration and onset of fluorescence decline was longer upon ARCA treatment compared to harringtonine treatment ($\sim$10 min. vs $\sim$ 5 min.; Fig. 1I), whereas the duration of the run-off is comparable for both treatments (Fig. 1J). These data show that inhibition of eIF4E by ARCA treatment does not instantly induce inhibition of translation on all mRNAs. Instead, mRNAs remain able to recruit new ribosomes for 5-10 minutes after inhibition of eIF4E. Notably, directly from the moment of ARCA micro-injection onwards no new SunTag foci were observed, indicating that ARCA can bind eIF4E rapidly and that the mRNA-to-mRNA variation of the ARCA effect is not caused by variable ARCA-eIF4E binding. We previously estimated the translation initiation rate on these SunTag reporter mRNAs at $\sim$2.5 ribosomes/minute (Yan et al., 2016). Thus, a single eIF4E-cap interaction is stable for several minutes and we can now predict that binding of a single eIF4E to a cap can drive translation initiation of multiple ($\sim$12-25) ribosomes.

Translation initiation inhibition by ha4E-BP1 results in blinking on single mRNAs

To analyze the dynamics of translating initiation in more detail way, we wished to manipulate eIF and ribosome assembly a different manner. For this, we took advantage of the translation regulatory protein 4E-BP, that competes with eIF4G for binding to eIF4E (Fig. 2A; Peter et al., 2015). Specifically, we used a hyperactive variant of 4E-BP1 (ha4E-BP1), that is non-degradable and cannot be phosphorylated by mTOR (Hoek et al., 2019; Yanagiya et al., 2012). Indeed, overexpression of ha4E-BP1 in STAb-U2OS cells with the SunTag reporter mRNAs led to a strong reduction in the number of mRNA molecules on which SunTag translation was observed (Hoek et al., 2019). Furthermore, the translation rate was strongly reduced on the mRNAs that underwent translation (Fig. S1A), confirming that ha4E-BP1 overexpression induces potent translation inhibition.

In the absence of ha4E-BP1, the SunTag fluorescence was stable during most time-points of individual traces (Fig. 2B), in agreement with our previous observations (Yan et al., 2016). However, on some mRNAs we also observed rapid decline of SunTag fluorescence, which was followed by recovery to full SunTag fluorescence intensity levels (Fig. 2C). These repeated cycles of translation shut-off and reinitiation were referred to as 'blinking phenotype' as the SunTag signal on these mRNAs displayed a blinking pattern. Blinking was extremely prevalent upon overexpression of ha4E-BP1 (Fig. 2D), suggesting that the blinking phenotype depends on eIF assembly and cap-recruitment dynamics.

We reasoned that analysis of the blinking phenotype on single mRNAs in cells with overexpression of ha4E-BP1 can be used to study the stability of the link between cap-eIF4E-eIF4G. If the link is intact, ribosomes are recruited to the mRNA, which is observed as translation on the mRNA (i.e. the blink 'on' time). If the link is lost in ha4E-BP1 overexpressing cells, establishment of a new link is hampered by the high levels of ha4E-BP1, which competes with eIF4G for binding to eIF4E, leading to impairment of translation on the mRNA (i.e. the blink 'off' time). Only when a new
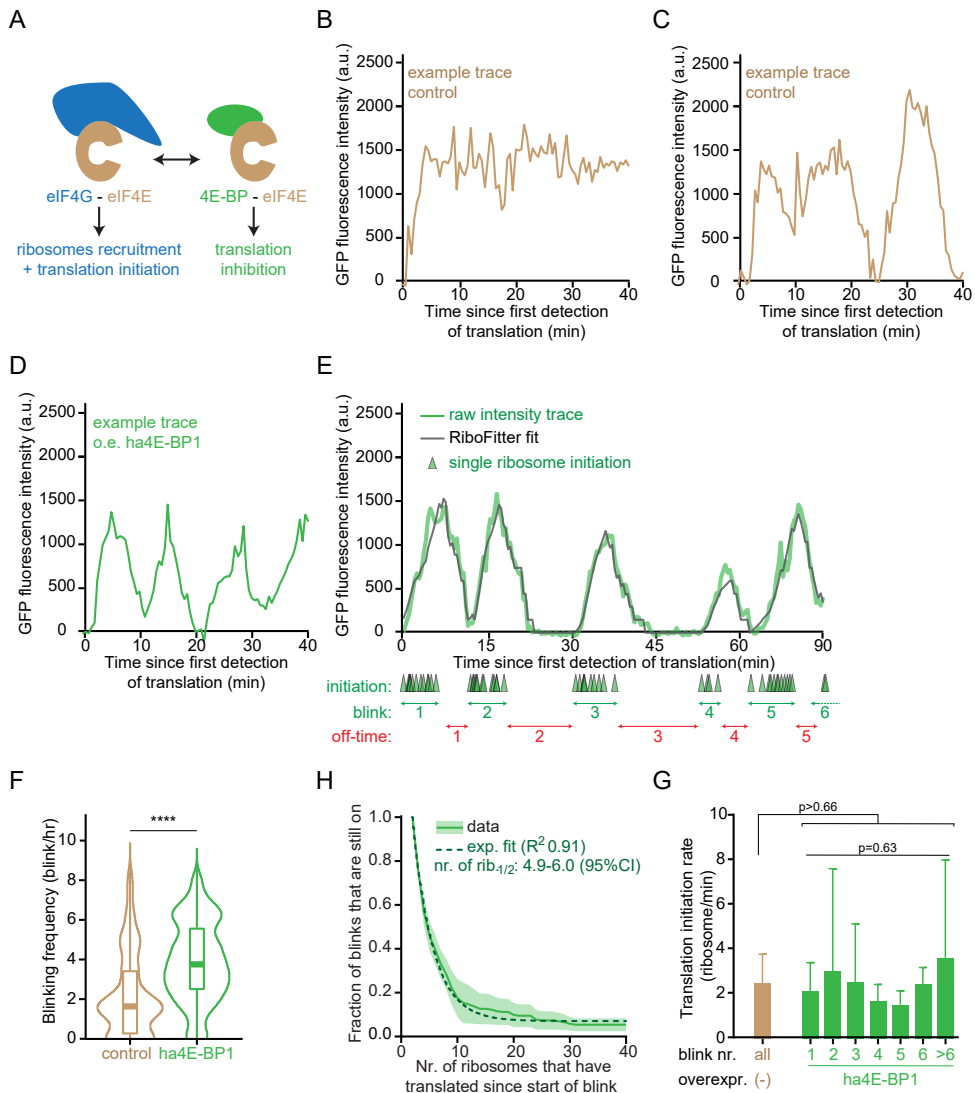
Figure 2. Translation initiation by ha4E-BP1 results in blinking phenotype.

A) Schematic illustrating competition between eIF4G and 4E-BPs for binding to eIF4E. B-D) Fluorescence intensity traces of 3 example SunTag mRNAs expressed in a STAb cell with (D) or without (B, C) ha4E-BP1 overexpression. E) An example fluorescence intensity trace and its RiboFitter fit of a SunTag mRNA in a STAb cell with overexpression of ha4E-BP1. Triangles underneath the x-axis indicate the translation initiation moment of single ribosomes. Green and red arrows at the bottom indicate blinks and off-time. F) Violin plot of the blinking frequency. Statistic based on two-tailed Mann-Whitney test: **** p<0.0001. H) Kaplan-Meier distribution curve of the number of ribosomes per blink. Solid line and shaded area indicate mean and SEM respectively. Dashed line indicates best one-component exponential decay fit. G) Bar graph of the mean translation rate of indicated blinks. Error bars indicate SD. Statistics based on Dunnett's multiple comparison test (control vs. ha4E-BP1) or ANOVA test (1st vs nth blink). The number of experimental repeats, cells, mRNAs, blinks, and/or ribosomes are listed in Table S1.

cap-eIF4E-eIF4G link is formed (i.e. if one eIF4G molecule manages to bind to eIF4E), translation can occur on the mRNA again, leading to the detection of a new blink. Consequently, the duration of a blink reflects how long the cap-eIF4E-eIF4G link was intact.

Our previously established computation pipelines were used to quantitatively investigate the blinking of single mRNAs in more details (TransTrack and RiboFitter, see details in (Boersma et al., 2019; Khuperkar et al., 2020)). In brief, each translating ribosome contributes to the SunTag fluorescence observed on an mRNA and RiboFitter determines the number of ribosomes and the moment of initiation of each ribosome by regenerating the fluorescence intensity trace based on the fluorescence from a single ribosome (called the theoretical single ribosome trace; Fig. S1B, C; see Methods). Based on the RiboFitter outputs of each mRNA, mRNAs with pervasive blinking were characterized by periods of translation initiation and periods without any initiation (see Methods for exact parameters of 'blink' and 'off-time'). Consistent with the visual inspection of the traces, blinking took place more frequently in cells with ha4E-BP1 overexpression compared to control cells (3.8 hr^{-1} vs. 1.6 hr^{-1} (medians); Fig. 2F). Note that the blinking frequency in control cells (1.6 hr^{-1}) is higher compared to the previously determined blinking frequency (0.3 hr^{-1}; (Yan et al., 2016)), as the RiboFitter-based analysis of blinking phenotype is more sensitive compared to the previously performed manual analysis.

Interestingly, no strong differences in initiation rates were observed between individual blinks in ha4E-BP1 overexpressing cells, i.e. during the 'on' times of the blinks (2.0-2.5 ribosomes/min.; Fig. 1G). The translation initiation rate during any blink in ha4E-BP1 overexpression cells was also similar to the translation initiation rate in the control cells (Fig. 1G, S1A). Based on these data, we propose that an mRNA transition between two states: 1) a translation on-state, during which the mRNA molecule is capable of recruiting ribosomes at a constant rate and 2) a translation off-state, during which no translation initiation takes place. Upon ha4E-BP1 overexpression, the translation initiation rate during the on-state is unaffected. Instead, the ha4E-BP1 induces more and longer off-states, resulting in frequent blinking. Per blink, multiple ribosomes were recruited to the mRNAs (Fig. 2H), and blinks lasted approximately 3.5 minutes (Fig. S1E). Thus, a single cap-eIF4E-eIF4G connection is stable for multiple minutes and can recruit multiple small ribosomal subunits and is thus responsible for translation initiation of several consecutive ribosomes.

DISCUSSION
Dynamics of ribosome recruitment

Inhibition of translation initiation with eIF4E inhibitor ARCA revealed that a single eIF4E typically binds to an mRNA molecule for multiple minutes. Similarly, inhibition of eIF4E/eIF4G showed that a single cap-eIF4E-eIF4G link is responsible for ribosome recruitment for several minutes. Collectively, we identify that binding of a single eIF4E and/or eIF4G to a cap can facilitate the recruitment and translation of multiple ribosomes.

Furthermore, we observe variation in the ribosome recruitment dynamics on single mRNA; on some mRNAs the ability to recruit ribosomes is lost rapidly after ARCA

treatment, whereas some other mRNAs remain capable of recruiting mRNA for tens of minutes after ARCA treatment. Likewise, duration of individual blinks in ha4E-BP1 overexpressing cells is variable between individual blinks. We previously reported that translation initiation rate on single mRNAs can fluctuate over time. Taken together, the variations in ribosome recruitment due to dynamics of the link between cap-eIF4E-eIF4G-small ribosomal subunit may contribute to temporal fluctuations as observed by in single-molecule mRNA translation imaging.

Limitations of current study

Interpretation of the current study is complicated by limited understanding of the exact mode of action of the reagents. For example, we used ARCA to inhibit eIF4E-cap binding (Fig. 1A), but it is unclear whether ARCA prevents freely diffusing eIF4E from binding to a cap or whether ARCA could also induce eIF4E release from mRNAs that already had an eIF4E bound prior to ARCA. Regardless of the exact mode of action, single-molecule imaging of mRNA translation upon ARCA treatment can be used to study cap-eIF4E binding. If ARCA only prevents freely diffusing eIF4E from binding to a cap, the time between drug administration and onset of fluorescence decline reflects the stability of the cap-eIF4E interaction. If ARCA can also induce release of an already-bound eIF4E (i.e. if ARCA reduces the cap-eIF4E stability), the time between drug administration and onset of fluorescence decline would result in an underestimation of the stability of the cap-eIF4E interaction. In the current study, we have not (yet) corrected this potential underestimation. We concluded that the interaction between cap and eIF4E is stable for several minutes, but our conservative interpretation may thus underestimate how stable the interaction is.

Similarly, it is unclear how ha4E-BP1 overexpression exactly affects eIF4E-eIF4G binding. For example, recent single-molecule analysis of eIF4E in the presence or absence of 4E-BP1 indicated that 4E-BP1 may induce eIF4E release from the cap (Gandin et al., 2021). Consequently, our conclusion on the stability of the eIF4E-eIF4G interaction based on ha4E-BP1 overexpression is conservative as well. If ha4E-BP1 can interfere with multiple interactions near the cap of an mRNA, the actual stability of the eIF4E-eIF4G interaction may be longer.

Despite limited understanding of the exact mechanism of the reagents, our conservative interpretation indicates that the interactions between cap-eIF4E-eIF4G can facilitate recruitment of multiple small ribosomal subunits.

Future directions

To date, we tested the dynamics of eIFs and ribosome recruitment on mRNAs from one reporter gene in one cell type (U2OS cells). A recent study found that the mechanism of translation initiation may differ between cell types (Bohlen et al., 2020), indicating that analysis of eIF dynamics and ribosome recruitment should be analyzed in various cell types. Several additional factors could impact efficiency of ribosome recruitment or temporal heterogeneity of translation and are therefore prime candidates to be included in follow-up studies. For example, the presence of chemical modifications near the cap or RNA structures in a 5'UTR impact the translation initiation (reviewed in Sonneveld et al., 2020). Moreover, our reporter mRNAs are capped and we therefore only studied eIF4E and eIF4G dynamics on capped mRNAs

and we cannot analyze cap-independent small ribosomal subunit (Lee et al., 2016). Furthermore, mRNAs with 5' terminal oligopyrimidine (TOP) motifs (TOP mRNAs) are particularly sensitive to mTOR-4E-BP1 mediated translation regulation (Jefferies et al., 1994; Patursky-Polischuk et al., 2009; Thoreen et al., 2012). An interesting follow-up experiment could focus on the blinking phenotype on TOP mRNAs.

Likewise, we have not yet explored the contributions of other eIFs in driving temporal fluctuations in translation initiation, like helicase eIF4A, or small ribosomal subunit scaffold eIF3. Inhibition of eIF4A can result in reduced small ribosome recruitment or in stalling during 5'UTR scanning and premature initiation, and the strength of this effect varies between genes (Iwasaki et al., 2016). The large eIF3 complex establishes the connection between eIF4G and the small ribosomal subunit and can remain bound to the small ribosomal subunit during scanning and even during the first translation elongation steps (Archer et al., 2016; Roy et al., 2010; Szamecz et al., 2008), although the exact timing and mechanism of eIF3 release are unknown. It would therefore be interesting to interrogate the involvement of eIF4A and/or eIF3 in ribosome recruitment and in temporal variation in translation initiation.

We recently reported that start site selection fidelity can change between single mRNA of the same gene or over time on the same mRNA (Boersma et al., 2019). It would be interesting to test the effect of ARCA or ha4E-BP1 on MashTag reporter mRNAs to combine analysis of eIF dynamics, small ribosomal subunit recruitment, and start site selection.

Although the exact translation initiation rate and thus the number of ribosomes recruited to an mRNA may differ greatly depending on the conditions and between mRNAs from different genes, the current study shows that the interactions between cap, eIF4E, and eIF4G are stable and that multiple small ribosomal subunits can be recruited by a single cap-eIF4E-eIF4G link. Alternative assays, such as fluorescence correlation spectroscopy and single-particle tracking of single eIFs also provide insight in translation initiation (Gandin et al., 2021). We expect that live-cell single-molecule eIF analysis and translation initiation on single mRNAs is a powerful combination to understand the full complex process of translation initiation, including eIF assembly, ribosomes recruitment, start site selection, and ribosome assembly.

AUTHOR CONTRIBUTIONS

SB: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Visualization, and Writing-Original draft; TAH: Software, Formal analysis; MET: Conceptualization, Supervision and Project management.

METHODS
RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
DMEM	Gibco	Cat# 31966021
Leibovitz's L15 medium	Gibco	Cat# 21083-027

Opti-MEM	Sigma-Aldrich	Cat# 11058-021
Fetal Bovine Serum (FBS)	Sigma-Aldrich	Cat# F7524
Penicillin-Streptomycin	Gibco	Cat# 15140-122
Fugene 6	Promega	Cat# E231A
Cy5-azide	Lumiprobe	Cat# A3030
ARCA	TriLink Biotechnologies	Cat# N-7003
Nucleoside analog – 2'-Deoxyguanosine monohydrate	Sigma-Aldrich	Cat # D7145-25MG
Puromycin	ThermoFisher Scientific	Cat# 12122530
Harringtonine	Cayman Chemical	Cat# 15361
Experimental Models: Cell Lines		
U2OS SunTag translation imaging cells (C19)	Yan et al. 2016	N/A
Recombinant DNA		
pTT5-BFP-NLS-P2A-ha4EBP1	Hoek et al., 2019	pSB140
pHR-NLS-BFP	Tanenbaum lab	#338
Software and Algorithms		
ImageJ	NIH	https://imagej.nih.gov/ij/
Micromanager	Micro-Manager 1.4.22	https://micro-manager.org/
NIS-Elements Imaging software	Nikon	https://www.microscope.healthcare.nikon.com/en_EU/products/software
Graphpad Prism 8	GraphPad Software Inc	http://www.graphpad.com/scientific-software/prism/
MATLAB R2012b	The Mathworks, Inc.	https://nl.mathworks.com/products/matlab.html
R 3.5.1	R Project for Statistical Computing	http://www.r-project.org/
TransTrack (MatLab)	Boersma et al., 2019	https://github.com/TanenbaumLab
RiboFitter (R)	Boersma et al., 2019	https://github.com/TanenbaumLab
Other		
96-well glass bottom imaging plates (Matriplates)	Brooks Life Science Systems	Cat# MGB096-1-2-LG-L
Sterile cell culture dishes with glass bottom 35/10mm (CELLview)	Greiner Bio-One	Cat# 627860
FemtoTips needles	Eppendorf	Cat #930000035

EXPERIMENTAL MODEL AND SUBJECT DETAILS
Cell lines

Human U2OS cells used for imaging were grown in DMEM (4.5 g/L glucose, Gibco) supplemented with 5% fetal bovine serum (Sigma-Aldrich) and 1% penicillin/streptomycin (Gibco) and cells were grown with 5% CO_2 at 37°C. All live-cell imaging experiments were performed using a previously generated U2OS reporter cell line (C19) (Yan et al., 2016). C19 is a monoclonal cell line stably expressing 1) TetR (to enable doxycycline-induced expression), 2) PCP-2xmCherry-CAAX (for labeling and tethering of single mRNA molecules), 3) STAb-sfGFP (for labeling of SunTag peptides). Additionally, C19 cells express reporter mRNAs (24xSunTag-kif18b-24xPP7) upon induction with doxycycline. Of note, the C19 monoclonal cell line has relative low expression of the reporter mRNAs and is therefore suited for relatively long tracing of individual mRNAs. Cells were confirmed to be mycoplasma negative.

METHOD DETAILS
Cell culture and transfections for imaging

For live-cell imaging, C19 cells were seeded at ~40-45% confluency 1 day before imaging in a 35/10mm glass-bottom plate (CELLview, Greiner) or in a 96-well glass-bottom plate (Matriplates, Brooks Life Science Systems) for the micro-injections or blinking pattern experiments respectively. For the blinking pattern experiments, ~300ng of the plasmid expressing ha4E-BP1 or the control plasmids ware transfected using Opti-MEM (Sigma-Aldrich) and Fugene (Promega). Notably, the plasmids were transfected <10 hrs before the live-cell imaging, as prolonged expression of ha4E-BP1 may result in toxicity.

Before imaging, the culture medium on the cells was replaces with imaging medium (pre-warmed CO_2-independent Leibovitz's-15 medium (Gibco) containing 5% fetal bovine serum (Sigma-Aldrich) and 1% penicillin/streptomycin (Gibco)) including doxycycline (1 mg/ml) (Sigma-Aldrich) to induce expression of the reporter mRNAs. Live-cell acquisition was started 30-45 minutes after induction of reporter expression.

Micro-injections during imaging

To examine the effect of non-cell permeable ARCA on translation, ARCA was introduced by micro-injection into the cells.

Micro-injection mixtures

The micro-injection mixtures were prepared in PBS and contained 25mM ARCA (TriLink Biotechnologies) or 25mM nucleoside analog (Sigma-Aldrich). To enable selection of successfully injected cells, a dye (Cy5, 25mM, Lumiprobe) was included in the micro-injection mixtures. As the volume of the micro-injection mixture introduced into a cell upon micro-injection is variable, the exact concentration of the introduced reagents in the injected cells is unknown. To limit the contribution of variable micro-injection volumes to the assay, very high concentrations of the reagents were used in the micro-injection mixtures. For example, the used ARCA concentration (25mM) is several orders of magnitude higher than the ARCA concentration required to inhibit translation in *in vitro* translation assays (K_i = 3.07 ± 0.31 µM, (Jemielity et al., 2003))).

Micro-injection settings

The micro-injection mixtures were loaded into FemtoTips needles (Eppendorf), after which the tip of the needles was removed by gently breaking the needle. The needle was manually positioned just above cells of interest, such that the mixture would be injected into the cells using pre-set injection movements of the needle. Micro-injections were performed at 100-200 µPa and 200 µm/s. Prior to the start of image acquisitions, the micro-injection settings were verified and optimized on several cells to ensure that >70% of selected cells would be injected and cell ruptures were minimal.

Microscopy equipment

Imaging experiments were performed using a Nikon TI inverted microscope with MicroManager software (Edelstein et al., 2010) or NIS Element Software (Nikon), equipped with a temperature-controlled hood, with a perfect focus system (Nikon), with a Yokagawa CSU-X1 spinning disc, with an iXon Ultra 897 EM-CCD camera (Andor), and with a InjectMan4 micro-injection set-up (Epppendorf). Images related to the micro-injection experiments were acquired using a 40x 0.95 NA air-objective and images related to the blinking pattern experiments were acquired using a 100x 1.49 NA oil-immersion objective.

Image acquisition

Cells for imaging were selected ~30 minutes after the induction of the reporter mRNAs. Only cells with >1 and <20 translating mRNA at the start of the time-lapse were selected for imaging. In the blinking pattern experiments, cells were additionally selected based on successful transfection, as observed by expression of nuclear-localized BFP. In all experiments, images of a single Z-plane ware acquired every 30 seconds, focused on the bottom of the cells.

In the blinking pattern experiments, GFP and mCherry images were acquired using 500 ms exposures for 90 minutes.

In the micro-injection experiments, GFP images were acquired using 1000-1250 ms exposure. The images were acquired for 5-10 minutes before treatment with drugs or micro-injections. The acquisition was paused for ~ 90 seconds, to add the translation inhibitor harringtonine (3 µg/ml; Cayman Chemical) to the imaging medium or to perform micro-injections. After drug administration, the time-lapse acquisition was continued for 30 minutes. To test whether fluorescent foci present at the end of an imaging time-lapse represent translation foci, the inhibitor puromycin (0.1 mg/ml; ThermoFischer Scientific) was added to the imaging medium and the acquisition was continued for another 5-10 minutes.

To determine the fluorescence intensity corresponding to the plateau intensity of a single ribosome trace (i.e. nascent chain with 24xSunTag peptides), images of a single time-point were acquired to measure the intensity of mature protein. As mature protein cannot be detected using the relatively long exposure of 500 ms, cells of interest were imaged using 2 settings: A) short exposure (20 ms, and B) normal exposure (500 ms) and regular laser power (6.25-fold less laser power compared to setting A).

QUANTIFICATION AND STATISTICAL ANALYSIS
Post-acquisition processing of microscopy data

To correct for photobleaching, either the Fiji plugin Photobleaching or the photobleaching option in TransTrack were used. Both applications perform the bleaching correction based on the total fluorescence in the field of view at each time-point. The bleaching rate is determined by fitting the fluorescence over time with an exponential decay distribution. If the decay is >10% during the entire time-lapse, the bleaching rate is used to correct each time-point.

Measuring fluorescence intensities of translation sites of micro-injection experiments

Only cells that were successfully micro-injected were included in the analysis based on observing the dye (Cy5) in the injected cell and survival of the infected cell. Cells that looked stressed or died within ~45 minutes after injection were excluded from analysis.

To determine the effect of the drugs on translation of individual mRNA molecules, single mRNAs were tracked before and after drug treatment. Only translating mRNAs that were observed > 5 minutes before drug administration were included in the analysis. In contrast to our previous experiments using the SunTag-kif18b-24xPP7 translation reporters, the mCherry signal corresponding the PP7 structures in the 3'UTRs of the reporter mRNAs can not be used to observe and track individual mRNAs, as the signal is not bright enough to detect an mRNA using the 40x air objective. However, tethering of mRNAs to the membrane can be used to identify a translation mRNA based on the immobility of a GFP spot. To exclude translation-independent aggregates or autofluorescent foci, only GFP foci that either ran-off after drug treatment or were sensitive to puromycin treatment at the end of the time-lapse were tracked and measured.

The fluorescence intensity of single mRNAs was measured manually in an ROI of 6x6 pixels centered on the GFP spot. For each spot, the intensity was background corrected based on the fluorescence intensity of the whole cell. If a spot could not be detected for a single time-point, the intensity was determined at the same location as the last detected spot. If the spot was absent for >1 time-point (indicating complete run-off), tracking of that spot was stopped and the intensity value 0 a.u. was added manually to the intensity trace of that spot till the end of the time-lapse.

To compare different datasets, each trace was normalized to the fluorescence intensity 1.5 minutes before drug administration.

Measuring fluorescence intensities of translation sites blinking pattern experiments

Tracking of individual mRNAs, measuring their GFP fluorescence intensity, and correcting for background fluorescence was performed using the software package TransTrack (Boersma et al., 2019; Khuperkar et al., 2020). Tracking was done based on the mRNA-based mCherry signal. Maximally 10 mRNAs per cell were randomly selected and included in the analysis.

Calculating drug-induced translation run-off kinetics

The GFP fluorescence intensity traces of single mRNAs upon ARCA or harringtonine-induced translation run-off consisted of two parts: 1) a constant part, during which the fluorescence intensity remained stable, and 2) a declining part, during which the fluorescence intensity gradually reduced to zero (Fig. 1G). The first part (the constant part) represents stable translation and indicates that new ribosomes are still initiating translation. The second part (the declining part) represent loss of translation initiation; ribosomes that had already initiated will complete translation, but not new ribosomes are initiating translation. The transition between the two parts reflects the moment that inhibition of translation initiation is started. For some mRNAs, no declining part was observed in the trace, and these mRNAs are

excluded from further analysis. To calculate the duration of the two parts and thus determine the timing between drug administration and starting of translation initiation inhibition, we fitted the intensity traces to a model. The model is based on previous fitting of comparable intensity traces and consists of a constant plateau and a linear decay (Hoek et al., 2019). Different versions of the model were fit to each trace, in which the transition moment ranged from start to end of the trace. The transition moment that fit best between model and trace was used as the moment that inhibition of translation initiation is started. The constant part was defined as the time between drug administration and the starting of inhibition. The declining part was defined as time between start of translation initiation inhibition and end of the trace.

Analysis of blinking pattern

Interpretation of the blinking pattern of single mRNAs upon ha4E-BP1 expression required determination of the number of ribosomes translating an mRNA at each time-point and parameters to define a blink.

Determine moment of translation initiation of single ribosomes

The previously developed algorithm RiboFitter (Boersma et al., 2019; Khuperkar et al., 2020), was used to determine the number of ribosomes translating an mRNA. In brief, the experimentally observed fluorescence intensity trace is constructed *in silico* by combining theoretical single ribosome traces. In an iterative manner, the best fit between observed and reconstructed trace is identified. The read-out consists of information on the number of ribosomes translating an mRNA at each time-point and the moment of translation initiation of each ribosome.

To use RiboFitter, we needed to determine the theoretical single ribosome trace under the used imaging conditions (Fig. S1B). The theoretical single ribosome trace is defined by 3 parameters: 1) the build-up time, which reflect the time that a ribosome is translating the SunTag peptide array and during which the intensity increases linearly, 2) the plateau time, which reflect the time that a ribosome is translating the coding sequence downstream of the SunTag array and during which the intensity remains constant, and 3) the plateau intensity, which corresponds to the intensity of 24xSunTag repeats. The timing parameters 1 and 2 were calculated using the previously determined elongation speed of 10 nt/sec (Yan et al., 2016), and reporter length and formulas as described previously (Boersma et al., 2019) (Fig. S1B).

The plateau intensity (parameter 3) was determined by measuring the fluorescence intensity of mature (i.e. released from the ribosome) proteins, as described previously (Khuperkar et al., 2020). Both during the plateau phase of translation and after release from the ribosome, the peptide chain contains 24xSunTag peptides, and the intensity of mature proteins is therefore similar to the plateau intensity. As mature proteins can not be observed using the regular acquisition settings, two acquisition settings were used: A) short exposure time with high laser power, and B) regular exposure time with regular laser power. Using setting A, the mature proteins can be observed as freely diffusing non-mRNA bound foci. The intensity of crisp single mature proteins was measured, and background corrected using TransTrack. The absolute intensity of single mature proteins determined using setting A does not

reflect the plateau intensity of single ribosomes using setting B. To correct for the difference between the acquisition settings, we calculated a conversion factor based on the intensity of the same translation sites imaged using both settings. Next, this factor was used to convert the intensity of single mature proteins under setting A to the intensity of single mature proteins corresponding the regular acquisition settings (setting B), which was then used as plateau intensity of a theoretical single ribosome trace (Fig. S1B, C).

Define parameters of single blinks

Single blinks were defined based on a minimal number of ribosomes within a blink and a minimal off-time between blinks using the TransTrack output on the exact moment of initiation of single ribosomes.

Based on visual inspection of traces, a blink is defined as at least 3 consecutive ribosomes translating an mRNA within 2 minutes. Additionally, based on visual inspection of traces with blinks and in accordance with our previous analyses (Hoek et al., 2019), a blink ends if there is a period of > 2 minutes without a new blink. To test whether the 2 minutes threshold led to adequate reporting on the high blinking frequency upon ha4E-BP1 expression, we calculated the Δinitiation time (i.e. timing between initiation of a ribosome and the previously initiated ribosome) and compared the Δinitiation time distribution in control and ha4E-BP1 cells. In accordance with the visual inspection, blinking is very frequent upon ha4E-BP1 expression compared to control (23% vs. 3% of Δinitiation time > 2minutes; Fig. S1D) and can therefore be used as a threshold to distinguish blinks.

Comparison of the 1st vs. the nth blink

Only a selection of mRNAs was included in the comparison of number of ribosomes, the blink duration, and the translation rate of the 1st blink to the 2nd, 3rd, etc. blink. Only mRNAs that landed in the off-state (i.e. no SunTag fluorescence was detected in the first time-point that the mRNA was detected) were included in the analyses.

Statistical analyses and generation of graphs

All statistical analyses and graphs were performed and generated using Prism GraphPad (v8.2.1). Statistical tests are explained in figure legends.

To determine the median number of ribosomes per blink and the median duration of blinks and the off-times, the distributions of the number of ribosomes per blink (Fig. 2H), and the timing of the blinks and off-times (Fig. S1E, G) were fitted to a one-component exponential decay distribution. During blink analysis, a minimal threshold of 3 ribosomes, and 2 minutes off-times were used to annotate blinks. Therefore, the same constraints were used while fitting the decay distributions. The median and 95% confidence intervals were derived from the 50%-value of the decay distribution and the error of the fits.

SUPPLEMENTAL ITEMS

Table S1 – related to all figures. Number of experimental repeats, cells, and spots analyzed per experiment.

Overview of the number of experimental repeats, cells, and mRNAs that were used to generate all the graphs. If relevant, (re)plotting of the same datasets is indicated in the last column.

Fig.	panel	Description	color of dataset	repeats	data-points	units of	(re) plotting in figure
1	1C	single mRNA traces	yellow/orange/red/brown				1E, H
	1D	vehicle	black	5	21, 56	cells, mRNAs	-
	1D	nucleoside analog	green	3	8, 11	cells, mRNAs	-
	1D	ARCA	red	4	10, 35	cells, mRNAs	1F, I, J
	1D	harringtonine	blue	3	12, 66	cells, mRNAs	1F, I, J
	1E	single mRNA traces	yellow/orange/red/brown				1C, H
	1F	ARCA	red	4	10, 27	cells, mRNAs	1D, I, J
	1F	harringtonine	blue	3	12, 61	cells, mRNAs	1D, I, J
	1H	single mRNA trace	brown				1C, E
	1I	ARCA	red	4	10 ,27	cells, mRNAs	1D, F, J
	1I	harringtonine	blue	3	12, 61	cells, mRNAs	1D, F, J
	1J	ARCA	red	4	10, 27	cells, mRNAs	1D, F, I
	1J	harringtonine	blue	3	12, 61	cells, mRNAs	1D, F, I
2	2F	control	brown	2	12, 52	cells, mRNAs	S1A
	2F	ha4E-BP1 o.e.	green	3	24, 71	cells, mRNAs	S1A
	2G	control	brown	2	12, 52, 5789	cells, mRNAs, ribosomes	-
	2G	1st blink	green	3	24, 33, 33	cells, mRNAs, blinks	S1F
	2G	2nd blink	green	3	24, 33, 26	cells, mRNAs, blinks	S1F
	2G	3rd blink	green	3	24, 33, 20	cells, mRNAs, blinks	S1F
	2G	4th blink	green	3	24, 33, 13	cells, mRNAs, blinks	S1F
	2G	5th blink	green	3	24, 33, 8	cells, mRNAs, blinks	-
	2G	6th blink	green	3	24, 33, 3	cells, mRNAs, blinks	-
	2G	> 6th blink	green	3	24, 33, 4	cells, mRNAs, blinks	-
	2H	nr of ribosomes per blink	green	3	24, 71, 205	cells, mRNAs, blinks	S1E, G
S1	S1A	control	brown	2	12, 52	cells, mRNAs	-
	S1A	ha4E-BP1 o.e.	green	3	24, 71	cells, mRNAs	-
	S1C	single mature proteins sites to calculate conversion factor	grey not plotted	3 3	36, 447 31, 72	cells, proteins cells, sites	-
	S1D	control	brown	2	12, 52, 5789	cells, mRNAs, ribosomes	-

	S1D	ha4E-BP1 o.e.	green	3	24, 71, 2647	cells, mRNAs, ribo-somes	-
	S1E	ha4E-BP1 o.e.	green	3	24, 71, 205	cells, mRNAs, blinks	2H, S1G
	S1F	duration 1st blink	purple	3	24, 33, 33	cells, mRNAs, blinks	2G
	S1F	duration 2nd blink	magenta	3	24, 33, 26	cells, mRNAs, blinks	2G
	S1F	duration 3rd blink	grey	3	24, 33, 20	cells, mRNAs, blinks	2G
	S1F	duration 4th blink	orange	3	24, 33, 13	cells, mRNAs, blinks	2G
	S1G	off-time	red	3	24, 71, 205	cells, mRNAs, blinks	2H, S1E

I

II

III

IV

V

VI

&
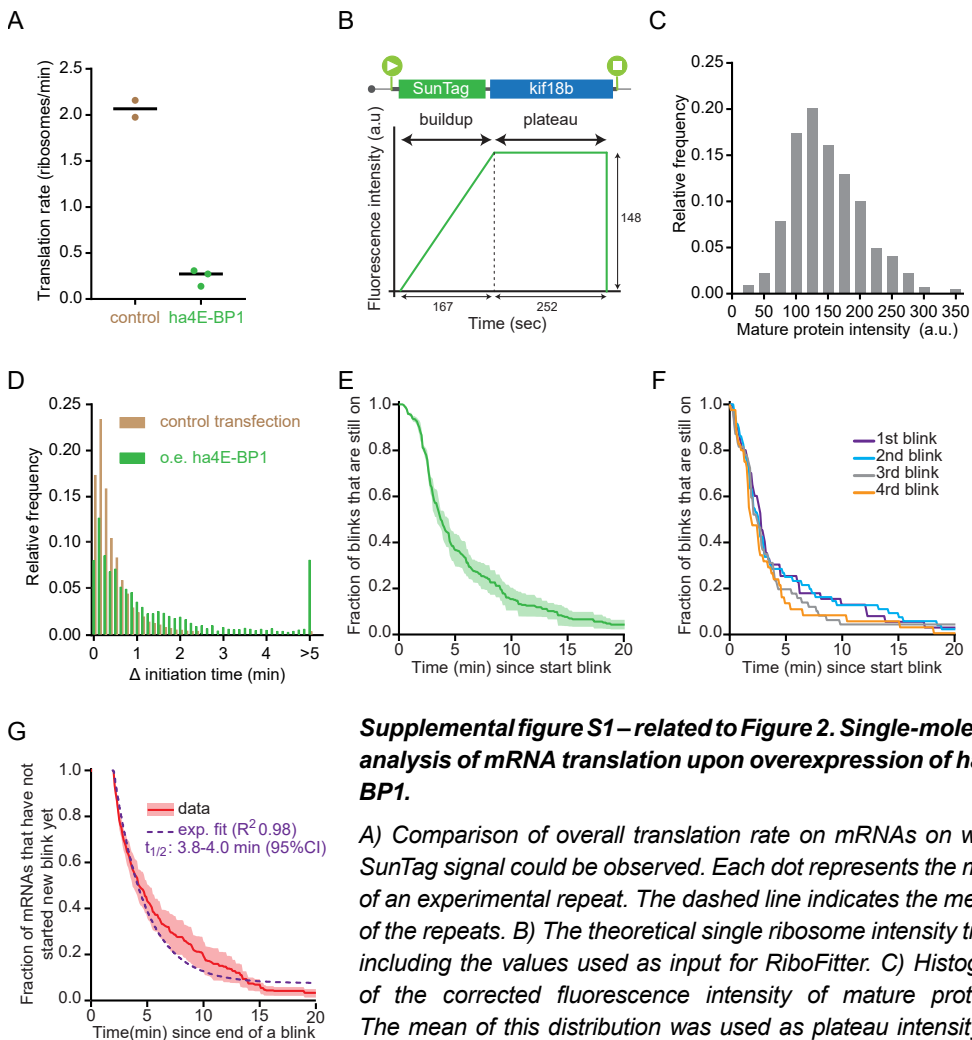
Supplemental figure S1 – related to Figure 2. Single-molecule analysis of mRNA translation upon overexpression of ha4E-BP1.

A) Comparison of overall translation rate on mRNAs on which SunTag signal could be observed. Each dot represents the mean of an experimental repeat. The dashed line indicates the median of the repeats. B) The theoretical single ribosome intensity trace, including the values used as input for RiboFitter. C) Histogram of the corrected fluorescence intensity of mature proteins. The mean of this distribution was used as plateau intensity. D) Histogram of the time between initiation of a ribosome and the initiation of the previous ribosome, based on RiboFitter. E-G) Kaplan-Meier distribution curves of the duration of all blinks (E), the 1st-4th blink (F), or the off-time between blinks (G) in ha4E-BP1 overexpressing cells. Solid lines and shaded areas indicate mean and SEM respectively (E, G). Dashed line indicates best one-component exponential decay fit (G). The number of experimental repeats, cells, mRNAs, blinks, and/or ribosomes are listed in table S1.
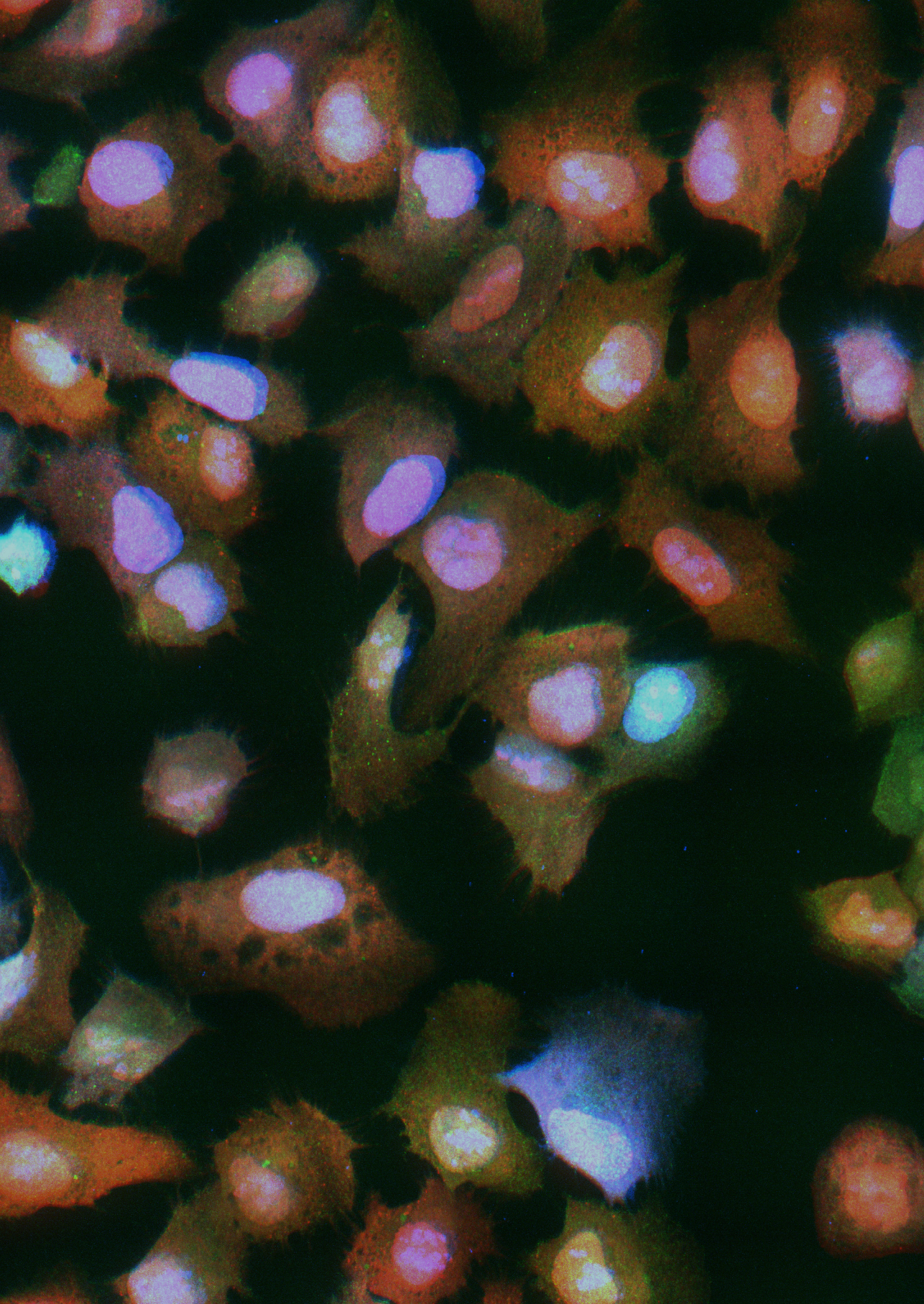
Chapter 4

Translation and Replication Dynamics of Single RNA Viruses

Sanne Boersma[1], Huib H. Rabouw[1,2,$], Lucas J. M. Bruurs[1,$], Tonja Pavlovič[1], Arno L. W. van Vliet[2], Joep Beumer[1], Hans Clevers[1], Frank J. M. van Kuppeveld[2,#], and Marvin E. Tanenbaum[1,#]

[1] Oncode Institute, Hubrecht Institute–KNAW and University Medical Center Utrecht, Utrecht, The Netherlands.

[2] Virology Division, Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, 3584 CL Utrecht, the Netherlands.

[$] Co-second authors

[#] Co-corresponding authors

Background illustration: example image of global translation measurements in infected cells
Related to: Chapter4-Fig. 4A

SUMMARY

RNA viruses are among the most prevalent pathogens and represent a major burden on society. While RNA viruses have been studied extensively, little is known about the processes that occur during the first several hours of infection due to a lack of sensitive assays. Here, we develop a single-molecule imaging assay, virus infection real-time imaging (VIRIM), to study translation and replication of individual RNA viruses in live cells. VIRIM uncovered a striking heterogeneity in replication dynamics between cells and revealed extensive coordination between translation and replication of single viral RNAs. Furthermore, using VIRIM we identify the replication step of the incoming viral RNA as a major bottleneck of successful infection, and identify host genes that are responsible for inhibition of early virus replication. Single-molecule imaging of virus infection represents a powerful tool to study virus replication and virus-host interactions that may be broadly applicable to RNA viruses.

HIGHLIGHTS & GRAPHICAL ABSTRACT

- *Single-molecule imaging assay to study translation and replication of +RNA viruses*

- *Early picornavirus infection occurs in five distinct phases*

- *Heterogeneity in translation and replication of viral RNAs*

- *Replication of incoming vRNA represents major target for host antiviral activity*
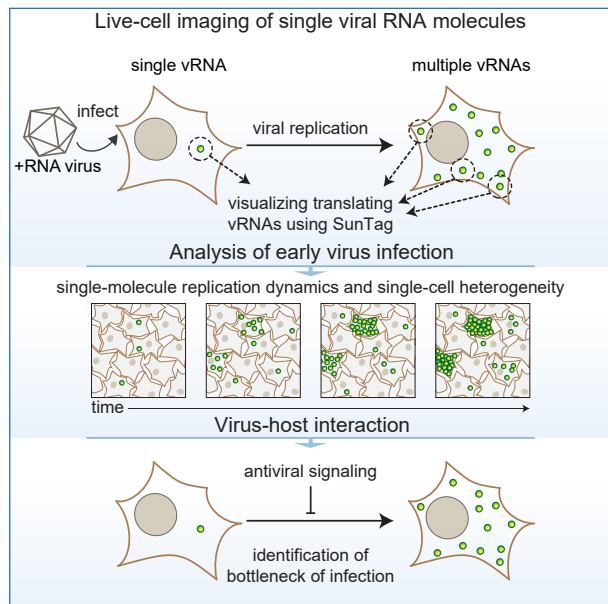
INTRODUCTION

The group of positive-strand RNA (+RNA) viruses comprises many virus families, including important pathogens of humans and animals, like coronaviridae (e.g. MERS-CoV, SARS-CoV2), flaviviridae (e.g. Zika, Dengue, HCV), caliciviridae (e.g. norovirus) and picornaviridae (e.g. poliovirus, coxsackievirus, rhinovirus, and other emerging enteroviruses such as EV-A71 and EV-D68). +RNA viruses cause a major health and economic burden on society, and very few treatment options currently exist for the majority of +RNA virus infections.

Most +RNA viruses contain a single-strand positive-sense RNA genome that can be directly translated into viral proteins upon release into the cytoplasm of a host cell. Upon synthesis, viral proteins execute various functions, such as viral RNA (vRNA) replication, modification of host cell processes to facilitate virus amplification, and repression of antiviral signaling in the host cell. After translation of the incoming vRNA (i.e. the vRNA that infected the host cell), the newly-synthesized viral RNA-dependent RNA polymerase (RdRp) generates negative-sense RNA (-RNA), which in turn is used as template for synthesis of additional +RNAs. These new +RNAs can enter a new round of translation and replication or can be encapsulated to form new infectious virus particles (Baggen et al., 2018; Barrows et al., 2018; Thorne and Goodfellow, 2014; de Wit et al., 2016).

As vRNA molecules can engage in multiple processes (translation, replication, and/or packaging), tightly controlled switching between these dynamic processes is likely important for viral reproduction. Even for the incoming vRNA a translation-to-replication switch is essential to initiate viral replication in newly-infected cells. While some factors have been identified that may contribute to the switch from translation to replication (Ahlquist et al., 2003; Sean et al., 2009), few mechanistic insights into the this switch are currently available.

Cells have sophisticated mechanisms to detect and counteract viral infection, including protein sensors that detect long double-stranded (viral) RNA (dsRNA), which is formed during replication of +RNA viruses. Upon detection of viral dsRNA, host cell signaling leads to rapid activation of innate antiviral pathways, such as the interferon (IFN) induction pathway. Subsequent IFN signaling leads to upregulation of IFN induced genes (ISGs), which are critical to limit the reproduction of the virus (Samuel, 2001; Schoggins et al., 2011; Stetson and Medzhitov, 2006). Viruses, in turn, actively counteract antiviral signaling pathways. For example, many picornaviruses produce proteases that target host dsRNA sensors or members of the IFN signaling cascade to prevent detection of the virus and concomitant production of antiviral signaling molecules (Wang et al., 2018). Moreover, many RNA viruses shut down host translation and transcription, which may also hamper the antiviral response (Chase and Semler, 2012; Walsh et al., 2013). Therefore, early viral detection may improve the likelihood of mounting an effective antiviral response in an infected cell. Thus, the outcome of a viral infection is likely determined by competition between viral translation/replication kinetics and host-cell antiviral signaling kinetics. Interestingly, substantial cell-to-cell heterogeneity has been observed for antiviral signaling, even in a homogeneous population of cells in culture (Doğanay et al., 2017; Patil et al., 2015; Zawatzky et al., 1985), suggesting that cellular and/or viral heterogeneity may be an important aspect of virus-host competition.

Currently available assays are suboptimal to study viral translation and replication

dynamics or virus-host competition for multiple reasons. First, most current assays (e.g. western blot, PCR, immunofluorescence, engineered GFP-expressing viruses, etc) are not sufficiently sensitive to detect the virus during the first few hours of infection, when viral translation and replication, and antiviral responses are initiated. It has been particularly challenging to interrogate the incoming viral particle, as it contains only a single vRNA molecule and thus can easily evade detection. Second, most assays require cell lysis or fixation, and therefore do not provide real-time measurements of live, infected cells. As a result, it is difficult to correlate molecular events that occur early during infection with the eventual outcome of infection. Third, many cells are often required for a single measurement, which is particularly problematic for the analysis of dynamic processes that are not synchronized in time. If multiple cells in a population are infected at different times, an ensemble method is inadequate to study temporally defined events, such as replication of the incoming vRNA or initiation of antiviral signaling. Moreover, the highly heterogeneous response to viral infection represents an additional problem for ensemble methods. Fourth, most assays assess only a single parameter of the viral infection (e.g. vRNA levels or viral protein levels). However, viral translation and replication are interconnected; a translation defect will result in the production of less polymerase, which may reduce the replication rate. Therefore, single parameter assays have a limited ability to uncover mechanistic insights into regulation of either translation or replication specifically. Finally, ensemble methods are also unable to assess spatial information of viral infection. Therefore, new tools are urgently required to provide real-time, spatially resolved single-molecule measurements of viral translation and replication, and virus-host interactions.

RESULTS
Single-molecule analysis of viral translation and replication

To analyze early events during virus infection, we aimed to develop a live-cell imaging assay to visualize individual vRNAs. We applied our previously developed SunTag fluorescence imaging system (Tanenbaum et al., 2014), which allows single-molecule detection in live cells. The SunTag system consists of an array of small peptides (SunTag peptides) and a fluorescently labeled intracellular antibody (scFv-GFP; SunTag antibody (STAb)) that can bind to the SunTag peptides. We and others have previously shown that the SunTag system can be used to visualize translation of single mRNAs (Morisaki et al., 2016; Pichon et al., 2016; Wang et al., 2016; Wu et al., 2016; Yan et al., 2016); when multiple SunTag peptides are introduced at the N-terminus of an open reading frame (ORF), STAbs can bind to the SunTag peptides co-translationally, as soon as they emerge from the ribosome, thus fluorescently labeling the translating mRNA (Fig 1A). As single mRNA molecules are generally translated by multiple ribosomes simultaneously, translating mRNAs are often associated with a high level of SunTag fluorescence, and can easily be distinguished from single 'mature' (i.e. ribosome released) proteins based on fluorescence intensity. In contrast, single-molecule analysis of translation cannot be achieved using GFP encoded by the mRNA; fluorescent proteins like GFP have a relatively long 'maturation time' (i.e. the time between synthesis and fluorescence) (Balleza et al., 2018), so most GFP molecules do not become fluorescent until after translation is

completed and the GFP molecule has been released from the mRNA. In the latter scenario GFP fluorescence is not associated with the mRNAs and does not report directly on the translational status or the localization of individual translating mRNAs.

We reasoned that the SunTag translation imaging system could also be employed to visualize translation of single vRNAs, allowing tracking of viral infections in space and time with single-molecule sensitivity (Fig. 1A). We engineered a coxsackievirus B3 (CVB3), a representative member of the enterovirus genus, with 5 SunTag peptide repeats at the N-terminus of the viral polyprotein (SunTag-CVB3) (Fig. 1A). The SunTag array was stably maintained in the viral RNA through multiple passages (Fig. S1A), although it caused some reduction in overall vRNA levels, similar to other inserts in CVB3, such as GFP (Fig. S1B-D) (Andino et al., 1994; Feuer et al., 2002; Lanke et al., 2009)). Shortly after infection of human U2OS cells stably expressing the STAb (referred to as 'STAb cells') with SunTag-CVB3 at low MOI (MOI=0.25), one or more bright GFP foci could be observed in infected cells (Fig. 1B). Single molecule FISH (smFISH) analysis showed that SunTag GFP foci co-localized with CVB3 +RNA (Fig. 1C), and GFP foci rapidly disappeared upon administration of the translation inhibitor puromycin (Fig. 1D), confirming that GFP foci represent nascent polypeptides associated with translating vRNAs, rather than mature proteins. Quantitative analysis of GFP foci intensities revealed that individual GFP foci correspond to ~90 SunTag peptides (Fig. S1E). As a single ribosome translating the vRNA is associated with only 5 SunTag peptides, these results indicate that GFP foci represent vRNAs simultaneously translated by many ribosomes.

When cells were followed by time-lapse microscopy, we generally observed an increase in the number of GFP foci over time in individual cells (Fig. 1B), indicative of viral replication. The number of GFP foci observed in live-cell imaging experiments correlated well with the number of vRNAs as assessed by single-molecule FISH (smFISH) in the same cells after fixation (R^2 = 0. 90, Fig. 1E, S1F). The increase in GFP foci was strongly attenuated by an inhibitor of the RdRp, $3D^{pol}$ (GPC-N114, (van der Linden et al., 2015)), confirming that an increase in number of GFP foci in single cells reflects vRNA replication (Fig. 1F).

To assess whether SunTag-CVB3 replicated with kinetics similar to wildtype CVB3, we developed a live-cell sensor of infection by CVB3 that does not rely on the SunTag. Previous reports have shown that nuclear pore integrity becomes impaired upon infection with CVB3, resulting in increased exchange of nuclear and cytoplasmic components (Belov et al., 2000; Flather and Semler, 2015; Gustin and Sarnow, 2001). To test whether we could leverage impaired nuclear transport as a marker for infection, we generated STAb cells stably expressing BFP fused to a nuclear localization signal (NLS-BFP). We found that translocation of BFP from the nucleus to the cytoplasm occurs extremely rapidly after infection with SunTag-CVB3, within minutes of translation of the incoming vRNA (Fig. 1G, H, movie S1). The start of BFP translocation can therefore be used as a proxy for the moment of infection. Using live-cell imaging of NLS-BFP localization (to assess the moment of infection) combined with smFISH of the same cells after fixation (to assess viral replication), we compared the number of vRNAs over time for both SunTag and wildtype CVB3. While the number of vRNAs over time was highly heterogeneous between cells for both wildtype and SunTag-CVB3, the average number of vRNAs per cell over time
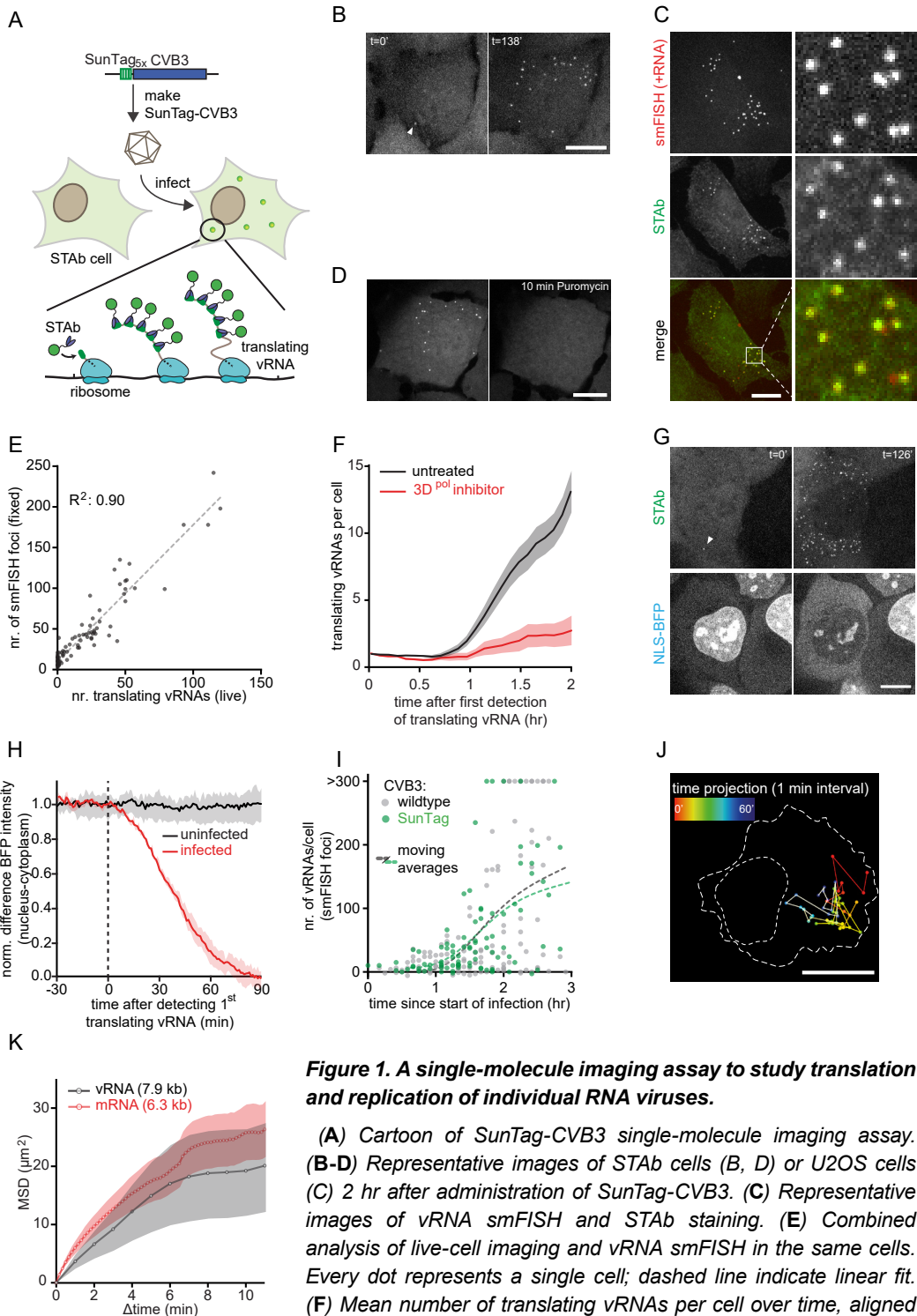
Figure 1. A single-molecule imaging assay to study translation and replication of individual RNA viruses.

(**A**) Cartoon of SunTag-CVB3 single-molecule imaging assay. (**B-D**) Representative images of STAb cells (B, D) or U2OS cells (C) 2 hr after administration of SunTag-CVB3. (**C**) Representative images of vRNA smFISH and STAb staining. (**E**) Combined analysis of live-cell imaging and vRNA smFISH in the same cells. Every dot represents a single cell; dashed line indicate linear fit. (**F**) Mean number of translating vRNAs per cell over time, aligned at first detection of a vRNA, either with or without pre-treated with

Figure 1 continued:

*10 µM 3D^pol inhibitor (GPC-N114). (**G**) Representative images of NLS-BFP STAb cells 1 hr after SunTag-CVB3 administration. (**H**) Difference in BFP fluorescence intensity between nucleoplasm and cytoplasm. Data is aligned at the start of phase1 (dashed line) and normalized to the values of the 3 min before start of phase1. (**I**) Combined analysis of live-cell imaging and vRNA smFISH in the same cells. Every dot represents a single cell. (**J**) Time projection of a single translating vRNA. Color indicates time in min since first detection of the vRNA; dotted lines indicate cell and nuclear outline in first time-point. (**K**) Diffusion kinetics of translating vRNA or mRNA molecules. Time in minutes since first detection of a translating vRNA (arrow head) is given in (B) and (G). Shaded areas in F, H, K indicate SEM. Scale bars, 15 µm. See also Movie S1 and Figure S1. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.*

was similar for both viruses, indicating that insertion of the SunTag into the viral genome does not affect early vRNA replication (Fig. 1I). Furthermore, by combining smFISH with immunofluorescence, we found that the ratio of vRNAs to viral proteins is also similar for wildtype and SunTag CVB3 (Fig. S1H), indicating that the SunTag also does not hamper viral translation.

Quantitative comparison SunTag-CVB3 with previously established methods to detect viral infection (staining of viral dsRNA and fluorescence generated by GFP-CVB3) confirmed that previous methods could reliably detect viral infection only at 3-5 hr after infection (Fig. S1I-L), when the vRNA has already undergone (several rounds of) replication. Together, these results show that the SunTag-CVB3 faithfully recapitulates wildtype viral infection dynamics, and uniquely reports on early events during infection. We refer to this single-molecule virus imaging assay using SunTag as VIRIM (Virus Infection Real-time IMaging).

Localization and mobility of vRNAs during early infection

Using the SunTag-CVB3 we first examined localization and mobility of translating vRNAs. During the first 2-3 hr, we did not observe any preferential localization of translating vRNAs in the cytoplasm. vRNAs moved rapidly throughout the cytoplasm and showed a mobility that was similar to the mobility of host mRNAs of comparable length (Fig. 1J, K, S1M). At later stages during infection (2-4 hr post-infection (p.i.)), we did observe a small subset of vRNAs that became immobilized in the vicinity of the nucleus (Fig. S1N, O). All GFP foci (both mobile and immobile) rapidly disappeared upon treatment with puromycin, (Fig. S2A), indicating that both types of foci represent translating vRNAs. We conclude that during early infection the majority of translating vRNAs are not localized to specific sub-cellular sites, but rather are freely diffusing through the cytoplasm.

Heterogeneity and dynamics of replication

Long-term time-lapse imaging of SunTag-CVB3-infected cells revealed a remarkable recurring pattern in viral replication, including five distinct phases (Fig. 2A-C, movie S1); infection phase1 starts with the appearance of a single GFP spot (at MOI <1), representing translation of the incoming vRNA. Phase1 is followed by a period without GFP foci, referred to as phase2. Phase2 may represent vRNA replication (synthesis of a -RNA and subsequent synthesis of multiple new +RNAs templated from
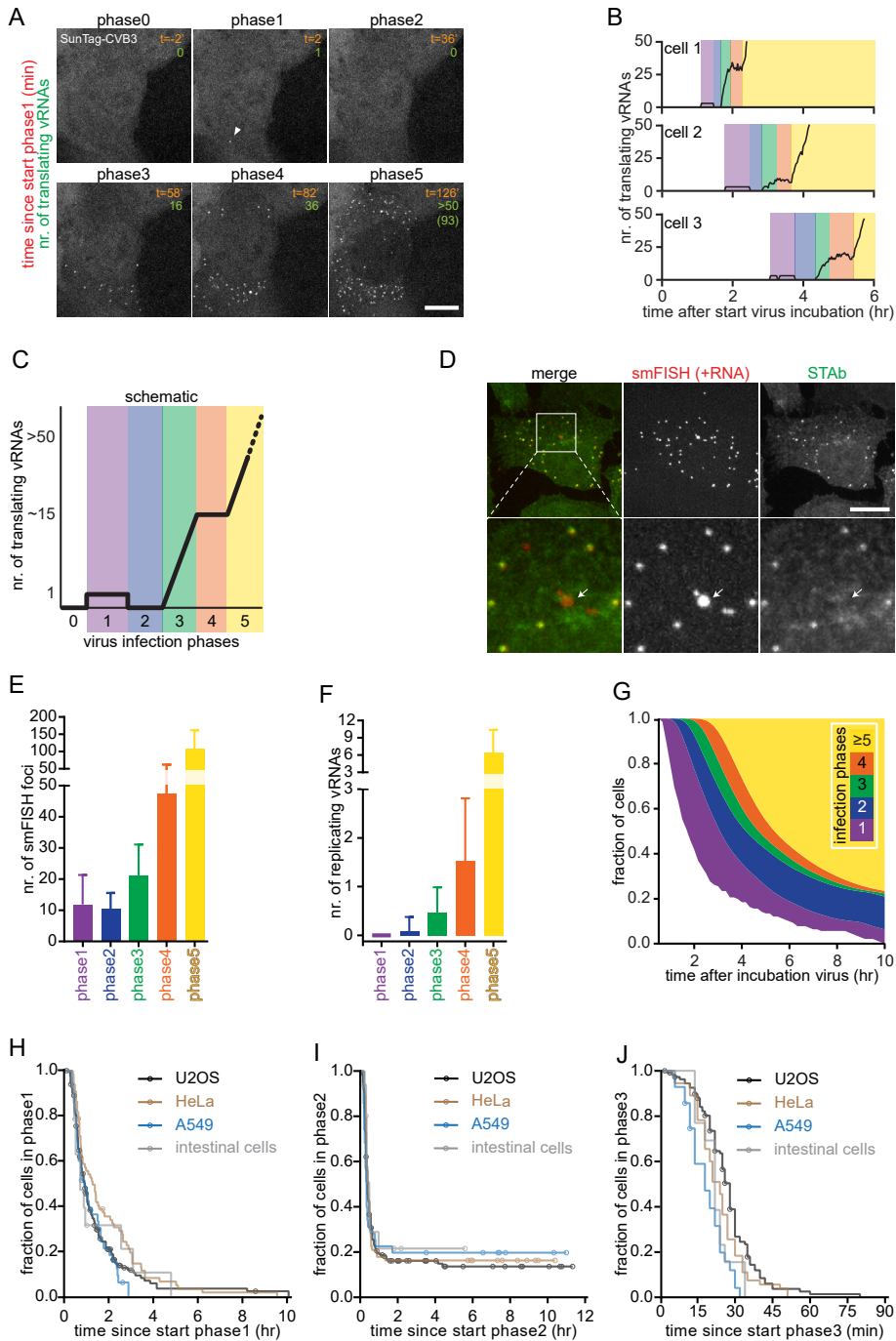
Figure 2. Single-cell dynamics and heterogeneity of virus replication.

(**A, B**) Representative images (A) and example quantifications (B) of time-lapse movies of STAb cells infected with SunTag-CVB3. Example images from the same time-lapse movie are also used in Fig. 1G and Movie S1. (**C**) Cartoon of infection phases in single cells. (**D**) Representative

Figure 2 continued:

I

II

III

IV

V

VI

&
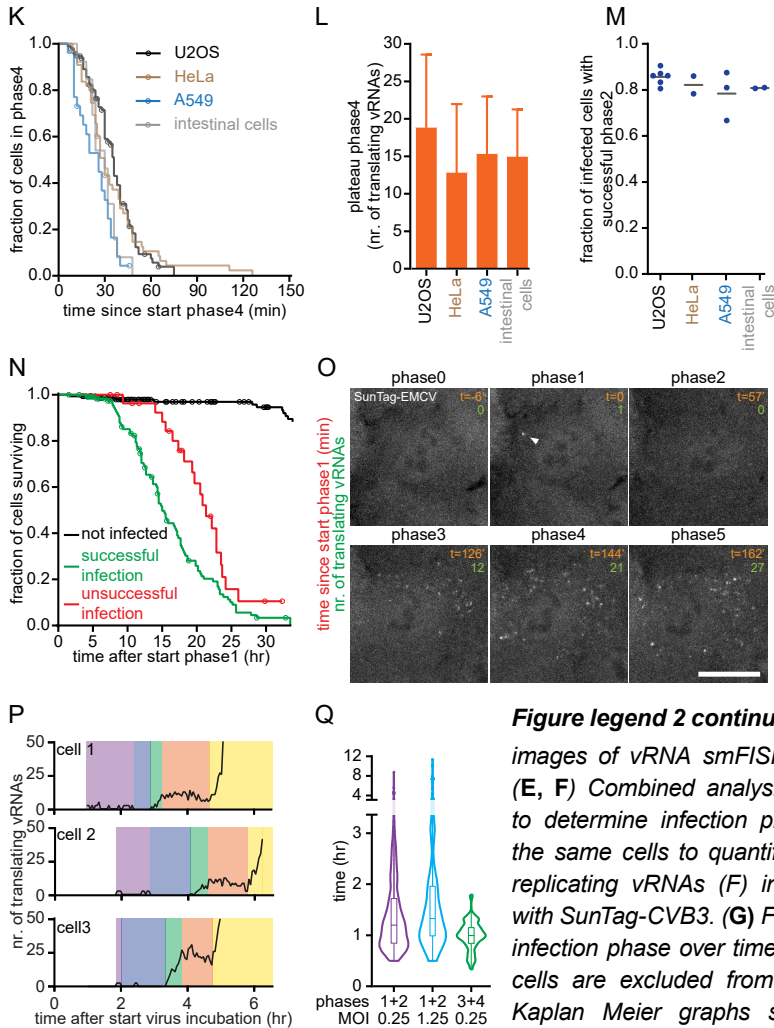
Figure legend 2 continued:

images of vRNA smFISH and STAb staining. (**E, F**) Combined analysis of live-cell imaging to determine infection phase and smFISH in the same cells to quantify smFISH foci (E) or replicating vRNAs (F) in STAb cells infected with SunTag-CVB3. (**G**) Fraction of cells in each infection phase over time. Note that uninfected cells are excluded from quantification. (**H-K**) Kaplan Meier graphs showing durations of infection phases. (**L**) Mean number of translating

vRNAs during phase4. (**M**) Fraction of infected cells with successful phase2, based on plateau in Kaplan Meier curve of the duration of phase2. Every dot indicates a repeat and lines indicate mean. (**N**) Kaplan Meier survival curve. (**O, P**) Representative images (O) and example quantifications (P) of representative time-lapse movies of STAb cells infected with SunTag-EMCV. (**Q**) Violin plot of combined timing of phases1 and 2 or phases3 and 4. Colors (B, C, P) illustrate infection phases. Note that data points during phases 0, 1, and 2 (B, P) are increased 3-fold to aid visualization of data that is very close to the x-axes. Arrow heads (A, E) indicates the first translating vRNA; arrow (D) indicates a replicating vRNA. Error bars (E, F, L) indicate SD;; circles (H-K, N) indicate last analyzable time-point for individual cells. Scale bars, 15 µm. See also Movie S1-3 and Figure S2, S3. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.

the -RNA), as vRNA translation is shut down during replication (Barton et al., 1999; Gamarnik and Andino, 1998). Consistent with this, the average duration of phase2 was in line with the expected time required to synthesize -RNA and +RNA based on *in vitro* measurements of replication speed (Arnold and Cameron, 2000; Barton and Flanegan, 1997). Phase3 starts with the re-appearance of a GFP spot, followed by a rapid increase in the number of GFP foci, likely due to translation of newly synthesized +RNAs. During phase3, additional +RNAs are likely synthesized from the –RNA that was produced in phase2. In phase4, the number of GFP foci remains constant, indicating that the newly synthesized +RNAs are undergoing translation in preparation of a new round of replication. On average 15-20 translating vRNAs were observed during phase4, consistent with a previous observation that +RNAs typically outnumber –RNAs by approximately 20-fold (Dave et al., 2019). During phase4, a subset of vRNAs is expected to undergo replication, yielding new –RNAs and subsequent new +RNAs. In phase5, a second rapid increase in the number of foci is observed, likely reflecting synthesis and translation of the additional +RNAs. Phase5 continues until the number of foci per cell exceeds the detection limit.

Quantitative analysis revealed that foci calling, which underlies annotation of infection phases, was highly accurate with only 0.4% false-positives and 7.5% false-negatives (Fig. S2B). Moreover, GFP foci in all infection phases disappeared upon treatment with the translation inhibitor puromycin, confirming that the GFP foci exclusively represent translating vRNAs (Fig. S2A). Note that at the end of phase5 GFP, which previously localized in both cytoplasm and nucleus, accumulated in the cytoplasm, likely reflecting an excess of mature SunTag protein in the cytoplasm, which sequestered the STAb (Fig. 2A, movie S1). The excess of SunTag peptides over STAb molecules may cause a lower binding stoichiometry of the STAb to SunTag peptide arrays and interfere with quantitative interpretation of GFP foci intensities. However, we found that STAb binding stoichiometry was not affected during early infection, at least until ~60 foci per cell were present (i.e. beyond phase5) (Fig S2C), demonstrating that STAb binding stoichiometry was constant through phases1-5.

To confirm that replication of the incoming vRNA occurred during phase2 and 3, we combined VIRIM with CVB3 +RNA smFISH analysis of the same cells (Fig. S1F, 2D). This analysis revealed that the number of smFISH foci increased starting from phase3 onwards (Fig. 2E), consistent with vRNA replication in phase2 and 3. Interestingly, close inspection of the smFISH foci revealed that two types of smFISH foci could be distinguished in infected cells based on the smFISH intensity (Fig. 2D, S2D); most smFISH foci showed a uniform, low intensity, while a small subset of foci was much brighter (> 2.5 fold). During replication multiple polymerases can simultaneously use the –RNA as a template, resulting in multiple co-localization nascent +RNAs, suggesting that these bright foci may represent replicating vRNAs. The number and fluorescence intensity of the bright foci was indeed reduced in cells treated with the 3D[pol] inhibitor (Fig. S2F-G), and the bright foci rarely co-localized with STAb fluorescence (1.2%, n=329 bright foci, 2 repeats; Fig. 2D), indicating that the bright smFISH foci indeed represent replicating vRNAs. Quantitative analysis revealed that these replicating vRNAs were not detected during phase1, but became visible from phase2 onwards (Fig. 2F), further confirming that replication initiates during phase2.

Although the five infection phases were observed in most infected cells, the duration

of the individual phases, as well as the number of translating vRNAs during each phase was heterogeneous between cells (Fig. 2B). In particular, substantial heterogeneity was observed in the duration of phase1, ranging from 12 min to > 4 hr. As a result, cells in all infection phases co-exist within a population of cells at any moment (Fig. 2G). Such diverse cells are pooled and averaged in ensemble measurements, highlighting the importance of a real-time single-molecule imaging approach, like VIRIM. Similar infection phases were observed in HeLa cells, as well as in cells derived from the airway epithelium (A549 cancer cells) and gastrointestinal tract (primary human intestinal organoids), which represent natural targets of CVB3 (Fig. 2H-M; S2E-M; (Baggen et al., 2018)). Together, these results suggest that the infection phases and their timing may be a universal phenomenon for CVB3.

When analyzing the replication phases in more detail, we found that 15-20% of the SunTag-CVB3 infected cells were arrested during phase2 (Fig. 2I, M, movie S2), indicating that the incoming vRNA did not undergo replication to produce new vRNAs in those cells. In all cases in which phase2 was completed successfully, subsequent phases were also successful, indicating that phase2 is the most vulnerable phase of the viral life cycle that is key for the success of infection. Notably, most cells in which no detectable viral replication occurred eventually still died, albeit slightly slower than cells with regular replication (Fig. 2N; movie S3). Cells containing unsuccessfully replicated virus may play an important role in antiviral immunity *in vivo*, for example through cytokine production or through release of viral antigens, without release of viral progeny.

To determine whether we could similarly assess viral replication kinetics of other viruses, we also employed VIRIM to encephalomyocarditis virus (EMCV), a member of the cardiovirus genus within the picornavirus family (Baggen et al., 2018). SunTag-EMCV was viable and could be stably propagated in human cells. After infection of STAb-U2OS cells with SunTag-EMCV, we observed rapidly diffusing GFP foci, similar as with SunTag-CVB3. SunTag-EMCV foci were on average ~3-fold dimmer than SunTag-CVB3 foci (Fig. S3K). While EMCV GFP foci were more difficult to detect, GFP foci calling could still be performed with high accuracy (Fig. S3L). Dimmer translating vRNA foci suggest that the translation efficiency of EMCV is lower than that of CVB3 (the vRNA is occupied by fewer translating ribosomes). Nonetheless, analysis of the number of translating vRNAs over time revealed that EMCV replication followed a similar pattern as CVB3 replication. (Fig. 2O, P). The infection phases may thus be a general phenomenon of picornaviruses. These results show that VIRIM may be widely applicable to study translation and replication dynamics of +RNA viruses.

The translation-to-replication switch

During phase1, a virus translates its genome in preparation of vRNA replication (phase2/3). Similarly, newly synthesized vRNAs are translated in phase3 and 4 in preparation of replication that occurs during phase4/5. Interestingly, we found that the median duration of phase1+2 (translation and replication of incoming vRNA) was similar as the median duration of phase3+4 (translation and replication of daughter vRNAs) (60 vs. 72 minutes; Fig. 2Q), even though far more viral proteins and vRNAs are present in the cell during phase3+4. These results suggest that the timing of vRNA replication is not determined by the number of vRNAs or viral proteins in the cell, but rather is intrinsic to individual vRNAs (i.e. controlled in cis). Consistent with
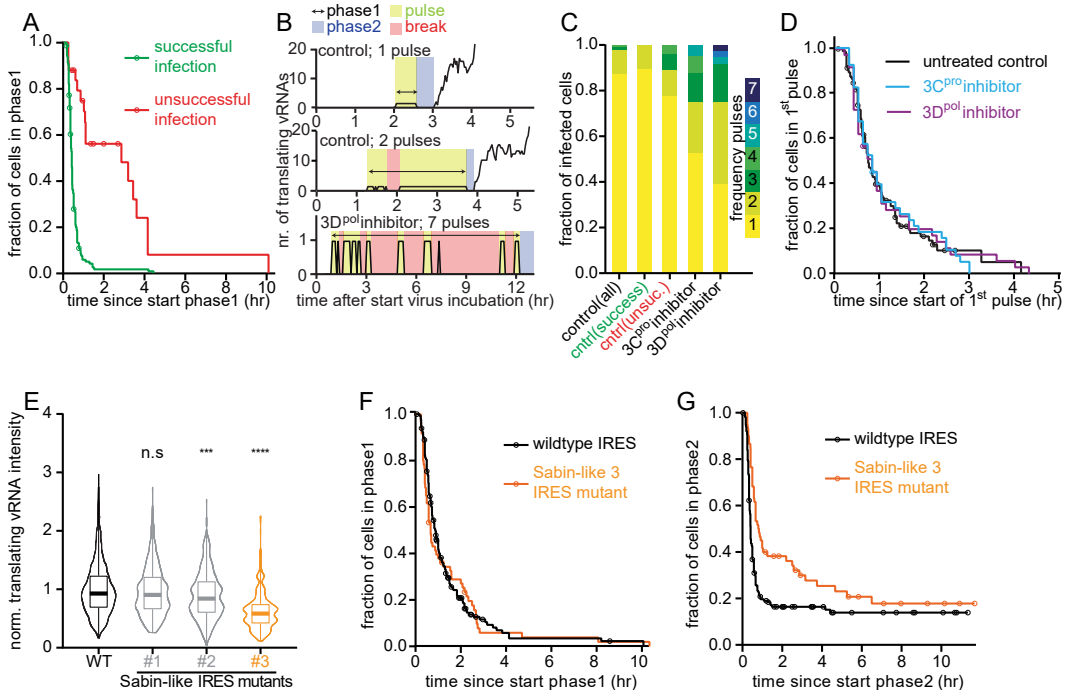
Figure 3. Coordination between vRNA translation and replication.

(**A**) Kaplan Meier graphs showing durations of infection phase1. (**B**) Examples of phase1 pulses in single cells. Colors illustrate individual infection phases, pulses, and breaks. (**C**) Frequency of pulses. 3D^{pol}inhibitor: GPC-N114 (10 μM); 3C^{pro}inhibitor: Rupintrivir (10 μM). (**D**) Kaplan Meier graphs showing duration of first pulses. (**E**) Violin plots of fluorescence intensities of translating vRNAs, normalized to mean of wildtype (WT) IRES virus. (**F, G**) Kaplan Meier graphs showing duration of phase1 (F) and phase2 (G). Data plotted in black (F, G) are replotted from Fig. 2J, K for comparison. Circles (A, D, F, G) indicate last analyzable time-point for individual cells. ***, $p < 0.001$; ****, $p < 0.0001$; n.s., not significant (based on two-tailed unpaired Student's t-test). Scale bar, 15 μm. See also Figure S4. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.

this, a similar median duration of phase1+2 was observed when using a higher MOI (Fig. 2Q) that resulted in a substantial number of cells (~35%) infected by multiple viruses (see STAR Methods; Fig. S3M, N).

As replication of an incoming vRNA depends on newly-translated viral polymerase, we wondered whether the duration of phase1 (which determines the number of viral polymerase proteins produced) is predictive of the success of replication during phase2. We found that cells with unsuccessful replication (phase2 arrest) generally did have an extended phase1 duration (Fig. 3A). In most cells, phase1 consists of an uninterrupted period in which only a single translating vRNA is observed (Fig. 2C, 3B). However, in a subset of cells (15-20%) we observed a single translating vRNA, followed by a period (> 12 min) without detectable GFP foci during phase1, followed by another period with only a single translating vRNA (we refer to these uninterrupted periods of a single GFP spot as translation 'pulses'; Fig. 3B). While multiple

translation pulses were rare in cells with successful replication, they were more frequently observed in cells with unsuccessful replication (Fig. 3C). Reappearance of GFP foci after their initial disappearance was unlikely due to a second infection of the same cell (see STAR Methods; Fig. S4A). Instead, we speculated that reinitiation of translation occurs as result of a failure in replication. Indeed, inhibition of viral replication using the 3Dpol inhibitor resulted in an extended phase1 with additional translation pulses (Fig. 3B, S4B). Similar effects were observed with Rupintrivir (3Cpro inhibitor), which prevents proper processing of the viral polymerase, and as such, also acts as a replication inhibitor (Fig. 3 C, S4B, C). The infection rate was not affected by replication inhibitors, excluding secondary infections as a major cause for the observed increase in translation pulses (Fig. S4D). Notably, in the majority (~65%) of untreated infected cells that underwent a second translation pulse, replication occurred successfully, indicating that the viral strategy to reinitiate translation upon failed replication frequently results in successful replication.

Interestingly, the average duration of each translation pulse was similar for both untreated cells and cells treated with replication inhibitors (Fig. 3D, S4F-K). Replication itself is thus not required to shut down translation. Instead, these findings suggest that an independent phase1 'timer' exists that regulates the translation-to-replication switch. During phase1, the replication machinery (including 3Dpol) that is essential for vRNA replication during phase2 is synthesized. We hypothesized that the phase1 'timer' may reflect production of a threshold amount of viral protein. To test this hypothesis, we reduced viral translation rate through introduction of a point mutation in the viral internal ribosome entry site (IRES). Introduction of an IRES mutation (Sabin-like 3 mutation) reduced viral translation by ~30% (Fig. 3E), consistent with previous studies (Ben M'hadheb-Gharbi et al., 2006; Svitkin et al., 1985, 1990). Surprisingly however, the IRES mutation did not alter the phase1 'timer' that regulates the translation-to-replication switch (Fig. 3F, S4L, M), indicating that translation shut-down occurs at a set time, independently of the amount of viral protein that has been synthesized. Interestingly, IRES-mutant viruses did show a defect in viral replication during phase2 (Fig. 3G), suggesting that entry into the replication phase with a reduced level of viral proteins impairs replication.

Host vs. viral translation

During infection, many RNA viruses shut down host translation, although the specific mechanism varies between viruses (Walsh et al., 2013). Enteroviruses inhibit host translation through cleavage of the translation initiation factor eIF4G by the viral 2A[pro] protein, which was suggested to stimulate viral translation (Hambidge and Sarnow, 1992; Kräusslich et al., 1987; Lamphear et al., 1995). To study the coordination between host and viral protein synthesis directly, we examined viral translation rates along with host translation efficiency over time in single cells. Viral translation rates were assessed based on GFP foci intensity, while host protein synthesis was determined using a fluorescently labeled methionine analog to label newly synthesized proteins (Estell et al., 2017) (Fig. S5A; see STAR Methods). Strikingly, global host protein synthesis was already substantially reduced during infection phase1, which occurs before replication of the incoming vRNA (Fig. 4A-C, S5B). The rapid global reduction in protein synthesis rates most likely reflects an inhibition of translation rather than transcription, as a global decrease in protein synthesis rates
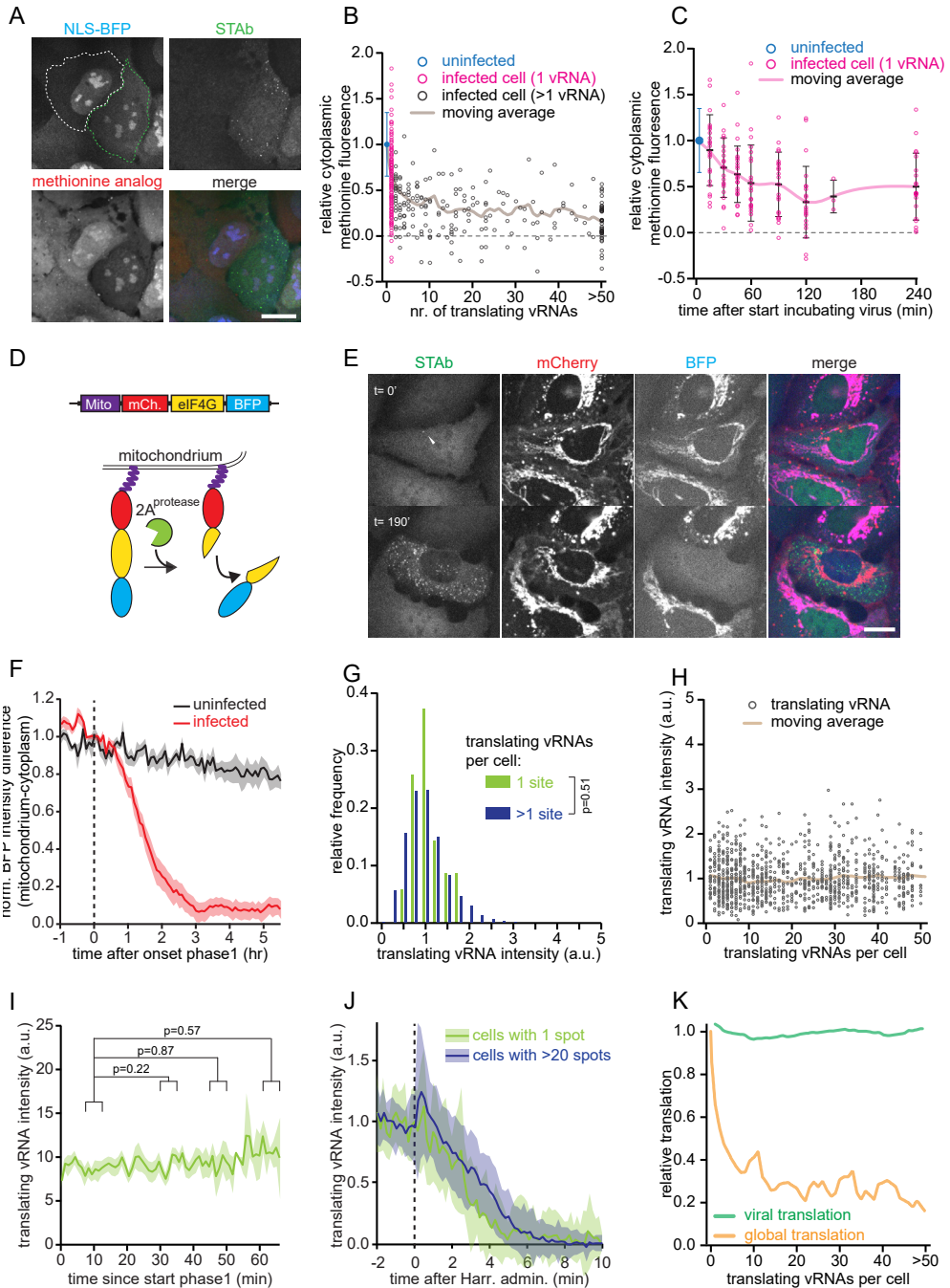
Figure 4. Viral translation efficiency is unaffected by shutdown of host cell translation.

(**A**) *Representative images of uninfected (white outline) and infected (green outline) NLS-STAb cells stained for incorporated methionine analog to indicate global translation rates. (***B, C***) Fluorescence intensities of methionine analog normalized to the mean of uninfected cells (blue, set to 1) from the same sample and to a puromycin treated control (dotted line, set to 0). Data in pink (cells*

Figure 4 continued:

*with a single translating vRNA) is replotted in (C). Every dot represents a single infected cell. (**D**) Cartoon of eIF4G cleavage reporter. (**E**) Representative images of STAb cells expressing the eIF4G cleavage reporter at the indicated time (min) since first detection of a translating vRNA (arrow head). (**F**) BFP fluorescence intensity difference between mitochondria and cytoplasm. Intensity difference traces are aligned to the start of phase1 (dashed line) and normalized to the values of the 15 min before start of phase1. (**G, H**) Comparison of GFP fluorescence intensity of translating vRNA in infected cells with a single or multiple translating vRNAs. (**I, J**) GFP fluorescence intensity of single translating vRNAs over the course of phase1 (I) or after administration of 3 μM harringtonine (J). Intensity traces (J) are aligned to the moment of drug administration (dashed line) and normalized to the mean intensity over 2 minutes before drug treatment. (**K**) Comparison of viral translation and host cell protein production based on the moving averages from (B) and (H). Error bars and shaded areas indicate SD (B, C, I, J) or SEM (F). Statistics based on two-tailed unpaired Student's t-test (G) or paired Wilcoxon test (I). Scale bars, 15 μm. See also Movie S4 and Figure S5. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.*

would be difficult to achieve so quickly through transcription inhibition. To assess how the timing of host translation shutdown relates to the kinetics of eIF4G cleavage, we developed a live-cell eIF4G cleavage biosensor and examined the timing of eIF4G cleavage during early infection. In this biosensor, a fusion of mCherry-eIF4G-BFP was tethered to the outer mitochondrial membrane (Fig. 4D). Cleavage of eIF4G results in dissociation of BFP fluorescence from mCherry-labeled mitochondria. Potent eIF4G cleavage was observed during viral infection (Fig. 4E, S5C movie S4), with half of the eIF4G biosensor cleaved ~1.5 hr after the start of phase1 (Fig. 4F, S5D). Together, these results show that host translation is rapidly shut down after infection. In contrast, viral translation efficiency was similar in all phases of infection (Fig. 4G, H), including before and after host translation shutdown (Fig 4I, J; see STAR Methods). Therefore, despite the accumulation of cleaved eIF4G and host translation shutdown, viral translation remained constant (Fig. 4K). Our kinetic analysis also revealed that host shut-down occurs before (substantial) dsRNA is present in the cell, and therefore could pre-empt production of antiviral proteins.

Virus-host interactions

Innate antiviral responses, most notably IFN signaling, play a key role in repressing the spread of most RNA viruses (Belkowski and Sen, 1987; Samuel, 2001; Schoggins et al., 2011; Stetson and Medzhitov, 2006). While the effects of IFN signaling on gene expression have been well documented (Schneider et al., 2014; Schoggins, 2018), its effects on early viral infection are largely unknown. We stimulated the innate antiviral state by treating cells with interferon α2 (referred to as IFN) and assessed the timing, efficiency, and heterogeneity of infection. IFN signaling did not alter the fraction of cells over time that became infected (i.e. showed at least 1 GFP spot) (Fig. 5A), but led to a striking increase in cells in which the incoming vRNA was arrested during phase2 (replication phase) (Fig. 5B-E). Surprisingly, the subset of cells (~40%) in which the incoming vRNA did undergo successful replication, progressed normally through the later infection phases (Fig. 5C-F), albeit with a slight reduction in the number of vRNAs in phase4 (Fig. 5G). Together, these results show that the IFN-induced antiviral state predominantly acts to prevent replication
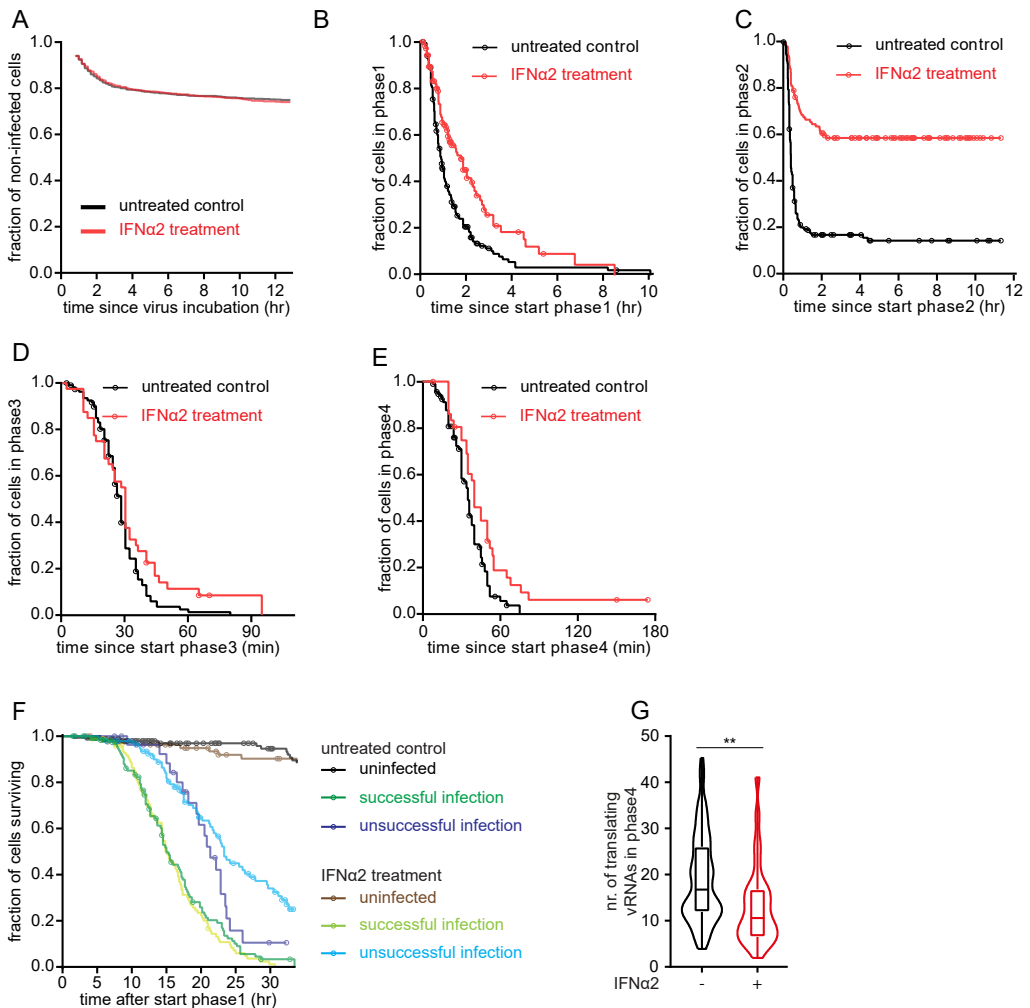
Figure 5. Potent IFN-induced inhibition of replication of the incoming vRNA.

(**A**) *Kaplan Meier graphs of the fraction of uninfected cells remaining after incubation with SunTag-CVB3 with or without IFNα2 pretreatment. Data is corrected for the fraction of cells that were infected before the start of the movies, as indicated by the gap at the start of each graph. Data plotted in black is replotted from Fig. S3M. (**B-E**) Kaplan Meier graphs showing durations of infection phases. Data in black is replotted from Fig. 2H-K for comparison. (**G**) Violin and boxplots of the mean number of translating vRNAs per infected cell during phase4. Data plotted in black is replotted from Fig. 2L. **, p < 0.01 based on unpaired student's t-test. (**F**) Kaplan Meier graphs of the fraction of surviving cells that are either infected or uninfected after indicated treatment. Groups were further subdivided based on whether infection resulted in successful replication during phase2 or not. For comparison, data plotted with dark colors (no pre-treatment control) is replotted from Fig. 2H. Circles (B-F) indicate last analyzable time-point for individual cells. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.*

of the incoming vRNA, with little effect on other aspects of early infection.

Since IFN promotes an antiviral state through upregulation of ISGs (Schneider et al., 2014; Schoggins, 2018), we set out to identify ISG(s) responsible for the suppression of replication of the incoming vRNA. Using RNA-seq, we identified 37 genes that were upregulated >3-fold upon IFN treatment in U2OS cells (Fig. 6A), which included many well-known anti-viral genes (Fig. S6A). We adapted the VIRIM assay to facilitate screening of many experimental conditions (see STAR Methods; the adapted VIRIM assay lacks single-molecule detection sensitivity) and tested the involvement of 28 of these ISGs in suppressing replication of the incoming virus. We identified 6 ISGs (IFIT1, OAS1, OAS3, STAT1, HELZ2, and C19orf66) whose depletion partially neutralized the repression of vRNA replication by IFN (Fig. 6B, S6B). Identification of STAT1, the transcription factor of ISGs, validated our screen (Schneider et al., 2014). Combined knockdown of multiple ISGs relieved the phase2 arrest even more (up to ~60%; Fig. 6C, S6B-G), demonstrating that multiple antiviral mechanisms act in parallel to block replication of the incoming vRNA. Notably, the well-known ISG Protein kinase R (PKR) was present in our screen, but did not inhibit replication of the incoming vRNA (Fig. 6B, S6H, I). While all proteins identified in our screen have been implicated in antiviral signaling previously (Fusco et al., 2017; Kumar et al., 2014; Li et al., 2016; Suzuki et al., 2016), none of these proteins have been implicated in inhibition of early vRNA replication, illustrating the potential of VIRIM in analyzing antiviral mechanisms.

The observed phase2 arrest upon IFN treatment could be caused by an inhibition of vRNA replication, or could result from RNA decay of the incoming vRNA (thus preventing its replication). To explore the role of RNA decay in the phase2 arrest, we first depleted the 5'-to-3' exonuclease Xrn1, as well Dis3L, an essential subunit of the exosome (the major 3'-to-5' exonuclease complex).Neither Xrn1 nor Dis3L depletion affected the phase2 arrest induced by IFN (Fig. S6J-L). In our screen we did identify OAS1 and OAS3 (Fig. 6B, S6B), whose activity is required for activation of the well-known antiviral gene RNAse L (Li et al., 2016), and RNAse L is known to stimulate RNA decay through endonucleolytic cleavage during viral infection (Chakrabarti et al., 2011; Silverman, 2007). Therefore, we tested the role of RNAse L in the IFN-induced phase2 arrest. Indeed, knockdown of RNAse L also mitigated the phase2 arrest induced by IFN, as assessed by VIRIM and the adapted VIRIM assay (Fig. 6C, D, S6J, M), suggesting that RNAse L-mediated RNA decay may be important for the IFN-induced phase2 arrest. RNAse L depletion did, however, not decrease the duration of phase1 (Fig. 6E), suggesting that the RNAse L-induced RNA decay is limited to phase2 and is thus likely triggered by vRNA replication.

Interestingly, in addition to the strong phase2 arrest upon IFN treatment, we observed a modest extension of phase1 (Fig. 5B). Closer examination revealed that the phase1 extension was caused by an increase in the number of translation pulses during phase1 (Fig. 6F, G, S6N), indicative of a replication failure (see Fig. 3A-C). Thus, IFN likely acts through multiple mechanisms to suppress viral replication, including both vRNA degradation and inhibition of replication.

Finally, we further explored the role of one of the identified ISGs, IFN-induced protein with tetratricopeptide repeats 1 (IFIT1). IFIT1 is a repressor of viral infection and is thought to inhibit viral translation through binding to triphosphate-containing 5' viral ends (Daffis et al., 2010; Kumar et al., 2014; Pichlmair et al., 2011). However,

I

II

III

IV

V

VI

&

Figure 6. IFN acts through multiple parallel mechanisms to block replication of the incoming vRNA.

(**A**) *Scatter plot of gene expression changes after IFNα2 treatment. Every dot indicates a single gene, and red dots indicate genes with at least 3-fold increased expression upon IFNα2 treatment in both repeats. (**B, C**) Fraction of infected cells with successfully replicating virus after transfection with the indicated siRNAs, and/or treatment with IFNα2. Every dot represents an independent experiment and black lines indicate the means. Grey control data is based on experiments without siRNA transfections; blue control data is based on non-targeting siRNAs. Dashed lines represent the means of controls. (**D, E**) Kaplan Meier graphs showing duration of infection phase2 (D), or phase1 (E). Data in black and red is replotted from Fig. 5B, C, for comparison. (**K**) Frequency of pulses. Control data is replotted from Fig. 3C. (**G**) Kaplan Meier graphs showing duration of phase1*

Figure 6 continued:

pulses. Data in black is replotted from S3D, for comparison. (I) Kaplan Meier graphs showing duration of infection phase2. Data in black and red is replotted from Fig. 5C for comparison. Circles (D, E, G, I) indicate last analyzable time-point for individual cells. *, p < 0.05; **, p < 0.01; ***, p < 0.01; p < 0.001; ****, p < 0.0001; ns, not significant, Dunnett's multiple comparisons test (B, C), or Gehan-Breslow-Wilcoxon test (D, I). † indicate not-significant genes, which were included in follow-up analysis. See also Figure S6. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.

enteroviruses lack a triphosphate group at the 5' end of their genome (Baggen et al., 2018). Recently, an additional translation-independent function was proposed for IFIT1, although mechanistic details remain unknown (Mears et al., 2019). We reproduced the siRNA phenotype of IFIT1 using an additional IFIT1 siRNA (Fig. S6O), confirming the specificity of IFIT1 knockdown. Next, we performed live-cell analysis using VIRIM of early infection in IFN-treated and IFIT1-depleted cells revealing a reduction in phase2 arrest, without substantial changes to the other infection phases (Fig. 6I). Together, these results reveal that the replication phase of the incoming vRNA is the key point of repression by IFN signaling and that IFN likely acts through multiple parallel mechanisms.

DISCUSSION

Insights into viral replication

Using VIRIM, we found that early CVB3 infection consists of five phases, each reflecting a distinct set of molecular events in the enteroviral life cycle. Similar infection phases were observed in various cell types and for the cardiovirus EMCV and thus may represent a universal phenomenon for picornaviruses. Through detailed analysis of these infection phases, we were able to provide multiple insights into viral translation and replication: 1) we found that the transition from translation to replication is controlled by a 'timer', which likely acts in *cis* on each vRNA individually (Fig 2Q). 2) Translation shutdown is induced independent of vRNA replication, as shutdown occurred with identical kinetics in cells treated with replication inhibitors (Fig. 3D). 3) Upon failed replication, vRNAs frequently reinitiate translation for another attempt to replicate (Fig. 3B, C). As our results suggest that replication inhibition of the incoming vRNA is a major point of attack of antiviral signaling (see below), reinitiation of translation followed by a second attempt at replication may represent a novel mechanism by which viruses combat the antiviral response. Interestingly, this second 'pulse' of translation shows similar kinetics to the first one, indicating that the translation-to-replication timer is reset after every replication attempt. 4) The onset of the translation-to-replication switch is not dependent on the amount of viral protein that has been produced, as CVB3 with a mutated IRES that shows reduced translation initiates replication with similar kinetics as wildtype CVB3. (Fig. 3F). Yet, the efficiency of virus replication is lower for IRES-mutant viruses, suggesting that they enter the replication phase with insufficient viral protein to execute replication with high fidelity (Fig. 3G). We speculate that the timing of the translation-to-replication switch (i.e. duration of phase1) has evolved to occur as soon as possible, yet only after sufficient levels of viral proteins have been produced to ensure productive replication.

Insights into virus-host competition

To combat viral infection, host cells must rapidly detect and inhibit a virus upon infection. However, viruses have evolved various mechanisms to evade detection and/or counteract antiviral pathways (Wang et al., 2018). Our study reveals that this competition between picornaviruses and their host is initiated very early during infection. Advanced technologies, like VIRIM or the recently developed assay to identify host proteins interacting with vRNAs (Kim et al., 2020) are ideally suited to study early viral replication. Here, we applied VIRIM to study virus-host interactions during the first hours of infection. We find that CVB3-induced host cell modifications occur extremely rapidly upon infection: eIF4G cleavage, host cell translation shutdown, and impairment of nuclear transport are all initiated within minutes of initial infection, at the start of phase1, and generally reach completion 1-3 hr after initial infection (Fig. 1G, H, 4B, C). Very few (<100) viral proteins have been synthesized when host cell modifications initiate, highlighting the remarkable potency of enteroviral proteases (which drive many of the abovementioned host cell modifications). The swift host shutdown may provide a key advantage to the virus; substantial host cell inhibition is already induced *before* replication of the incoming vRNA occurs, and maximal host inhibition is generally achieved at the moment when the larger burst of replication occurs (phase 4/5). Since the formation of dsRNA during viral replication is a key trigger of an antiviral response, shut down of host protein synthesis before formation of dsRNA, may be an effective viral strategy to limit production of antiviral signaling molecules. These findings on virus-host interaction kinetics may explain in part why only a small subset of infected cells mounts a strong IFN response (Doğanay et al., 2017; Patil et al., 2015; Zawatzky et al., 1985).

Many RNA viruses specifically target the host cell translation machinery (Chase and Semler, 2012; Walsh et al., 2013). As discussed above, it is possible that host translation shut-down counteracts antiviral signaling. An additional consequence of host cell translation shutdown is that the translation machinery (e.g. ribosomes) becomes available exclusively for vRNA translation, possibly boosting vRNA translation. Previous reports suggested that eIF4G cleavage stimulates vRNA translation (Hambidge and Sarnow, 1992; Kräusslich et al., 1987; Lamphear et al., 1995). Here, we examined the relationship between eIF4G cleavage and host translation shutdown with viral translation rates using VIRIM. We found that vRNA translation rates are similar before and after host translation shutdown and eIF4G cleavage (Fig. 4K), indicating that viral translation is not directly affected by these processes. Our results do not exclude, however, the possibility that host cell translation shutdown boosts viral translation at later stages during infection, when the number of vRNAs is substantially larger and translation machinery may become limiting.

A key aspect of antiviral signaling is the production of IFN, which induces the expression of a large set of ISGs and strongly represses virus spreading. While the set of genes that is upregulated upon IFN signaling has been well documented, the function of the majority of ISGs in combatting enterovirus infection is largely unknown. Using VIRIM, we find that IFN signaling causes a strong and specific inhibition of phase2, i.e. replication of the incoming vRNA (Fig. 5C). In all cases where the incoming vRNA replicated successfully, substantial viral replication

was observed during the subsequent infection phases (Fig. 5D-G). These findings indicate that replication of the incoming vRNA represents a major point of attack for antiviral signaling. To provide insights into the mechanism of this attack, we screened for ISGs that inhibit early enteroviral replication. We provide evidence that multiple mechanisms, including replication-triggered RNA decay through OAS-RNAse L, as well as inhibition of vRNA replication play important roles in mediating the effects of IFN on early viral replication. Taken together, these results uncover a major target of antiviral signaling (replication of the incoming vRNA) and identify multiple genes involved in this process. More broadly, this study shows how VIRIM can be used to dissect antiviral signaling. Whether replication of the incoming vRNA is also a bottleneck in infection of other +RNA viruses is an important topic for future research. VIRIM may also be a useful tool to dissect the mechanisms of antiviral drugs and potentially screen for novel drugs with specific modes of action during early infection.

Success-rate of infection

Previous work has shown that multiple viral particles (in some cases even hundreds) are needed for (detectable) infection of a cell (Klasse, 2015), suggesting that productive infection by individual viral particles fails in the majority of cases. Little is known about the limiting step(s) in the viral life cycle that is responsible for this bottleneck(s). Using VIRIM, we discovered that in 15-20% of the cells replication of the incoming vRNA fails (Fig. 2I, N), resulting in elimination of the viral infection. This number dramatically increased upon activation of the IFN signaling pathway (Fig. 5C). In addition, we found that ~40% of vRNAs were not undergoing translation (Fig. 1E). It is possible that vRNAs cycle between a translated and non-translated state. Alternatively, a subset of vRNAs may be defective in translation, which could also contribute to the failure of some viral particles in host cell infection. VIRIM will be a valuable tool to study the success-rate of different steps in the viral life cycle of many different viruses, in different cell types or in distinct cell states (e.g. stimulated with different signaling molecules).

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS
Conceptualization, SB, HHR, FJMvK and MET; Methodology, SB, HHR, LJMB, FJMvK and MET; Validation: SB, HHR; Formal analysis, SB; Investigation, SB, HHR, LJMB, TP, and ALWvV; Resources, JB, and HC; Data curation, SB, HHR, LJMB; Writing-Original draft, SB and MET; Writing – Review & Editing, all authors; Visualization,

I

II

III

IV

V

VI

&

SB; Supervision, FJMvK, and MET; Funding acquisition, FJMvK, and MET.

COMPETING INTERESTS
HC is inventor of several patents related to organoid technology; his full disclosure is given at https://www.uu.nl/staff/JCClevers/.

STAR METHODS
RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-dsRNA	English & Scientific Consulting	Cat #J2
Mouse polyclonal anti-3Dpol	Oh et al., 2009	N/A
Mouse monoclonal anti-tubulin	Sigma-Aldrich	Cat# T9026
Rabbit polyclonal anti-eIF4GI	Bethyl Laboratories	Cat# A300-502A
Mouse monoclonal anti-mCherry	Invitrogen	Cat# MA5-32977
Donkey polyclonal anti-mouse-Cy5	Jackson ImmunoResearch	Cat# 715-175-151
Goat polyclonal anti-mouse IRDye680	LI-COR	Cat# 926-68070
Goat polyclonal anti-rabbit IRDye800	LI-COR	Cat# 926-32211
Virus strains		
CVB3	Wessels et al., 2006	N/A
EMCV	Duke and Palmenberg, 1989	N/A
eGFP-CVB3	Lanke et al., 2009	N/A
SunTag-EMCV	This study	N/A
SunTag-CVB3	This study	N/A
SunTag-CVB3Sabin-like1	This study	N/A
SunTag-CVB3Sabin-like2	This study	N/A
SunTag-CVB3Sabin-like3	This study	N/A
Chemicals, Peptides, and Recombinant Proteins		
DMEM	Gibco	Cat# 31966021
Leibovitz's L15 medium	Gibco	Cat# 21083-027
Opti-MEM	Sigma-Aldrich	Cat# 11058-021
Methionine-free DMEM	Gibco	Cat# 21013024
Advanced DMEM/F12	Thermo Fisher scientific	Cat# 12634-010
TryplE	Thermo Fisher scientific	Cat# 12605010
Fetal Bovine Serum (FBS)	Sigma-Aldrich	Cat# F7524
Penicillin-Streptomycin	Gibco	Cat# 15140-122
Glutamine	Sigma-Aldrich	Cat# G6392-1VL
Cysteine	Sigma-Aldrich	Cat# C7352-10MG
FuGENE 6	Promega	Cat# E231A
Lipofectamine 2000	ThermoFisher	Cat #11668019
Lipofectamine RNAiMAX	Invitrogen	Cat# 13778-075
Polyethylenimine	Polysciences Inc	Cat# 23966

Polybrene	Santa Cruz Biotechnology, Inc	Cat# sc-134220
Propidium Iodide	Sigma-Aldrich	Cat# P4170
Doxycycline	Sigma-Aldrich	Cat# D9891-1G
Zeozin	Invitrogen	Cat# R25001
Puromycin	ThermoFisher Scientific	Cat# 12122530
Harringtonine	Cayman Chemical	Cat# 15361
GPC-N114	Van der Linden et al., 2015	N/A
Rupitrivir	Sigma-Aldrich	Cat# PZ0315
Interferon α2	Sigma-Aldrich	Cat# IF007
scFv-sfGFP-StrepII	This study	N/A
Glucose oxidase	Sigma-Aldrich	Cat# G2133-10KU
Catalase	Sigma-Aldrich	Cat# C3515-10MG
Cultrex Basement Membrane Extract (BME), Growth Factor Reduced, Type 2	R&D Systems, Bio-Techne	Cat# 3533-001-02
Protease inhibitor cocktail	Roche	Cat# 11697498001
Rho kinase inhibitor	Calbiochem	Cat# 555550
TRIsure	Bioline	Cat# 38033
Atto633-NHS	Atto-Tec	Cat# AD 633-31
Cy5-azide	Lumiprobe	Cat# A3030
Amino-11-ddUTP	Lumiprobe	Cat# 15040
L-Homopropargylglycine	Sigma-Aldrich	Cat# 900893-100MG
Desthiobiotin	IBA Life Science	Cat# 2-1000-001
Digitonine	Sigma-Aldrich	Cat# D141
Paraformaldehyde	Aurion	Cat# 15710
Formamide	ThermoFischer	Cat# AM9342
Critical commercial assays		
T7 RiboMax	Promega	Cat# P1320
HiScribe	New England Biolabs	Cat# E2040S
Superscript III reverse transcriptase	Invitrogen	Cat# 18080093
Tetro reverse transcriptase	Bioline	Cat# BIO-65050
Terminal deoxynucleotidyl Transferase	ThermoFisher	Cat# EP0162
iQ SYBR Green SuperMix	Bio-Rad	Cat# 1708885
Nucleospin RNA	Macherey-Nagel	Cat# 740990.50
Zymo RNA cleanup	ZymoResearch	Cat# R2061
Deposited data		
Raw and analyzed RNA sequencing -/+ IFN	Gene expression omnibus https://www. ncbi.nlm.nih.gov/geo/	GSE159280
Raw data of imaging experiments	Mendeley data	http://dx.doi.org/10.17632/9sxbk6cvn9.1
Experimental Models: Cell Lines		
U2OS cells	Tanenbaum lab	Cat# HTB-96
A549 cells	Van Kuppeveld lab	Cat# CCL-185
BHK-21 cells	Van Kuppeveld lab	CCL-10
HeLa cells	Tanenbaum lab	Cat# CCL-2
RPE1 cells	Tanenbaum lab	Cat# CRL-4000
HEK293T cells	Tanenbaum lab & Van Kuppeveld lab	Cat# CRL-3216

Oligonucleotides		
See Table S2 for all sequences of smFISH probes, siRNAs, or qPCR oligos	This study	N/A
Recombinant DNA		
Plasmids used in this study	Tanenbaum lab	https://www.tanenbaumlab.org/plasmids
Software and Algorithms		
ImageJ	NIH	https://imagej.nih.gov/ij/
Micromanager	Micro-Manager 1.4.22	https://micro-manager.org
NIS-Elements Imaging software	Nikon	https://www.microscope.healthcare.nikon.com/en_EU/products/software
Graphpad Prism 8	GraphPad Software Inc	http://www.graphpad.com/scientific-software/prism/
Other		
StrepTactin Sepharose beads	IBA life sciences	Cat# 2-1201-002
Zeba desalting column	VWR	Cat# GE17-0851-01
96-well glass bottom imaging plates-(Matriplates)	Brooks Life Science Systems	Cat# MGB096-1-2-LG-L

RESOURCE AVAILABILITY
Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Marvin Tanenbaum (M.Tanenbaum@hubrecht.eu).

Materials availability

The unique/stable reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

Data and code availability

The RNA sequencing data of this study has been deposited in the Gene Expression Omnibus (GEO) under accession code GSE159280. A selection of raw imaging data is made available through Mendeley data: http://dx.doi.org/10.17632/9sxbk6cvn9.1.

EXPERIMENTAL MODEL AND SUBJECT DETAILS
Cell lines

Human U2OS, HeLa, RPE1, HEK293T cells used for imaging, lentivirus production, and CVB3 production were grown in DMEM (4.5 g/L glucose, Gibco) supplemented with 5% fetal bovine serum (Sigma-Aldrich) and 1% penicillin/streptomycin (Gibco). Human A549 cells and Hamster BHK-21 for imaging and EMCV production were cultured in in DMEM (4.5 g/L glucose, Gibco) supplemented with 10% fetal bovine serum (Sigma-Aldrich) and 1% penicillin/streptomycin (Gibco). All cells were grown with 5% CO_2 at 37°C. Cells were confirmed to be mycoplasma negative.

Intestinal organoids

Tissue from the human small intestine was obtained from the UMC Utrecht with informed consent of the patient. The patient was diagnosed with small intestinal cancer that was resected. A sample from non-transformed, normal mucosa was obtained for organoid culture used in the study. The study was approved by the UMC Utrecht ethical committee (Utrecht, The Netherlands) and was in accordance with the Declaration of Helsinki and according to Dutch law. This study is compliant with all relevant ethical regulations regarding research involving human participants. Human small intestinal cells were isolated, processed and cultured as described previously (Sato et al., 2011).

METHOD DETAILS
Plasmids

The sequences of plasmids used in this study can be found via www.tanenbaumlab. org/plasmids. The eIF4G1 coding sequence was amplified from Addgene #45640.

Cell line generation

For generation of cell lines stably expressing transgenes, lentiviral transduction was used. To produce lentivirus, HEK293T cells were transfected using Polyethylenimine (PEI) with the lentiviral plasmid of interest and packaging vectors psPax and pMD2. The cell culture medium was refreshed 1 day after transfection. The supernatant containing the lentivirus was collected 3 days after transfection. To make stable cell lines, cells (U2OS, HeLa, or A549) were seeded at ~35% confluency one day before infection. To infect cells, the viral supernatant was added to the cells along with Polybrene (10 mg/ml) (Santa Cruz Biotechnology Inc) and the cells were spin-infected for 90-120 minutes at 2000 rpm at 25°C. After spin-infection, the medium was refreshed and cells were cultured for minimally 2 days before further analysis was performed. To isolate cells that express the fluorescent transgenes, cells were FACS-sorted. To generate U2OS and HeLa cell lines expressing STAb in which all cells expressed the STAb at similar levels, single cells were sorted into 96-wells plates to generate monoclonal cell lines. To generate a cell line expressing nuclear BFP, a previously-generated monoclonal cell line stably expressing TetR, STAb, and PP7-2xmCherry-CAAX ((Yan et al., 2016) hereafter referred to as STAb U2OS cells) was infected with lentivirus encoding nuclear BFP. A polyclonal population of BFP-positive cells was sorted that expressed BFP at similar levels. To generate an A549 cell line expressing STAb, cells were infected with the STAb lentivirus and a polyclonal population of GFP positive cells was sorted. Cells that had similar GFP expression levels as those in the STAb-U2OS cells were selected.

To generate a cell line expressing the eIF4G cleavage reporter, we could not use lentiviral transduction, as the eIF4G cleavage reporter is too large for efficient lentivirus production. Instead, the cell line expressing the eIF4G cleavage reporter was generated by transfecting (Fugene 6; Promega) the STAb-U2OS cells with a plasmid encoding the eIF4G cleavage reporter. One day after the transfection, the medium was refreshed and selection for stable integration of the eIF4G cleavage reporter was initiated (0.4 mg/ml Zeocin (Invitrogen)). Selection was performed

for 14 days during which the Zeocin-containing medium was refreshed every 5 days. Note that eIF4G cleavage reporter was expressed from an inducible promoter to prevent possible toxicity due to long-term expression. Expression of the eIF4G cleavage reporter was induced by incubating cells with doxycycline (1 µg/ml; Sigma-Aldrich) for 12-24 hr before an experiment.

Generation of STAb intestinal organoids

For generation of intestinal organoids stably expressing STAb, lentiviral transduction was used, as described previously (Koo et al., 2012). To generate a polyclonal population of STAb expressing intestinal cells, organoids were dissociated by incubating 5 min with TrypLE (TryplE Express; Life Technologies) and GFP positive cells were FACS-sorted. Only cells with GFP fluorescence comparable to STAb U2OS were sorted. After sorting, cells were cultured for three days in the presence of a Rho kinase inhibitor (Rho kinase inhibitor, 10µM, Calbiochem) before further analysis was performed.

CVB3 and EMCV design and production

The CVB3 used in this study was derived from the pRibCVB3/T7 plasmid, which contains the cDNA of CVB3 strain Nancy driven by a T7 RNA polymerase promoter (Wessels et al., 2006). The EMCV used in this study was derived from the pM16.1 plasmid (Duke and Palmenberg, 1989),. Mutations in the IRES were introduced by site-directed mutagenesis of the pRIB infectious clone. To make SunTag-CVB3 or SunTag-EMCV, 5 copies of the SunTag-coding sequence were introduced in-frame upstream of the viral coding region, at a location that was previously successfully used for insertions in recombinant CVB3 (Lanke et al., 2009; see Fig. S1A and Data S1 for details). A cleavage site for the viral 3C protease was included between the SunTag and the rest of the viral polyprotein (ALFQG for CVB3, VFETQG for EMCV) to enable removal of SunTag from the viral protein VP4 (CVB3) or L protein (EMCV) and to prevent possible SunTag interference during virus particle assembly.

Virus stocks were made as described previously (Lanke et al., 2009; Wessels et al., 2006). In brief, the infectious clones were linearized using the MluI (CVB3) or BamHI (EMCV) restriction enzymes and used as template for in vitro transcription (T7 RiboMAX; Promega or HiScribe; New England Biology). RNA was purified (Nucleospin RNA; Machery-Nagel or ZymoResearch) and transfected into HEK293T or BHK-21 cells using Lipofectamine 2000 (ThermoFisher). After 2-3 days, when complete CPE (cytopathic effect) was observed, cells and supernatant were collected and freeze-thawed three times. The supernatant containing CVB3 was aliquoted and stored at -80°C. Virus titers were determined by endpoint titration, as described previously (Reed and Munech, 1938).

To confirm that recombinant viruses contained the correct inserts, viral RNA was isolated from the virus stocks (Nucleospin viral RNA; Machery-Nagel) and cDNA was synthesized using Superscript III reverse transcriptase (Invitrogen) according to manufacturer's protocol. Next, the region of interest was amplified by PCR (see Fig. S1A), and the correct size of the PCR product was verified by agarose gel electrophoresis and the sequence of the SunTag insert and IRES mutations were verified by Sanger sequencing.

siRNA transfections

All siRNAs used in this study were ordered as SMARTpool ONTargetPlus reagents from Dharmacon, except for: siXrn1 (AGAUGAACUUACCGUAGAAUU; (Hoek et al., 2019)), siDs3L (CCAUGUAACCGUAAGAAUA; (Staals et al., 2010)); siRNASEL_#2 (SantaCruz sc-45966), siIFIT1_#2 (Ambion s7150), siIFIT1_#3 (CCAGACAAUGGAUAUAUUAAG; (John et al., 2018)). STAb U2OS cells or NLS-BFP STAb U2OS cells were reverse-transfected with siRNAs at a final concentration of 10 nM siRNA using RNAiMAX transfection reagent (Invitrogen) and seeded in 48-well plates. After 24-48 hr, the cells were trypsinized and re-plated on either glass or plastic plates for imaging or for qPCR analysis, respectively. One day after re-plating, cells were imaged or harvested for qPCR. Where indicated, cells were treated with IFN (1000 U/ml IFNα2 (Sigma)) approximately 24 hr before analysis.

Quantitative RT-PCR (qPCR)

Virus replication

To compare replication kinetics between CVB3, GFP-CVB3, and SunTag-CVB3 viruses in either HeLa, U2OS, or STAb U2OS cells, the amount of viral RNA was determined over time after infection. Cells were plated in 384-well plate 1 day before infecting with the indicated recombinant viruses at an MOI (multiplicity of infection) of 1. 1 hr after inoculation, the virus-containing medium was removed, and cells were washed 3 times with PBS. At the indicated time points, cells were harvested and RNA was extracted using Nucleospin RNA (Machery-Nagel) according to manufactor's guidelines. qPCR was used to assess viral load over time.

siRNA knockdown efficiency

To determine the knock-down efficiency of siRNAs, siRNA-treated STAb cells were harvested and RNA was isolated using TRIsure (Bioline). Next, cDNA was synthesized using Random hexamers and Tetro Reverse Transcriptase (Bioline). qPCRs were performed using SYBR-Green Supermix (Bio-Rad) on a Bio-Rad Real-time PCR machines (CFX Connect Real-Time PCR Detection System). mRNA levels were determined by qPCR. If the Cq (quantitation cycle) of a sample was higher than the Cq of a water control, the sample was excluded from analysis. See Table S2 for sequence details of oligonucleotides used for qPCR. All RNA levels were normalized to GAPDH mRNA levels.

smFISH

Labeling of smFISH probes

Single-molecule Fluorescence In Situ Hybridization (smFISH) was performed as described previously ((Lyubimova et al., 2013; Raj et al., 2008)). Custom-made oligonucleotide probes (96) targeting the positive RNA strand of CVB3 were designed using the website www.biosearchtech.com (See Table S2 for sequence of the probes). Probes were labeled with Atto633-NHS (Atto-Tec), as described previously (Gaspar et al., 2018). In brief, Atto633-NHS was dissolved in DMSO and mixed with $NaHCO_3$ (final concentration 0.05 M; pH 8.4) and Amino-11-ddUTP (5

mM; Lumiprobe). The oligonucleotides were labeled by mixing 200 µM of each oligo with Terminal deoxynucleotidyl Transferase (TdT) buffer, 10 mM dye solution, and TdT (ThermoFischer) and incubating at 37°C overnight. Next, the labeled probes were precipitated using 100% ethanol, washed with 80% ethanol to remove free dye, and resuspended in nuclease-free water to a final concentration of 30 µM. Prior to hybridization, probes were diluted in TE to a working stock concentration of 1 µM.

Probe hybridization

STAb cells were plated on glass and incubated with SunTag-CVB3 as described in the section *'Live-cell imaging; cell culture before imaging'*. Two additional PBS washing steps were performed after virus incubation to remove SunTag-CVB3 adhering to the outside of cells. 2-4 Hr after inoculation with virus, live-cell imaging was terminated and cells were washed with PBS and fixed with 3.7% paraformaldehyde in PBS for 5 minutes at room temperature. After fixation, cells were washed twice with PBS and permeabilized with 100% ice-cold ethanol at 4°C for 30 minutes, cells were then washed twice for 15 minutes with smFISH wash buffer (2X SSC, 10 % formamide in DEPC-treated water). smFISH probes were incubated in hybridization buffer (10 nM probe concentration; 1% dextran sulfate, 2x SSC, 10% formamide in DEPC-treated water) at 37°C. After overnight incubation with smFISH probes, cells were washed twice for 1 hr at 37 °C and for 15 min at room temperature with smFISH wash buffer at room temperature. The wash buffer was replaced with imaging buffer (10 mM Tris, pH 8; 2x SCC; 0.63 % glucose, supplemented with glucose oxidase (Sigma) and catalase (Sigma)) and samples were stored at 4°C until imaging.

Combined smFISH and immunofluorescence

To combine smFISH and immunofluorescence, only a single wash was performed after hybridization of the smFISH probes. Then, the immunofluorescence staining was performed as described in the section *'Immunofluorescence of dsRNA and 3D^{polymerase}'*, starting with blocking until washing away the secondary antibody. Then, a final wash step using smFISH wash buffer was performed and the wash buffer was replaced with imaging buffer and stored at 4°C until imaging.

Immunofluorescence of dsRNA and 3D^{polymerase}

Cells infected with SunTag-CVB3 virus were first followed by time-lapse microscopy for 3-7 hr. and then cells were fixed with 3.7% paraformaldehyde for 30 minutes and permeabilized using PBS + 0.1% Triton X-100 for 5 minutes. Blocking was performed using PBS + 2% BSA + 50 mM NH_4Cl for 45 minutes. Anti-dsRNA antibody (J2, English & Scientific Consulting) or anti-3D^{polymerase} (Oh et al., 2009) diluted 1:1000 in blocking buffer was incubated for 45 minutes. After incubation with the primary antibody, three wash steps with block buffer were performed. Donkey anti-mouse Cy5 (Jackson lab) was used as a secondary antibody (1:200 in blocking buffer) and incubated for 45 minutes. After washing away the secondary antibody once with PBS, samples were kept in PBS at 4°C, until imaging.

Immunofluorescence with STAb

Purification of STAb

To purify STAb, RPE1 cells stably expressing STAb-StrepII were harvested, resuspended in lysis buffer (50 mM HEPES/KOH, pH 7.4, 200 mM KCl, 0.5 % Triton X-100, 1 mM PMSF, and EDTA free protease inhibitor cocktail (Sigma Aldrich)) on ice, and sonicated. To remove cell debris, the sample was centrifuged at 20000g for 1 hr at 4 °C and the supernatant was collected. To purify STAb-StrepII, the supernatant was incubated for 1 hr at 4 °C with StrepTactin Sepharose beads (IBA life sciences) that were pre-equilibrated with wash buffer (20 mM HEPES/KOH, pH 7.4, 200 mM KCl) and HEPES/KOH, 200 mM KCl, 0.5 % Triton X-100. After washing with wash buffer, the STAb-StrepII was eluted from the beads using elution buffer (20 mM HEPES/KOH, pH 7.4, 200 mM KCl, 5 mM desthiobiotin (IBA life Sciences)). The eluate was buffer exchanged using a Zeba desalting column (ThermoFisher) into storage buffer (20 mM HEPES/KOH, 200 mM KCl, 5 % glycerol). The purified protein was stored at -80 °C.

STAb staining

Immunofluorescence staining of SunTag using purified STAb was combined with smFISH as described in the section 'Combined *smFISH and immunofluorescence'*, using 1:100 STAb dilution in blocking buffer. As the purified STAb is already fluorescently labeled, no secondary antibody incubation step is required. Instead, immunofluorescence staining of SunTag was finalized by six wash steps with blocking buffer after primary antibody incubation.

Analysis of global translation efficiency

Global translation efficiency was determined by labeling newly-synthesized proteins using a methionine analog, followed by fluorescence labeling of the methionine analog based on an adapted protocol of a previously established assay (Estell et al., 2017). NLS-BFP STAb cells were seeded in 96-well glass-bottom plates. To deplete methionine, regular cell culture medium was replaced with methionine-free medium ((Met)-free DMEM (Gibco) supplemented with 2 mM glutamine (Sigma Aldrich), 0.02 mg/ml cysteine (Sigma Aldrich), 5% fetal bovine serum (Sigma Aldrich) and 1% penicillin/streptomycin (Gibco)). SunTag-CVB3 was then added to cells at indicated time-points and the methionine analog, L-Homopropargylglycine (HPG, 1 µM, Sigma) was added to cells 15 minutes before fixation. Several controls were taken along in these experiments: 1) cells that were not incubated with virus, 2) cells that were not treated with the HPG, or 3) cells that were treated with translation inhibitor puromycin (ThermoFisher Scientific; 0.1 mg/ml) starting 1 minute before incubation with HPG. HPG was washed away with ice-cold wash buffer (10mM HEPES/KOH pH7.4; 10mM NaCl; 5mM MgCl$_2$; 300mM sucrose) and cells were fixed in wash buffer with 3.7% PFA for 30 minutes at room temperature. After fixation, the cells were permeabilized with 0.015% digitonine (Sigma) in wash buffer for 5 minutes on ice. A click reaction was performed to fluorescently label HPG with Cy5 in freshly prepared click reaction buffer (100 mM Tris pH 8.8; 100 mM ascorbic acid; 1 mM CuSO4; 10 µM Cy5-azide (Lumiprobe)) for 15 minutes at room

temperature. The click reaction buffer was washed away 5 times and replaced with PBS and samples were stored on ice, until imaging (which occurred within 12 hr after finalizing the click reaction). Note that the protocol described above results in nuclear background fluorescence (Fig. S5A), which was not observed in a previous study, as in the earlier study permeabilization is performed prior to fixation (Estell et al., 2017). To prevent loss of cytoplasmic signal during permeabilization before fixation, cells were fixed prior to permeabilization in our experiments.

Western blot

STAb cells with inducible expression of the eIF4G cleavage reporter were seeded at 75% confluency in 6 cm dishes. One day after seeding, wild-type CVB3 virus (MOI=2) was added to each well. At the indicated time points post infection, cells were released by trypsin treatment, and subsequently lysed on ice in lysis buffer (40 mM Tris-HCl pH7.4, 150 mM NaCl, 10 mM EDTA, 1% NP-40, protease inhibitors [Roche]) for 30 min. For each sample, 100 µg of total protein content in 1xLSB was heated to 95^{0}C for 5 min, separated on a 7.5 % SDS-PAGE gel, and transferred to nitrocellulose membranes by wet transfer. Membranes were blocked in block buffer (PBS + 0.1% Tween-20 + 2% BSA) for 45 min, and then probed using primary antibodies mouse anti-tubulin (Sigma-Aldrich), rabbit anti-eIF4GI (Bethyl Laboratories), and mouse anti-mCherry (Invitrogen), used at 1:1000 dilution in block buffer for 1 hr. After three wash steps in wash buffer (PBS + 0.1% Tween-20), membranes were incubated in secondary antibodies goat anti-mouse IRDye680 (LI-COR) and goat anti-rabbit IRDye800 (LI-COR), used at 1:10,000 dilution in block buffer for 45 min. Membranes were washed twice in wash buffer and membranes were imaged using the Odyssey Imager (LI-COR).

RNA sequencing

STAb cells were seeded at 20% confluency in a 12-wells plate. One day after seeding, cells were incubated with 1000 U/ml IFN for 24h. As control, cells without IFN treatment were processed in parallel. Cells were harvested and RNA was isolated using TRISure (Bioline), according to manufacturer's guidelines. Further processing and sequencing was performed by Single Cell Discoveries, based on the CEL-seq2 protocol (Hashimshony et al., 2016; Muraro et al., 2016).

Live-cell microscopy

Microscopes

Imaging experiments were performed using either a 1) Nikon TI inverted microscope with NIS Element Software (Nikon), equipped with a perfect focus system, a Yokagawa CSU-X1 spinning disc, an iXon Ultra 897 EM-CCD camera (Andor), or 2) a Nikon TI2 inverted microscope with NIS Element Software, equipped with a perfect focus system, a Yokagawa CSU-X1 spinning disc and an Prime 95B sCMOS camera (Photometrics). Both microscopes were equipped with a temperature-controlled hood. For experiments involving long-term time-lapse analysis (>3 hr) a 60x 1.40 NA oil-immersion objective was used, while a 100x 1.49 NA oil-immersion objective was used for short-term analyses. A 100x 1.49 NA oil-immersion objective was used for

experiments using 2D organoid cells. The siRNA screen (Fig. 6B, C, S6B, H, J, O) was performed using a 40x 0.95 NA air objective.

Image acquisition

For long-term (>3 hr) time-lapse imaging to determine infection phase durations, x,y positions for imaging were selected randomly. Images were acquired every 1, 2, 3, or 5 minutes for 3-12 hours, using 50-70 ms exposure times. Multiple Z-slices (~10-15 planes with 0.8 µm steps) were imaged for GFP to ensure that the entire cell was imaged. If relevant, a single Z-slice in the middle of the cell was also imaged for BFP and/or mCherry. To determine cell survival, images were acquired every 10 minutes for 36 hr, using similar laser and Z-slice settings. To quantify success-rate of infection after siRNA-mediated knockdown of IFN-induced genes, a large-image comprised of 6-8 to 6-8 regular field-of-views (FOVs) was generated for each condition. Every large-image was acquired every 15-30minutes for 10-12h. Images of one Z-plane were acquired for BFP, and mCherry and 4 Z-slices (1.5 µm steps) were imaged for GFP.

For short-term live-cell imaging, x,y positions for imaging were selected based on the presence of translating vRNAs. For harringtonine run-off experiments, cells were selected based on the number of translating vRNAs (either 1 or 20-50) per cell and images were acquired every 10 seconds for 20-30 minutes, using 50 ms exposure and 10-15 Z-slices with 0.8 µm steps. To determine the intensity of translating vRNAs, images were acquired for a single time-point using 50-100 ms exposure times. Multiple Z-slices (12-20 with 0.5 µm steps) were imaged for GFP to ensure that the entire cell was imaged. Compared to the long-term imaging experiments, 2-fold higher GFP laser power was used.

To determine diffusion speeds of translating vRNAs, cells were selected based on the number of translating vRNAs. A single GFP Z-slice was acquired in the middle of cells, using maximal laser power with 20 ms exposure for 100 frames at ~40 ms interval (Fig. S1M). In a separate assay (Fig. 1J, K), the diffusion speed of translating vRNAs during infection phase1 was measured in randomly-selected cells and imaged every 1 minute for 4 hr using using 50-70 ms exposure times. Multiple Z-slices (12 planes with 0.8 µm steps) were imaged for GFP and a single Z-slice was imaged for BFP and mCherry. To determine the diffusion speed of translating cellular mRNAs, cells transfected with an mRNA reporter plasmid were selected that contained less than 10 translating reporter mRNAs per cells. Images of 10-15 Z-slices with 0.5 µm steps were acquired every 10 s using 50 ms exposure time.

Cell culture for imaging

Unless noted otherwise, live-cell imaging was performed by seeding STAb cells (U2OS, HeLa, or A549) or NLS-BFP STAb cells 1 day before imaging in a 96-well glass-bottom plate (Matriplates, Brooks Life Science Systems) at ~40-45% confluency. An MOI of ~0.25 was used for infection using SunTag-CVB3 virus and cells were incubated with virus-containing medium for ~30 minutes. After incubation with virus-containing medium, the virus-containing medium was replaced with imaging medium (pre-warmed CO_2-independent Leibovitz's-15 medium (Gibco) containing 5% fetal bovine serum (Sigma-Aldrich) and 1% penicillin/streptomycin (Gibco)).

Live-cell imaging was initiated 15-30 minutes after removal of virus-containing medium.

For experiments in which the diffusion of translating mRNA molecules was assessed, U2OS cells stably expressing TetR, and STAb were seeded at 45% confluency and reverse transfected with a plasmid expressing the reporter mRNAs using Fugene (Promega) 1 day before imaging. For experiments in which the GFP fluorescence intensity of individual 24xSunTag arrays was measure, STAb U2OS cells were seeded at 45% confluency and reverse transfected using Fugene (Promege) 1 day before imaging with a plasmid encoding the 24xSunTag-CAAX protein. One hour before the start of imaging, the cell culture medium was replaced with imaging medium. All live-cell imaging experiments were performed at 37°C.

2D-culture of organoid cells for live-cell imaging

To enable imaging of single translating vRNAs in intestinal organoid cells, organoid cells were cultured in 2D. 96-well glass-bottom plates (Matriplates, Brooks Life Science Systems) were coated for 30 minutes at 37°C with 100 µl coating mix (5% Basement Membrane Extract (Type 2, Cultrex reduced growth factor, R&D systems) dissolved in Advanced DMEM (Thermofisher Scientific)). The coating mix was removed and the wells were dried for 15 minutes at 37°C. Organoids grown in 3D were dissociated into single cells by incubating the organoids for 5 minutes with TrypLE (TrypLE Express; Life Technologies). Approximately 5000 cells were plated per well in expansion medium (see (Sato et al., 2011)). Virus incubation was performed as described in the section 'Cell culture for imaging'. Expansion medium was replaced with imaging medium (pre-warmed CO_2-independent Leibovitz's-15 medium (Gibco) containing 5% fetal bovine serum (Sigma-Aldrich) and 1% penicillin/streptomycin (Gibco)) 15-30 minutes before the start of live-cell imaging.

Drug treatment

In experiments in which cells were treated with IFN, 1000 U/ml IFNα2 (Sigma) was added approximately 24 hr before imaging. Virus-containing medium and imaging medium did not contain IFN. Where indicated, cells were treated with the inhibitors 3Ci: Rupitrivir (Sigma, 10 µM) or 3Di: GPC-N114 ((van der Linden et al., 2015), 10 µM) 30-60 minutes before addition of virus-containing medium and treatment of these drugs continued during incubation with virus-containing medium and during live-cell imaging.

To determine cell survival during time-lapse imaging, Propidium Iodide (PI, 20 µg/ml; Sigma) was added to the imaging medium.

To investigate viral translation dynamics, translation inhibitors puromycin (0.1 mg/ml; ThermoFischer Scientific) or harringtonine (3 µg/ml; Cayman Chemical) were added to the imaging medium at indicated time-points.

Image acquisition of fixed cells

Unless noted otherwise, imaging of fixed cells was performed using the 60x 1.40 NA objective. The x,y positions for imaging were selected either randomly or, in the cases

where smFISH and/or immunofluorescence were performed after live-cell imaging, the same x,y positions were used as those imaged during live-cell imaging. Images were acquired using 70 ms exposure. For GFP, Cy5, and Atto633, 15-20 Z-slices with 0.5 μm steps were acquired. For cells expressing NLS-BFP, a single Z-slice focused in the middle of the cell was imaged for BFP.

The 100x 1.49 NA objective was used to image fixed cells in which the methionine analog was stained. The x,y positions were selected randomly. Large-images of 6 to 6 (control conditions) or 9 to 9 (all other conditions) regular FOVs were generated. For BFP, GFP, and Cy5, images of 10-15 Z-slices with 0.8 μm steps were acquired using 100 ms exposure.

QUANTIFICATION AND STATISTICAL ANALYSIS
Post-acquisition processing of microscopy data

For all acquired Z-slices, maximal intensity projections were generated using NIS elements software. All downstream analyses were performed on the maximal intensity projections. For experiments in which fluorescence intensities were measured, correction for photobleaching was performed, unless noted otherwise, for each color individually, using the ImageJ plugin *Bleach correction,* using the *Exponential fit* option. In brief, the fluorescence intensity over time of the entire FOV was determined and fit with an exponential decay distribution. The bleaching rate was then calculated based on the exponential decay fit and was used to correct the fluorescence intensities of images at all time-points.

Counting the number of translating vRNAs per cell

In live-cell imaging experiments in which the duration of infection phases was determined, cells were selected for analysis based on the following criteria:

Only cells that were completely in the FOV for the entire duration of the analysis were included.

Maximally 5 cells per FOV were analyzed.

Cells with multiple translating vRNAs in the first time point of the movie were excluded from analysis.

If a cell undergoes mitosis during the movie, the cell is excluded from the analysis.

The number of translating vRNAs per cell was determined manually for each time point, starting from the first time point that a translating vRNA could be observed, until either individual vRNAs could no longer be distinguished anymore due to the large number of vRNAs in the cell or until the number of translating vRNAs exceeded 50 vRNAs/cell. All bright GFP spots were considered individual translating vRNAs. However, the following additional criteria were taken into account during counting of translating vRNAs:

Translating vRNAs are highly mobile (Fig. 1K, S1M) and a single vRNAs can therefore often be observed in multiple slices of a Z-stack at different x,y coordinates, resulting in two spots close together in the maximal intensity projection. To prevent double counting of the same translating vRNA in the maximal projection, two spots positioned very close together were counted as a single translating vRNA.

In many STAb-expressing cells, we also observed one or two distinct GFP foci adjacent to the nucleus, most likely indicating that the STAb bound with weak affinity to the centrosomes. These foci were independent of virus infection, as similar foci could also be observed in uninfected cells. These foci were excluded from all analyses.

Some cells contain GFP foci that were larger and far less mobile than translating vRNAs. These foci were observed in both infected and uninfected cells, and thus represent background foci. These background foci often co-localized with mCherry positive foci, indicative of autofluorescence. A combination of morphology, mobility, and dual positivity with mCherry was used to exclude these background foci during counting of translating vRNAs.

Of note, in experiments in which the GFP fluorescence intensity of translating vRNAs were measured, images were acquired for only a single time-point. The mobility of background foci could therefore not be taken into account to exclude background foci from analysis. However, we found that the morphology, size, mCherry fluorescence co-localization and subcellular location of such background foci was still sufficient to reliably distinguish the background foci from translating vRNAs.

Infection phases

Annotation of infection phases

Based on the number of translating vRNAs over time, infection phases 1 to 5 were called. Phase1 is defined as the period from the first detection of a translating vRNA until the disappearance of the translating vRNAs; phase2 is the period from disappearance of the first translating vRNA until the appearance of multiple translating vRNAs. The period during which the number of translating vRNAs increased after phase2 is referred to as phase3. Once the number of vRNAs plateaus, phase4 is called. Then, a second increase in the number of translating vRNAs is observed, which is referred to as phase5.

Of note, in some cases we observed a single translating vRNAs, which disappeared after some time, followed by the re-appearance of a single translating vRNA (rather than the appearance of multiple vRNAs, characteristic of phase2 followed by phase3). Such disappearance and reappearance of a single translating vRNA could be observed multiple times in rare cases (called phase1 pulses and breaks; see for example Fig. 3B, S3C). In these cases, phase1 is defined as the entire period from the first detection of a translating vRNA until the end of the last phase1 pulse (i.e. the last pulse before phase2), as indicated with an arrow in Fig. 3B, S3C. In analyses specifically focusing on these phase1 pulses, the timing of individual phase1 pulses and breaks was analyzed.

Since a small amount of noise (i.e. variations in the number of detected foci between adjacent time-points) was observed in the calling of the number of translating vRNAs per time-point (likely due to imperfect detection of vRNAs), the following additional criteria were taken into account for phase annotation:

Phase1, phase1 pulses, phase1 breaks, and phase2 were each called only if their duration was at least 12.5 minutes. Depending on the timing interval, 3-12 consecutive time points are needed to fulfill the 12.5 minutes requirement. A gap in phase 1 (i.e. one or multiple consecutive time-points without detection of a

translating vRNA) could be the result of noise in counting of translating vRNAs or a phase1 break. To test whether a 12.5 minutes threshold is sufficient to distinguish 'counting noise' from true biological changes in the infection phase, we analyzed the frequency and the number of consecutive time-points without a translating vRNA during phase1. For each cell, the frequency and duration of gaps of less than 12.5 minutes was determined (Fig. S3E). The majority of gaps consisted of a single time-point, indicating noise in translating vRNA counting. Gaps of 12.5 minutes or longer due to noise would not be expected based on the distributions of gap duration, but instead reflect a biological change in the number of GFP foci. A threshold of 12.5 minutes can thus be used to adequately annotate phase1, phase2, phase1 pulses, and phase1 breaks.

A moving average of the number of translating vRNAs over time was used to approximate both the end of phase3 and the start of phase5. Then, the raw data on the number of translating vRNAs over time was used to determine the exact first or last time-point. The end of phase3 is the last time-point that an increase relative to the previous time-point was observed. The start of phase5 is the first time-point that an increase relative to the previous time-point was observed.

Phase4 consists of the time points between the last time-point of phase3 and the first time-point of phase5.

Analysis of infection phases

For each infected cell, the duration of each infection phase was determined. Additionally, we determined whether the infection phase was complete (i.e. the start and end of an infection phase were observed during the movie). The distributions of infection phase durations are plotted as Kaplan-Meier plots, which allows inclusion of both complete and incomplete infection phases (for example Fig. 2H-K). In these Kaplan-Meier plots, the timing of an incomplete infection phase is plotted as a circle. For phase 3, 4, and 5 additional infection phase parameters were determined; the number of translating vRNAs per cell over time was plotted, and, for phases3 and 5, linear regression analysis of performed to determine the rate of vRNA synthesis. The mean number of translating vRNAs during phase4 was calculated as the average of all time points during phase4.

Independent repeats

Analyses of the durations infection phases were performed based on many (2-21) independent experiments. However, the number of cells per experiment varied greatly. Therefore, we did not display the mean values of the individual repeats, as experiments with a low number of analyzed cells make an inappropriately large contribution to this mean value. Instead, all analyzed cells were combined into one Kaplan-Meier graph. To verify that individual cells from individual repeats could be combined into one dataset, we confirmed the reproducibility of the infection phases by comparing the combined datasets to individual repeats. As some repeats contained a low number of analyzed cells, some repeats were combined into datasets of at least 30 cells (Fig. S2A-D).

Analysis of cells with multiple infections

Unless stated otherwise, experiments were performed using a low MOI (0.25) to ensure that single viruses infected individual cells. However, even at a low MOI, double infections (i.e. two viruses infecting a single cell) were observed in a small number of cells (~10.6 %). A double infection was called when a single cell contained two GFP foci for at least 12.5 min. To simply analysis of infection phases, double infected cells were excluded from all analyses, except from the experiments using a high (1.25) MOI (Fig. 2Q, S3M, N). For high MOI experiments double infections could be observed in many cells. To determine the duration between initial infection and translation of newly synthesized vRNAs, the time between first detection of a translating vRNA and rapid increase in translating vRNAs was determined (arrows in Fig. S3N). This time is referred to as combined duration of phase1+2.

Analysis to exclude that phase1 pulses are a result of double infections

Even though most of our experiments were performed using a low MOI (0.25), limiting the fraction of cells with double infections, in some infected cells we did observe double infections (i.e. multiple translating vRNAs present simultaneously during phase1; also see section 'Analysis of cells with multiple infections'). In addition to double infected cells, phase1 pulses were observed in some infected cells (Fig. 3C, S3C). In theory, two translation pulses in a single cell could represent either a single vRNA turning translation off and on again, or pulses could be the result of two sequential infections in a single cell. To examine whether phase1 pulses are likely to be due to double infections, we calculated the fraction of cells expected to be infected by multiple viruses (Ellis and Delbrück, 1939), the fraction of cells observed to be infected by multiple viruses, and the fraction of cells with translation pulses. If for example, 10% of cells are expected to contain double infections, but no double infections are called and 10% of infections are called as containing translation pulses, it is likely that such pulses in fact represent double infections. In contrast, if 10% of cells are expected to contain double infection and 10% of infections are called as double infections, while in addition another 10% of infections show pulses, we conclude that such pulses likely do not represent double infections, as we conclude that we are calling the expected number of double infections, and can thus accurately identify double infections.

The expected number of double infections can be calculated using the following equation (Ellis and Delbrück, 1939):

$$P(n) = \frac{m^n \times e^{-m}}{n!}$$

P(n) is the expected fraction of cells infected by n viruses using an MOI of m. With an MOI of 0.25, 77.9%, 19.5% and 2.6% of all cells is expected to be uninfected, infected by a single virus, or infected by >1 virus respectively. Thus, 11.7% (=2.6/(19.5+2.6)) of all infected cells is expected to be infected by >1 virus (Fig. S4A). We called double infections in 10.6% of the infected cells (Fig. S4A), a value very similar to the expected value. In addition, we called translation pulses in 11.4% of cells. This analysis thus strongly suggests that double infections are not primarily responsible for the observed translation pulses.

Correlation between durations of different infection phases

To investigate whether successful replication during phase2 (i.e. transition from phase2 to phase3) correlated with the duration of phase1 (Fig. 3A), all infected cells were subdivided based on the duration of phase2. Cells with successful replication were defined as cells in which phase2 is equal to or less than 60 minutes. If the end of phase2 was not observed, the infection was categorized as successful replication. We excluded cells for which we could not determine whether phase2 was successful; for example, if phase2 was incomplete and shorter than 60 minutes (for example, due to the end of the movie).

For analyses in which the duration of phase1+2 was compared to the duration of phase3+4 (Fig. 2Q), only cells were included in which both relevant infectious phases were completely observed (i.e. start and end observed). As a result, cells for which the end of phase2 was not observed (i.e. mostly cells with unsuccessful replication) are excluded from this analysis. The comparison between the combined duration of phase1+2 and phase3+4 was made based on the median of the distributions, instead of the average of the distributions, as the phase1+2 duration has a long tail towards longer times. This long tail is the result of the larger heterogeneity in phase1 duration (Fig. 2Q). During phase1+2, there is only a single vRNA that can undergo replication, so a delay for that vRNA molecule causes an increased translation+replication time. In contrast, during phase3+4 multiple newly synthesized vRNAs are translated and replicated. If only a single one of those vRNAs replicates successfully, new vRNAs are observed and phase3+4 is completed (start of phase5 is defined as the appearance of additional GFP foci). Thus, even if one vRNA is delayed in replication, another vRNA can still replication resulting in the completion of phase3+4. As a result, the shape of the distribution of phase3+4 duration is narrower compared to phase1+2 duration. Since we wanted to examine the duration of translation+replication of a 'typical' vRNA in phase1+2 with a vRNA in phase3+4, we chose the median in this analysis to prevent artificial skewing of the value by the outliers present only when a single vRNA is present in the cell.

Quantifying the number of smFISH foci per cell

Single vRNAs

To determine the number of vRNAs in single infected cells (Fig. 1E, I, 2E, S1F, G, S2D), the number of CVB3 smFISH spots in infected cells was determined. Loss of nuclear NLS-BFP was used as a read-out of cell infection. Spots with an intensity of over 2.5-fold of the median intensity of all spots were excluded from the analyses, as they represent sites of replication and do not represent single vRNAs (See section 'Quantifying the number of smFISH foci per cell-Replicating vRNAs'). For each replicate experiment the average number of CVB3 smFISH spots in uninfected cells (i.e. cells without loss of nuclear NLS-BFP) was determined (range 2 – 7 spots; see Fig. S1G) and this number was subtracted from the spot count of infected cells. Note that these 'background foci' are not observed if no virus incubation step was performed (Fig. S1F). These background foci therefore likely represent (intact) CVB3 particles adhering to the outside of the cell or in endosomes.

Replicating vRNAs

Based on the fluorescence intensity of smFISH foci, two types of smFISH foci were defined: 1) single vRNAs with a relatively low and homogenous distribution in fluorescence intensity, and 2) replicating vRNAs with >2.5x the fluorescence intensity of single vRNAs (Fig. S2D). The fluorescence intensity of single vRNA spots and replicating vRNAs was normalized to the mean of three smFISH foci in each cell.

Calculating mobility of translating RNAs

For analysis of the mobility of translating RNAs (either vRNAs or mRNAs), the ImageJ plugin *TrackMate* (v3.7.0) was used. *Manual Tracking* was used to generate tracks of single RNAs in Fig. 1J and K. Based on these tracks, the mean square displacement (MSD) was calculated for each RNA.

In a separate experiment, the mobility of translating vRNAs was determined in cells with more than one translating vRNA (Fig. S1M). The tracks were generated semi-automatically, using the *LoG detector* option of TrackMate. Spots were identified after applying a Gaussian filter and the intensity threshold was set manually. Next, tracks were generated using the *LAP tracker* option, allowing a maximal linking distance of 1 µm and up to 5 frame gaps in the track. All generated tracks were manually inspected. Tracks with less than 10 time-points were excluded from further analysis. Finally, the diffusion coefficient of each trace was calculated by fitting the first 10 time-points of the MSD curve.

Determining the fraction of translating immobilized vRNAs

To classify translating vRNAs based on their mobility, we visually determined whether a translating vRNA was mobile or immobile in 5 consecutive time points with 2 min interval. If a translating vRNA was visible in each time point at the same location, the vRNA was classified as immobile. The fraction of immobile vRNAs per cell was then calculated based on the number of immobilized sites and the total number of vRNAs for each cell as determined by the mean number of vRNAs counted over the 5 time points. If no immobilized vRNAs were observed, the fraction of immobilized vRNAs in that cell was classified as '0'. All cells, with and without immobilized vRNAs, were combined to calculate the mean fraction of translating immobilized vRNAs (Fig. S1O).

Translating vRNA intensity and vRNA translation efficiency

Fluorescence intensity of a translating vRNA

To analyze the translation efficiency of vRNAs over the course of infection, fluorescence intensities of translating vRNAs were measured manually in different cells with variable number of translating vRNAs. The fluorescence intensity was measured in an ROI of 8x8 pixels centered on the translating vRNA. In cells with less than 10 translating vRNAs, the intensity of all vRNAs was measured. In cells with more than 10 translating vRNAs, the intensity of 10 vRNAs was measured by pre-selecting a region of the cell and measuring all vRNAs within this region. For each translating vRNA, a local background subtraction was performed based on the mean

of the intensity of 3 random ROIs near the vRNA.

Analyzing GFP foci intensity

The fluorescence intensity of a translating vRNA depends on the number of nascent SunTag peptides (and thus the number of ribosomes) associated with the vRNA. To determine the number of nascent SunTag peptides per translating vRNA, we aimed to determine the precise number of GFP molecules associated with each GFP spot. To this end, the GFP fluorescence intensity of translating vRNAs was compared to the GFP fluorescence intensity of a SunTag 'stoichiometry reporter' with a known number of SunTag peptides. As it is difficult to unambiguously detect SunTag proteins with few SunTag peptides, a 24xSunTag-peptide array was selected, for which single molecules can be readily observed after expression in STAb cells. To further facilitate observation of the stoichiometry reporter, a C-terminal prenylation (CAAX) sequence was added to the 24xSunTag-peptie array to tether the proteins to the plasma membrane. Membrane tethering reduces mobility and limits localization to a 2D plane (when focusing on the bottom of the cell), facilitating intensity quantification. High expression of the SunTag-peptide arrays can interfere with unambiguous identification of single SunTag peptide arrays. Therefore, the highly repressive Emi 5'UTR (Tanenbaum et al., 2015) was used to ensure low expression of the stoichiometry reporter and only cells with low expression of the 24xSunTag-CAAX (<100 foci visible on bottom plasma membrane) were included in the analyses. Using identical imaging parameters as used to determine translating vRNA intensity, short time-lapse movie (<10 time-points; 25ms interval) were acquired of selected cells focused on the bottom of the cells. Only spots that were visible during 5 consecutive time-points were included in the analysis. The 24xSunTag-CAAX intensity was measured in an ROI of 8x8 pixels centered on the spot in the first time-point of the time-lapse. For each focus, a local background subtraction was performed based on the mean of the intensity of 3 random ROIs near the spot. The relative fluorescence intensity corresponding to a single SunTag peptide was calculated by dividing the mean fluorescence intensity of 24xSunTag-CAAX by 24. The fluorescence intensity of single translating vRNAs was measured as described in section 'Fluorescence intensity of a translating vRNA'. The number of nascent SunTag peptides associated with each vRNA was then determined based on the fluorescence intensity of a 1x SunTag and of the translating vRNA.

SunTag labeling stoichiometry during infection

The fluorescence intensity of a translating vRNA depends on the number of nascent peptides per vRNA and the intracellular concentration of unbound STAb. During infection of a SunTag-encoding virus, many SunTag peptides are synthesized, which recruit STAb. The amount 'mature' SunTag peptides may therefore affect the unbound cellular STAb concentration and thereby the translating vRNA intensity. Reduced brightness of vRNA foci due to limiting STAb availability may hamper analysis of the number of translating vRNAs and may complicate the analysis of translating vRNA intensities. To examine whether changes in the concentration of available STAb occur during early infection phases, the SunTag labeling stoichiometry in infected cells was determined. If the concentration of available STAb drops below a threshold,

the recruitment of STAb to the SunTag peptide array is reduced, leading to an altered fluorescence intensity of the SunTag peptide array. To this end, the fluorescence intensity of the stoichiometry reporter (24xSunTag-CAAX; see section 'Analyzing GFP foci intensity') was measured in infected cells, as described in 'Analyzing GFP foci intensity'. In the same cells, the number of translating vRNAs was quantified. The fluorescence intensity of 24xSunTag-CAAX was normalized to the mean intensity in uninfected cells.

Translation elongation rates

To compare the viral translation elongation rate at different time-points during infection, we performed harringtonine run-off experiments (Yan et al., 2016). The fluorescence intensity of single translated vRNAs was measured manually in an ROI of 5x5 pixel on the vRNA. For each vRNA a local background correction was performed by subtracting the local background fluorescence intensity. Translating vRNAs that showed an increasing intensity prior to harringtonine administration were excluded from analysis, as these translating vRNAs were likely not yet in steady-state translation. Where possible, translating vRNAs under the nucleus were selected for intensity measurements, as these vRNAs have low mobility and are thus more easily tracked over time. The intensity was measured for at least 10 minutes after harringtonine administration. If a vRNA could not be detected anymore, the ROI was positioned at the location where the translating RNA was last detected. The fluorescence intensity of each translating vRNA was normalized to the mean intensity of the same vRNA in the 2 minutes before harringtonine administration.

Measuring intensities of fluorescence reporters

eIF4G cleavage reporter

Cleavage of the eIF4G reporter was determined by the dissociation of BFP from mitochondria (as assessed by the mitochondrially-localized mCherry signal). The cleavage of the eIF4G cleavage reporter was quantified only in cells in which the start of phase 1 could be observed. In each of these cells, the BFP fluorescence intensity was measured in 3 ROIs (of 5x5 pixels) in the cytoplasm and in 3 ROIs on the mitochondria. In experiments with relatively short expression of the eIF4G cleavage reporter (12 hr doxycycline induction), the fluorescence intensities of BFP and mCherry increased over time, suggesting that expression of the eIF4G cleavage reporter had not yet reached steady state. To correct for variations in expression of the reporter over time, the BFP intensities were normalized based on the expression of mCherry. We first determined the fluorescence change in mCherry signal over time by calculating the mean mCherry intensity of 3 ROIs (5x5 pixels) on the mitochondria. The (auto fluorescence) background - the mean mCherry intensity in the cytoplasm was also measured in cells that did not express the eIF4G cleavage reporter – was determined and subtracted. Second, a correction factor for expression of the reporter was calculated for each time-point by dividing the mCherry intensity of that time-point by the mean mCherry intensity of the first five time-points of the movie. Third, the BFP fluorescence intensities of each time-point were then corrected by dividing the measured BFP intensities by the time-point-specific correction factor. Then, the difference between the mitochondrial

and cytoplasmic BFP intensities (after correction) was calculated by subtracting the cytoplasmic intensity from the mitochondrial intensity. Finally, the difference in BFP intensity over time was normalized to the mean difference in BFP intensity of the three time-points before the start of phase 1. The difference in BFP intensity between mitochondria and cytoplasm over time was also calculated for cells that were not infected (i.e. no translating vRNAs were observed). As the start of phase 1 could not be assigned in these cells, the difference in BFP intensity could not be normalized based on the start of phase 1. Instead, the difference in BFP intensity of control cells were aligned and normalized to the start of the movie.

Nuclear leakage of NLS-BFP

Loss of nuclear NLS-BFP during infection was analyzed in cells in which the start of phase 1 could be observed. The BFP fluorescence intensity in 3 ROIs (of 5x5 pixels) in the nucleoplasm and 3 ROIs in the cytoplasm was measured. Then, the mean cytoplasmic BFP intensity was subtracted from the mean nuclear BFP intensity. Finally, the difference in BFP intensity over time was normalized to the mean difference in BFP intensity of the three time-points before the start of phase 1. The difference in BFP intensity over time was also calculated for cells that were not infected (i.e. no translating vRNAs were observed). As the start of phase 1 could not be assigned in these cells, the difference in BFP intensity could not be normalized based on the start of phase 1. Instead, the difference in BFP intensity of control cells were aligned and normalized to the start of the movie.

Measurements of global host translation

All cells with translating vRNAs were included in the analysis. Additionally, at least 20 uninfected cells from the same images were included as controls. Mitotic cells were excluded from the analysis. The HPG-associated Cy5 fluorescence intensity (referred to as 'methionine fluorescence') of each cell was determined by calculating the mean intensity of 3 20x20 pixel ROIs in the cytoplasm of cells. If a cell crossed the boundaries of the large-image stich, the stitching area was excluded from measuring fluorescence intensities, as intensities were generally lower in these regions. Fluorescence background was measured in cells that were not subjected to HPG treatment. The methionine fluorescence intensities of infected cells were normalized to the mean intensity of uninfected cells (which was set to 1) and the mean intensity of puromycin treated cells (set to 0).

Determining the moment of infection

To determine the fraction of infected cells over time, the moment of infection was determined, based on the start of phase1. In some experiments, like when wild-type CVB3 was used, the first appearance of a translating vRNA could not be used to determine the start of phase1. In those instances, the moment of loss of nuclear NLS-BFP was used to determine the start of infection, because the start of phase 1 coincides with the loss of nuclear NLS-BFP (Fig. 1G, H, Movie S1-3). The time between inoculation with virus-containing medium and the first detection of loss of NLS-BFP was used. Cells that already displayed (partial) nuclear leakage at the first

I

II

III

IV

V

VI

&

time-point of the movie were excluded from analysis, as these cells were already infected before the start of the movie. In a separate analysis, the fraction of cells with nuclear leakage in the first time-point was determined to calculate the fraction off pre-imaging infected cells infected before the start of the movie. The Kaplan Meier plot with the fraction of uninfected over time was then corrected for pre-imaging infection by using the fraction of pre-imaging infected cells as starting point of the Kaplan-Meier plot.

Determining the moment of cell death

Movies acquired with propidium iodide in the imaging medium were analyzed to determine cell survival after infection. Per FOV, maximally 10 cells were selected randomly for analysis. Only cells that were not yet infected at the start of the movie were selected. Cells that went through mitosis during the movie were excluded from analysis. For each cell we scored whether the cell became infected or not during the movie, based on nuclear leakage of NLS-BFP. Then, we determined whether the infection lead to successful replication (i.e. phase3) within 6 hr of the start of phase1. If an infected cell moved out of the FOV within 6 hr after infection, the cell was excluded from analysis, as we could not determine whether or not the infection led to successful replication. Third, for each cell the moment of cell death was determined based on PI-positivity. If a cell did not die, or moved out of the FOV before dying, the cell was classified as survivor and the last time-point that the cell could be observed was used to determine the survival time. Finally, the fraction of surviving cells over time since infection was plotted in a Kaplan-Meier plot. For control cells that did not get infected, the survival was plotted relative to the start of the movie.

siRNA screen

To facilitate high-throughput analysis of viral replication success, we designed a rapid analysis pipeline. If an infection leads to successful replication, many new vRNAs are produced and translated, resulting in the synthesis of many more viral proteins and SunTag peptides (in case of SunTag-CVB3 virus). The SunTag peptides are located in the cytoplasm, while the STAb is located in both the nucleus and the cytoplasm. Therefore, production of high levels of viral SunTag protein results in recruitment of all STAb from the nucleus to the cytoplasm. During normal infection, this occurs in cells around ~2 hr after initiation of phase3 (See for example Movies S1, 3). Importantly, nuclear depletion of STAb is not observed in infected cells with unsuccessful replication (i.e. arrest during phase2; see Movies S2 and S3). Loss of nuclear GFP signal can be readily observed during time-lapse imaging, even when using a low magnification imaging objective, and is a robust read-out for successful vRNA replication. Usage of a low magnification objective allows for imaging of many more conditions simultaneously, and thus is especially useful for a quick screen.

All cells that were infected in the first 2 hr of a movie, as determined by the leakage of NLS-BFP from the nucleus, were included in the analysis. For each infected cells, we determined whether the infection led to successful replication, as determined by depletion of nuclear STAb, within 6 hr after the start of phase1. The fraction of infected cells with successful replication was then calculated by dividing the number of infected cells with successful replication by the total number of infected cells.

To assess the accuracy of using STAb nuclear loss as a read-out for successful viral replication, we determined the fraction of infected cells with successful replication with or without 24 hr IFN pre-treatment (data plotted in black in Fig. 6B). The fraction of infected cells with successful replication is 0.78 ± 0.05 (mean ± SD; n=7) and 0.31 ± 0.05 (mean ± SD; n=5) in the absence or presence of IFN treatment, respectively. The obtained values for the fraction of successful viral replication based on either loss of nuclear STAb or on single-molecule analysis of infection phases was comparable: without treatment 0.78 vs. 0.85 (no IFN) or 0.31 vs. 0.41 (with IFN). Note that the rapid analysis pipeline resulted in a slightly lower fraction of successful replication than the single-molecule VIRIM analysis, possibly due to too stringent scoring of loss of nuclear STAb. Thus, the analysis pipeline based on nuclear loss of STAb can be used to screen various conditions.

Analysis of RNA sequencing

To identify IFN-induced genes, cells with or without 24 hr IFN pre-treatment were compared. For each sample, paired-end reads from illumina sequencing were aligned to the human transcriptome, as described previously (Muraro et al., 2016). Genes with less than 10 reads in the datasets from the cells without IFN treatment were excluded from analysis. To enable comparison between the different samples, the read-count per gene was normalized to the average of the total read count of all samples (mean total read count per sample: 4535455). To determine the fold-induction of gene expression after IFN treatment, the ratio in gene expression in untreated and IFN pre-treated cells was calculated. All genes with at least 3-fold increase in expression upon IFN treatment in both experiments were selected (37 genes). To simplify the siRNA screen, a further selection of genes was made by excluding genes with well-described antiviral functions that are unrelated to inhibition of viral replication (9 genes excluded) and by excluding genes for which no SMARTpool ONTargetPlus reagents were available (1 gene excluded). As a result, 28 genes were included in the siRNA screen (for details on sequencing results and fold increase in expression after IFN treatment per gene see online data).

Statistical analyses and generation of graphs

All statistical analyses graphs were performed and generated using Prism GraphPad (v8.2.1). Statistical tests are explained in figure legends.

To facilitate visual inspection of general trends in heterogeneous data, moving averages were included in Fig. 1I, 4B, C, H, K. Moving averages were generated as locally weighted least squares (LOWESS) using the 'fit spline/lowess' function in GraphPad and using a 5 point smoothening window.

To determine the mean duration of phase1 and phase2, the distribution of phase1 and phase2 durations were fit to a one-component exponential decay distribution. During phase annotations, a minimal threshold of 12.5 minutes was used to annotate phase1 and phase2. Therefore, additional constrains (y=1 and x=0-12) were taken into account to generate the one-component exponential decay fits of phase1 and phase2.

A fraction of cells did not complete phase2, which was referred to as the fraction of

infected cells with 'unsuccessful replication'. To determine the fraction of cells with unsuccessful replication, the one-component exponential decay fit of phase2 was generated and the plateau of the fit was used as the fraction of cells with unsuccessful replication.

SUPPLEMENTAL ITEMS

Supplemental Movies S1-4 and Table S2 are available via https://doi.org/10.1016/j.cell.2020.10.019.

Table S1 – related to all figures. Number of experimental repeats, cells, and *mRNAs analyzed per experiment*.

Overview of the number of repeats per experiment, and of the number of cells and foci analyzed per experiment. Some datasets are plotted multiple times. If relevant replotting of the same datasets is indicated in the last column and in the corresponding figure legends.

Fig.	panel	Description	color	repeats	datapoints	units	replotting
1	1E	live-cell imaging and smFISH		4	93	cells	1I; 2E
	1F	untreated	black	17	57	cells	-
		3Di treated	red	2	36	cells	-
	1H	non-infected	black	3	23	cells	-
		infected	red	3	26	cells	-
	1I	live-cell imaging and smFISH Sun-Tag-CVB3	green	4	93	cells	1I; 2E
		live-cell imaging and smFISH wildtype CVB3	black	3	121	cells	-
	1K	viral RNA	black	3	24 & 24	cells & spots	-
		mRNA	red	2	10 & 28	cells & spots	-
S1	S1B	qPCR HeLa wildtype CVB3	black	3			-
		qPCR HeLa eGFP-CVB3	brown	3			-
		qPCR HeLa SunTag-CVB3	green	3			-
	S1C	qPCR U2OS wildtype CVB3	black	3			-
		qPCR U2OS eGFP-CVB3	brown	3			-
		qPCR U2OS SunTag-CVB3	green	3			-
	S1D	qPCR SunTag- U2OS wildtype CVB3	black	3			-
		qPCR SunTag- U2OS eGFP-CVB3	brown	3			-
		qPCR SunTag- U2OS SunTag-CVB3	green	3			-
	S1E	nr. of nascent SunTag peptides per vRNA	green	3	45 & 411 40 & 576	cells & vRNA spots cells & ST-CAAX spots (for normalization)	S3K
	S1G	smFISH 'background' per repeat - no virus incubation		3	28	cells	-

		smFISH 'background' per repeat - uninfected		3	106	cells	-
	S1H	smFISH and 3D IF SunTag-CVB3	green	3	48	cells	-
		smFISH and 3D IF wildtype CVB3	black	3	39	cells	-
	S1J	live-cell imaging and dsRNA staining	green	3	68 37	cells non-infected cells (for background correction)	-
	S1L	eGFP fluorescence and dsRNA staining	brown	2	199	cells	-
	S1M	diffusion cells with 1 spot		4	27 & 27	cells & spots	-
		diffusion cells with 3-20 spots		2	19 & 61	cells & spots	-
		diffusion cells with 25-40 spots		3	23 & 152	cells & spots	-
		diffusion cells with >45 spots		2	20 & 122	cells & spots	-
	S1O	immob fraction in cells with 3-20 spots		3	62	cells	-
		immob fraction in cells with 25-40 spots		3	62	cells	-
		immob fraction in cells with >45 spots		3	94	cells	-
2	2E	smFISH foci phase1	purple	4	12	cells	1E; 1I
		smFISH foci phase2	blue	4	23	cells	1E; 1I
		smFISH foci phase3	green	4	19	cells	1E; 1I
		smFISH foci phase4	orange	4	22	cells	1E; 1I
		smFISH foci phase5	yellow	4	18	cells	1E; 1I
	2F	replication sites phase1	purple	4	12	cells	-
		replication sites phase2	blue	4	23	cells	-
		replication sites phase3	green	4	17	cells	-
		replication sites phase4	orange	4	21	cells	-
		replication sites phase5	yellow	4	23	cells	-
	2H	duration phase1 U2OS	black	21	183	cells	3F; 5B; 6E; S3A; S4B
		duration phase1 HeLa	brown	2	66	cells	-
		duration phase1 A549	blue	3	35	cells	-
		duration phase1 intenstinal cells	grey	2	15	cells	-
	2I	duration phase2 U2OS	black	21	176	cells	3G; 5C; 6D; 6I; S3B
		duration phase2 HeLa	brown	2	66	cells	-
		duration phase2 A549	blue	3	35	cells	-
		duration phase2 intenstinal cells	grey	2	16	cells	-
	2J	duration phase3 U2OS	black	21	112	cells	5D; S3C; S4L
		duration phase3 HeLa	brown	2	55	cells	-
		duration phase3 A549	blue	3	28	cells	-
		duration phase3 intenstinal cells	grey	2	13	cells	-
	2K	duration phase4 U2OS	black	21	100	cells	5E; S3D; S4M
		duration phase4 HeLa	brown	2	55	cells	-

		duration phase4 A549	blue	3	26	cells	-
		duration phase4 intenstinal cells	grey	2	13	cells	-
	2L	plateau phase4 U2OS	black	21	100	cells	5G
		plateau phase4 HeLa	brown	2	55	cells	-
		plateau phase4 A549	blue	3	26	cells	-
		plateau phase4 intestinal cells	grey	2	13	cells	-
	2M	phase2 arrest U2OS	black	21	6 & 183	datasets & cells	S3A
		phase2 arrest HeLa	brown	2	66	cells	-
		phase2 arrest A549	blue	3	35	cells	-
		phase2 arrest intestinal cells	grey	2	52	cells	-
	2N	survival non-infected cells	black	3	160	cells	5F
		survival cells with successful replicaiton	green	3	120	cells	5F
		survival cells with unsuccessful replication	red	3	29	cells	5F
	2Q	combined duration phase1 and 2; MOI 0.25	purple	21	135	cells	-
		combined duration phase1 and 2; MOI 1.25	light blue	4	81	cells	-
		combined duration phase3 and 4; MOI 0.25	green	21	75	cells	-
S2	S2B	phase0		21	107 & 3189	cells & time-points	-
		phase1		21	116 & 3502	cells & time-points	-
		phase2		21	111 & 2748	cells & time-points	-
	S2C	SunTag labeling stochiometry		4	149 & 54	cells & uninfected cells (for normalization)	-
	S2D	smFISH foci		4	53 & 159	cells & smFISH foci	-
		replication sites		4	51& 51	cells & replication sites	S2E; S2F
	S2E	replication site intensity phase2	blue	4	2 & 2	cells & replication sites	S2D; S2F
		replication site intensity phase3	green	4	8 & 8	cells & replication sites	S2D; S2F
		replication site intensity phase4	orange	4	17 & 32	cells & replication sites	S2D; S2F
		replication site intensity phase5	yellow	4	23 & 143	cells & replication sites	S2D; S2F
	S2F	replication site intensity control	black	4	49 & 185	cells & replication sites	S2D; S2E
		replication site intensity 3Dpol inhibitor	pink	3	21 & 48	cells & replication sites	-
	S2G	replication sites control	black	4	91	cells	2F
		replication sites 3Dpol inhibitor	pink	3	77	cells	-
S3	S3A	duration phase1 combined into 1 dataset	black	21	183	cells	2H; 3F; 5B; 6E; S4B

		duration phase1 split into 6 datasets		21	30-31	cells	-
	S3B	duration phase2 combined into 1 dataset	black	21	176	cells	2I; 3G; 5C; 6D; 6I
		duration phase2 split into 6 datasets		21	30-31	cells	2M
	S3C	duration phase3 combined into 1 dataset	black	21	112	cells	2J; 5D; S4L
		duration phase3 split into 6 datasets		21	30-31	cells	-
	S3D	duration phase4 combined into 1 dataset	black	21	100	cells	2K; 5E; S4M
		duration phase4 split into 6 datasets		21	30-31	cells	-
	S3K	nr. of nascent SunTag peptides per vRNA CVB3	green	3	45 & 411 40 & 576	cells & vRNA spots cells & ST-CAAX spots (for normalization)	S1E
		nr. of nascent SunTag peptides per vRNA EMCV	blue	2	25 & 249 31 & 307	cells & vRNA spots cells & ST-CAAX spots (for normalization)	-
	S3L	phase0		2	8 & 77	cells & time-points	-
		phase1		2	8 & 159	cells & time-points	-
		phase2		2	8 & 240	cells & time-points	-
	S3M	infection MOI 0.25	black	4	709 & 602	cells & cells for pre-imaging correction	5A
		infection MOI 1.25	light blue	4	702 & 553	cells & cells for pre-imaging correction	-
3	3A	phase1 duration successful replication (phase2 ≤ 60 minutes)	green	21	146	cells	-
		phase1 duration unsuccessful replication (phase2 > 60 minutes)	red	21	27	cells	-
	3C	frequency pulses non-treated control		21	133	cells	6F
		frequency pulses 3Cpro inhibitor		2	40	cells	-
		frequency pulses 3Dpol inhibitor		2	36	cells	-
	3D	duration 1st pulse non-treated control	black	19	133	phase1 pulses	S4F
		duration 1st pulse 3Cpro inhibitor	blue	2	40	phase1 pulses	S4G
		duration 1st pulse 3Dpol inhibitor	purple	2	36	phase1 pulses	S4H
	3E	spot intensity wildtype IRES	black	3	184 & 1125	cells & spots	-
		spot intensity IRES mutant Sabin-like1	grey	3	48 & 279	cells & spots	-
		spot intensity IRES mutant Sabin-like2	grey	3	51 & 313	cells & spots	-
		spot intensity IRES mutant Sabin-like3	orange	3	53 & 325	cells & spots	-
	3F	duration phase1 wildtype IRES	black	21	183	cells	2H; 5B; 6E; S3A; S4B
		duration phase1 IRES mutant Sabin-like3	orange	3	58	cells	-
	3G	duration phase2 wildtype IRES	black	21	176	cells	2I; 5C; 6D; 6I; S3B

		duration phase2 IRES mutant Sa-bin-like3	orange	3	59	cells	-
S4	S4A	observed frequency >1 translating vRNA during phase1	orange	17	15 out of 141	cells	-
		observed frequency phase1 pulses	green	19	17 out of 133	cells	-
	S4B	duration phase1 control	black	21	183	cells	2H; 3F; 5B; 6E; S3A
		duration phase1 3Cpro inhibitor	blue	2	40	cells	-
		duration phase1 3Dpol inhibitor	purple	2	36	cells	-
	S4D	infection untreated control	black	2	160	cells	-
		infection 3Cpro inhibitor	blue	2	135	cells	-
		infection 3Dpol inhibitor	purple	2	144	cells	-
	S4E	noise 1 minute interval	grey	2	17 & 48	cells & gaps	-
		noise 2 minutes interval	blue	12	78 & 130	cells & gaps	-
		noise 5 minute interval	orange	5	27 & 28	cells & gaps	-
	S4F	non-treated control any pulse	grey	19	130 & 150	cells & pulses	6G
		non-treated control 1st pulse	orange	19	130	pulses	3D
		non-treated control 2nd pulse	green	19	17	pulses	-
	S4G	3Cpro inhibitor; any pulses	grey	2	40 & 76	cells & pulses	-
		3Cpro inhibitor; 1st pulse	orange	2	40	pulses	3D
		3Cpro inhibitor; 2nd pulse	green	2	19	pulses	-
		3Cpro inhibitor; 3rd pulse	blue	2	10	pulses	-
		3Cpro inhibitor; 4th pulse	purple	2	5	pulses	-
	S4H	3Dpol inhibitor; any pulses	grey	2	37 & 77	cells & pulses	-
		3Dpol inhibitor; 1st pulse	orange	2	37	pulses	3D
		3Dpol inhibitor; 2nd pulse	green	2	22	pulses	-
		3Dpol inhibitor; 3rd pulse	blue	2	9	pulses	-
	S3I	untreated control; any break	grey	19	130 & 149	cells & breaks	-
		untreated control; 1st break	orange	19	130	breaks	-
		untreated control; 2nd break	green	19	16	breaks	-
	S4J	3Cpro inhibitor; any break	grey	2	39 & 94	cells & breaks	-
		3Cpro inhibitor; 1st break	orange	2	39	breaks	-
		3Cpro inhibitor; 2nd break	green	2	23	breaks	-
		3Cpro inhibitor; 3rd break	blue	2	20	breaks	-
		3Cpro inhibitor; 4th break	purple	2	10	breaks	-
	S4K	3Dpol inhibitor; any break	grey	2	36 & 75	cells & breaks	-
		3Dpol inhibitor; 1st break	orange	2	36	breaks	-
		3Dpol inhibitor; 2nd break	green	2	22	breaks	-
		3Dpol inhibitor; 3rd break	blue	2	9	breaks	-
	S4L	duration phase3 wildtype IRES	black	12	112	cells	2J; 5D; S3C

		duration phase3 IRES mutant Sabin-like3	orange	3	46	cells	-
	S4M	duration phase4 wildtype IRES	black	21	100	cells	2K; 5E; S3D
		duration phase4 IRES mutant Sabin-like3	orange	3	46	cells	-
4	4B	methionine fluoresence; non-infected	blue	5	611	cells	4C; S5B
		methionine fluoresence; infected; 1 site	pink	5	169	cells	4C; S5B
		methionine fluoresence; infected; >1 site	black	5	213	cells	-
	4C	methionine fluoresence; non-infected	blue	5	611	cells	4B; S5B
		methionine fluoresence; 15 minutes after virus incubation	pink	3	23	cells	4B; S5B
		methionine fluoresence; 30 minutes after virus incubation	pink	3	23	cells	4B; S5B
		methionine fluoresence; 45 minutes after virus incubation	pink	3	27	cells	4B; S5B
		methionine fluoresence; 60 minutes after virus incubation	pink	4	25	cells	4B; S5B
		methionine fluoresence; 90 minutes after virus incubation	pink	3	22	cells	4B; S5B
		methionine fluoresence; 120 minutes after virus incubation	pink	3	25	cells	4B; S5B
		methionine fluoresence; 150 minutes after virus incubation	pink	1	3	cells	4B; S5B
		methionine fluoresence; 240 minutes after virus incubation	pink	4	21	cells	4B; S5B
	4F	eIF4G cleavage non-infected	black	4	5	cells	S5D
		eIF4G cleavage infected	red	4	26	cells	S5D
	4G	spot intensity in cells with 1 spot	green	4	35 &35	cells & spots	4H
		spot intensity in cells with > 1 spot	blue	4	159 & 1055	cells & spots	4H
	4H	spot intensity		4	194 & 1090	cells & spots	4G
	4I	phase1 intensity trace		3	43 & 43	cells & spots	-
	4J	harringtonine run-off in cells with 1 spot	green	3	11 & 11	cells & spots	-
		harringtonine run-off in cells with > 1 spot	blue	3	23 & 43	cells & spots	-
	4K	moving average virus translation	green	4	194 & 1090	cells & spots	4H
		moving average global translation	orange	5	993	cells	4B
S5	S5A	methionine fluorescence untreated cytoplasm		3	132	cells	-
		methionine fluorescence untreated nucleus		3	102	cells	-
		methionine fluorescence puro treated cytoplasm		3	174	cells	-
		methionine fluorescence puro treated nucleus		3	128	cells	-
	S5B	methionine fluorescence in untreated cells		5	215	cells	-

		methionine fluorescence in Puromycin cells		5	205	cells	-
		methionine fluorescence in uninfected cells after virus incubation		5	611	cells	4B, 4C
		methionine fluorescence in infected cells after virus incubation		5	382	cells	4B, 4C
	S5C	eIF4G cleavage reporter western blot		2			-
	S5D	eIF4G cleavage non-infected	black	4	5	cells	4F
		eIF4G cleavage successful viral replication	green	4	21	cells	4F
		eIF4G cleavage unsuccessful viral replication	red	4	5	cells	4F
5	5A	infection untreated	black	4	709 & 602	cells & cells for pre-imaging correction	S3M
		infection IFN	red	4	620 & 643	cells & cells for pre-imaging correction	-
	5B	duration phase1 untreated control	black	21	183	cells	2H; 3F; 6E; S3A; S4B
		duration phase1 IFN	red	7	94	cells	6E
	5C	duration phase2 untreated control	black	21	176	cells	2I; 3G; 6D; 6I; S3B
		duration phase2 IFN	red	7	94	cells	6E, 6I
	5D	duration phase3 untreated control	black	21	112	cells	2J; S3C; S4L
		duration phase3 IFN	red	7	36	cells	-
	5E	duration phase4 untreated control	black	21	100	cells	2K; S3D; S4M
		duration phase4 IFN	red	7	33	cells	-
	5F	survival non-infected cells untreated control	black	3	160	cells	2N
		survival cells with successful replication untreated control	dark green	3	120	cells	2N
		survival cells with unsuccessful replication untreated control	dark blue	3	29	cells	2N
		survival non-infected cells IFN	brown	3	88	cells	-
		survival cells with successful replication IFN	light green	3	99	cells	-
		survival cells with unsuccessful replication IFN	light blue	3	98	cells	-
	5G	spots in phase4 untreated control	black	21	100	cells	2L
		spots in phase4 IFN	red	7	32	cells	-
6	6A	RNA-sequencing IFN-induction		2			-
	6B	success-fraction infection	black or blue	each dot represents a repeat	≥ 50 per repeat per condition	cells	-
	6C	success-fraction infection	blue	each dot represents a repeat	≥ 50 per repeat per condition	cells	some controls also in S6B; S6J
	6D	duration phase 2 untreated control	black	21	176	cells	2I; 3G; 5C; 5I; S3B
		duration phase 2 IFN	red	7	94	cells	5C

		duration phase 2 siRNAseL + IFN	blue			cells	-
	6E	duration phase1 untreated control	black	21	183	cells	2H; 3F; 5B; S3A; S4B
		duration phase1 IFN	red	7	94	cells	5B
		duration phase 1 siRNAseL + IFN	blue			cells	-
	6F	frequency pulses untreated control		21	133	cells	3C
		frequency pulses IFN		7	94	cells	-
	6G	duration pulses untreated control	black	19	130 & 150	cells & pulses	S4F
		duration pulses IFN	red	7	94 & 151	cells & pulses	-
	6I	duration phase 2 untreated control	black	21	176	cells	2I; 3G; 5C; 6D; S3B
		duration phase 2 IFN	red	7	94	cells	5C
		duration phase 2 siIFIT1+IFN	green	3	75	cells	-
S6	S6B	success-fraction infection	blue	each dot represents a repeat	≥ 50 per repeat per condition	cells	some controls also in 6C; S6J
	S6C	qPCR siIFIT1		3			-
	S6D	qPCR OAS3		4			-
	S6E	qPCR C19Orf66		3			-
	S6F	qPCR OAS1		2			-
	S6G	qPCR HELZ2		3			-
	S6H	success-fraction infection	blue	3	≥ 50 per repeat per condition	cells	-
	S6I	qPCR PKR		4			-
	S6J	success-fraction infection	blue	3 and 2	≥ 50 per repeat per condition	cells	some controls also in 6C; S6B
	S6K	qPCR XRN1		2			-
	S6L	qPCR DIS3L		2			-
	S6M	qPCR RNAseL		3			-
	S6O	success-fraction infection	blue	3	≥ 50 per repeat per condition	cells	-

I

II

III

IV

V

VI

&
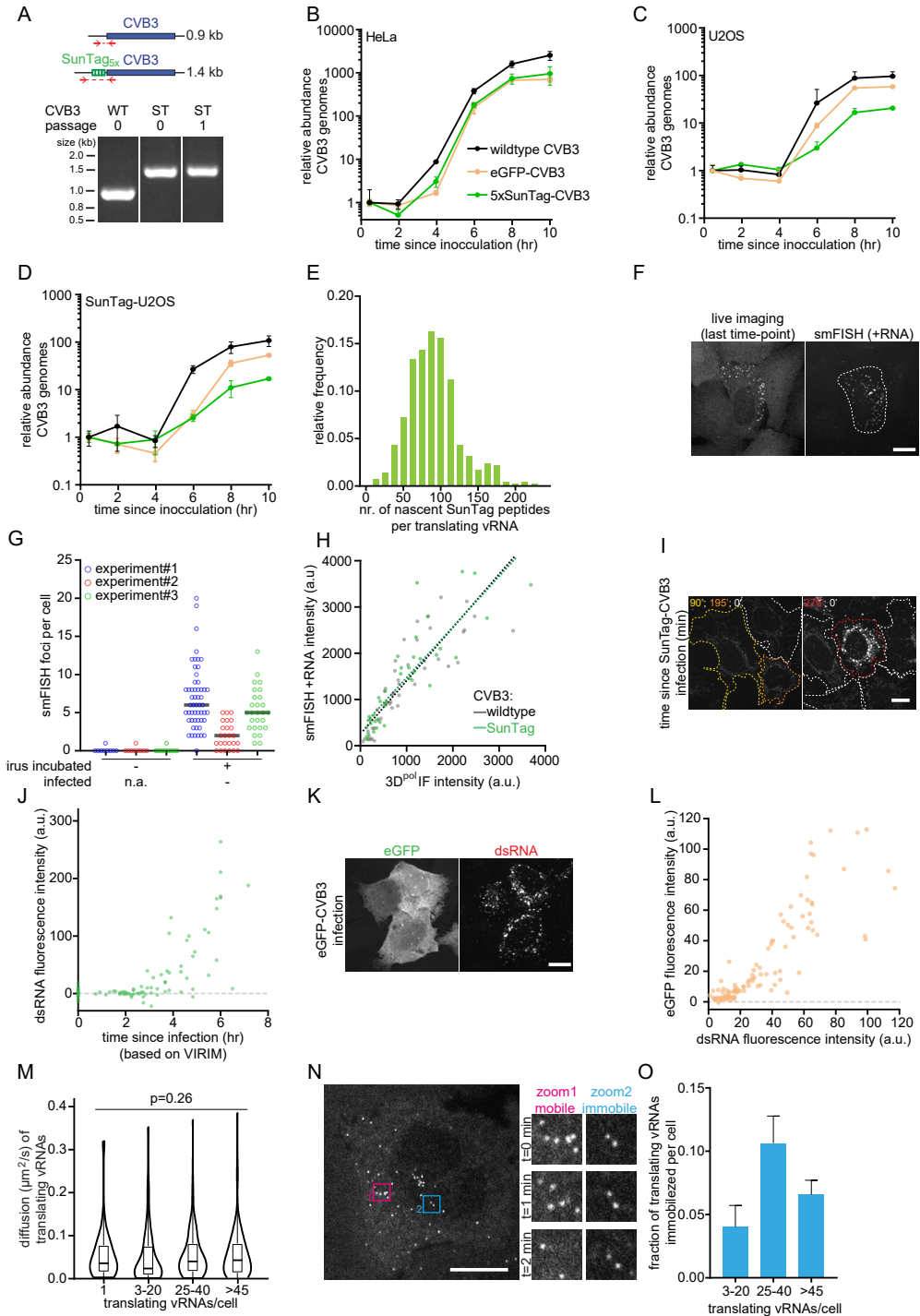
Supplemental Figure 1 – related to Figure 1. A single-molecule assay to visualize and analyze translation and replication of individual CVB3 vRNAs.

Figure S1 continued:

(A) Agarose gel analysis of the SunTag insert in the wild-type (WT) or SunTag (ST) CVB3 genome after RNA isolation, cDNA synthesis and PCR amplification of the indicated region. (B-D) Virus growth curves of the indicated viruses in indicated cell lines. Similar MOIs were used in all experiments. (E) Number of nascent SunTag peptides per translating vRNA focus based on the vRNA GFP fluorescence intensities compared to the GFP fluorescence intensity of 24xSunTag arrays expressed in STAb cells. (F) Representative images of STAb cell infected with SunTag-CVB3, during live-cell imaging (left) or after fixation and smFISH against +CVB3. (G) The number of smFISH +CVB3 'background' foci in cells not exposed to virus (left) or cells that were uninfected after incubation with SunTag-CVB3 (right). For each repeat, the mean number of smFISH foci in uninfected cells was used to correct the number of smFISH foci in infected cells in Fig. 1E, I, 2E. (H) Combined analysis viral protein synthesis (based on 3Dpol protein immunofluorescence) and viral load (based on fluorescence intensity of +CVB3 smFISH) in the STAb cells infected with indicated virus. Dashed lines indicate linear fits. (I, J) Representative images (I) and quantification (J) of combined analysis of live-cell imaging and dsRNA immunofluorescence of the same STAb cells infected with SunTag-CVB3. Color of outline (I) indicates the time between first detection of a translating vRNA and fixation. Cells in which no translating vRNAs were observed are indicated by a white outline and were used to correct for background fluorescence. (K, L) Representative images (K) and quantification (L) of combined analysis of GFP fluorescence and dsRNA immunofluorescence in the same U2OS cells infected with eGFP-CVB3. (M) Violin and boxplots of diffusion kinetics of translating vRNAs in cells that contain the indicated number of translating vRNAs. (N) Images of representative time-lapse movie of a STAb U2OS cell infected with SunTag-CVB3. Zooms indicate areas with mobile (pink) or immobilized (blue) translating vRNAs. (O) Bar graph of the fraction of immobilized translating vRNAs per cell. Every dot (G, H, J, L) indicates a single cell. Statistics is based on Kruskal-Wallis test. Error bars indicate SEM. Scale bars, 15 μm. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.
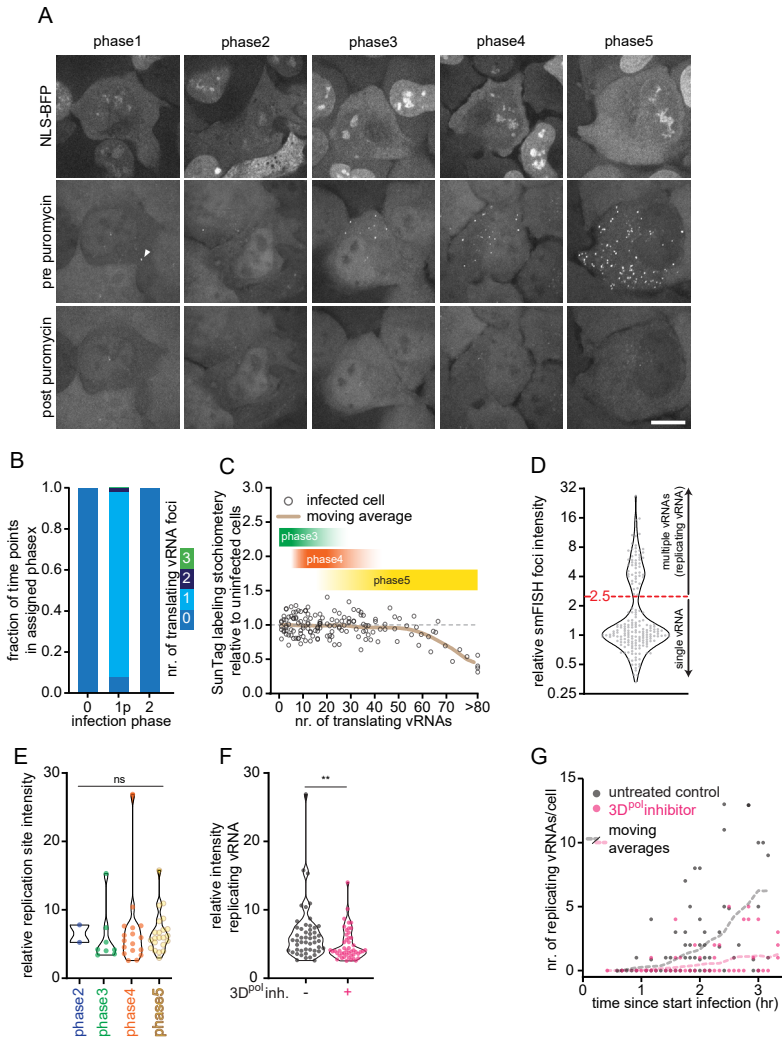
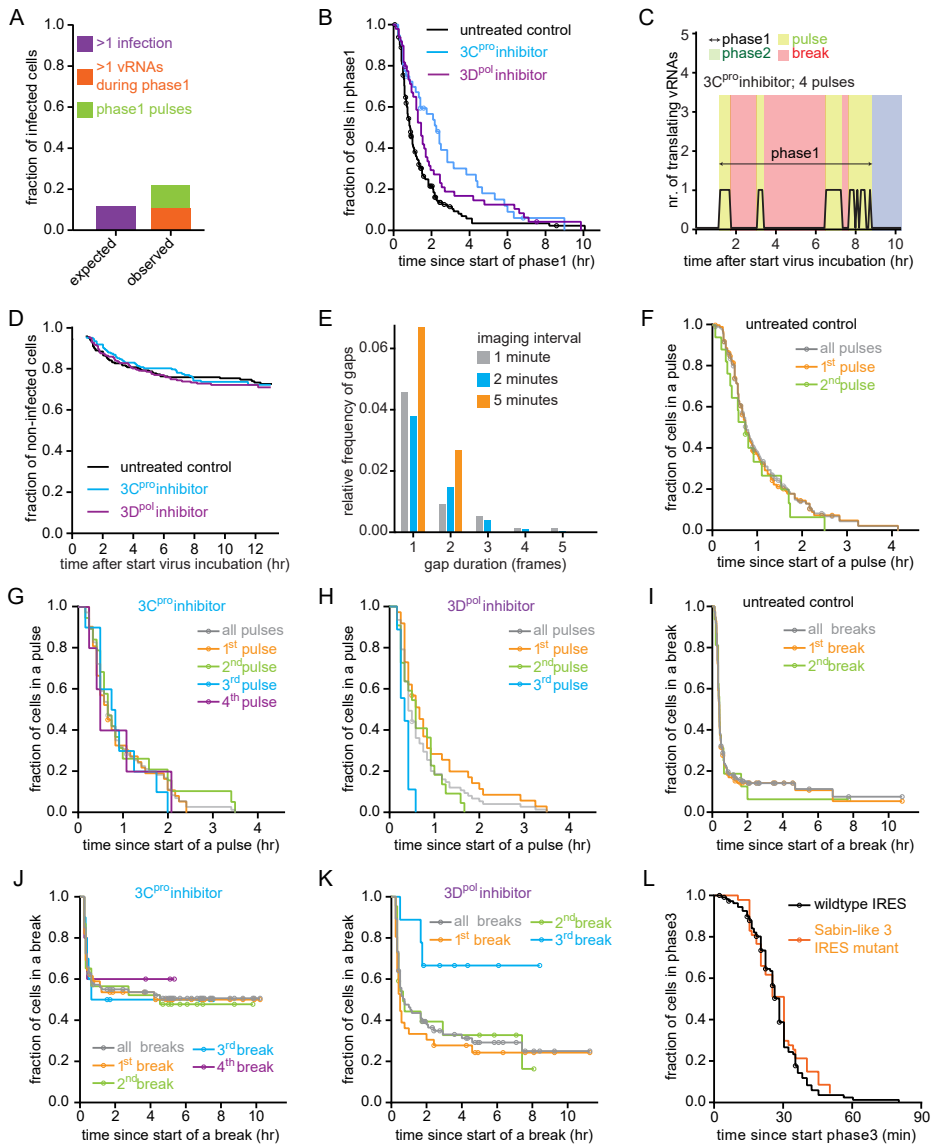**Supplemental Figure 2 - related to Figure 2. Live-cell imaging of translating vRNAs identifies infection phases.**

(A) Representative images of SunTag-CVB3 infected NLS-BFP STAb cells in different infection phases 3 hr after administration of virus before (middle panels) and 5 minutes after (bottom panels) puromycin administration. NLS-BFP (top panel) is shown as well to indicate that cells have been infection, based on loss of nuclear NLS-BFP fluorescence. Arrow head indicates first translating vRNA of during phase1. (B) Frequency of time-points during assigned phases during which indicated number of vRNAs were observed. Infection phase1p refers to phase1 pulses, indicating that only time-points assigned as a phase1 pulse were included in the analysis. (C) Combined analysis of the number of translating vRNAs and SunTag labeling stoichiometry in the same cells, normalized to the mean labeling stoichiometry in uninfected cells. SunTag labeling stoichiometry is based on GFP fluorescence intensity of 24xSunTag-peptide arrays expressed in the same cells. To illustrate corresponding phases, phase3, 4, and 5 are indicated based on the distribution of the number of translating vRNAs during phase4. (D) Fluorescence intensity of smFISH foci normalized to the mean of single vRNAs. Red dashed line indicates the intensity threshold of smFISH foci classified

Figure S2 continued:

as replicating vRNAs. (E) Replicating vRNA intensity normalized to the mean single vRNA smFISH. (F, G) Combined analysis of the same cells of live-cell imaging SunTag-CVB3 infection in SunTag U2OS cells to determine start of infection and smFISH to analyze replicating vRNAs. Replicating vRNA intensities are normalized to mean smFISH intensity of single vRNAs (F). 3Dpolinhibitor: GPC-N114 (10 µM). **, p < 0.01; ns, not significant, based on Kruskal-Wallis test (E) or unpaired Mann-Whitney test (F). Scale bars, 15 µm. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.
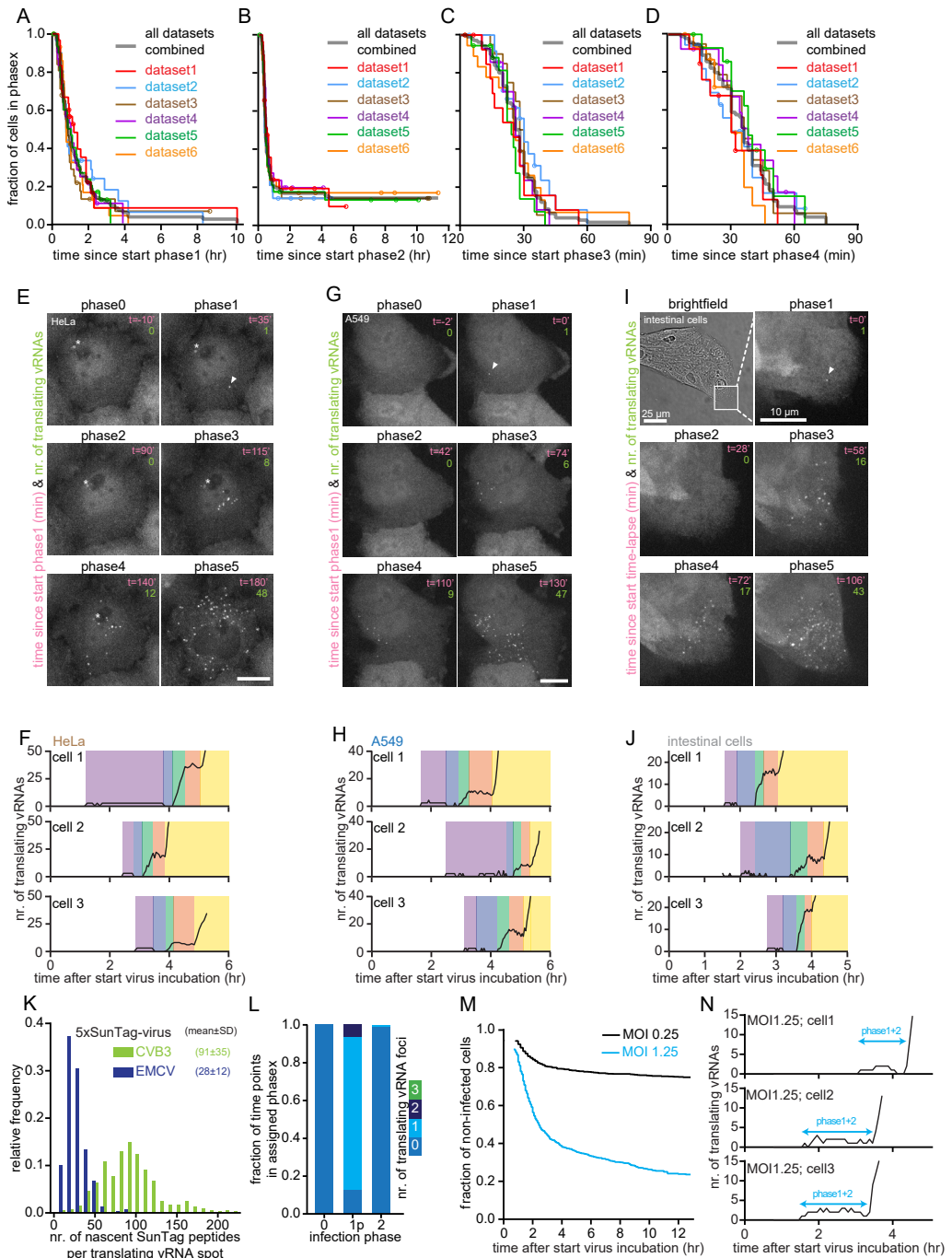
Supplemental Figure 3 – related to Figure 2. Reproducibility of single-cell dynamics and heterogeneity of viral replication.

(A-D) Kaplan Meier graphs showing durations of infection phases of all repeats combined (black) or separated into 6 datasets (colors). Circles indicate last analyzable time-point for individual cells. Data in black is replotted from Fig. 2H-K for comparison. (E-J) Representative images (E,

Figure S3 continued:

G, I) and example quantifications (F, H, J) of time-lapse movies of indicated STAb cells infected with SunTag-CVB3. (E, G, I) Arrow head indicates the first translating vRNA. Asterisk indicates a background spot. (F, H, J) Colors illustrate infection phases. Note that data points during phase0, 1, and 2 are enhanced 3-fold to aid visual inspection of data. (K) Number of nascent SunTag peptides per translating vRNA focus based on the vRNA GFP fluorescence intensities compared to the GFP fluorescence intensity of 24xSunTag arrays expressed in STAb cells. Data plotted in green is replotted from Fig. S1E for comparison. (L) Frequency of time-points during assigned phases during which indicated number of vRNAs were observed. Infection phase1p refers to phase1 pulses, indicating that only time-points assigned as a phase1 pulse were included in the analysis (See STAR methods). (M) Kaplan Meier graphs of the fraction of uninfected cells remaining after incubation with SunTag-CVB3. Data is corrected for the fraction of cells that were infected before the start of the movies, as indicated by the gap at the start of each graph. (N) Representative example quantifications of STAb cells infected by multiple viruses after incubation with SunTag-CVB3 at MOI 1.25, based on observing multiple vRNAs during phase1. As phase1 and2 of individual infections could not be distinguished, the duration between first detection of a translating vRNA and start of phase3 was determined as proxy for phase1+2 duration. Scale bars, 15 µm (E, G) or 10 µm and 25 µm (I). The number of experimental repeats and cells analyzed per experiment are listed in Table S1.
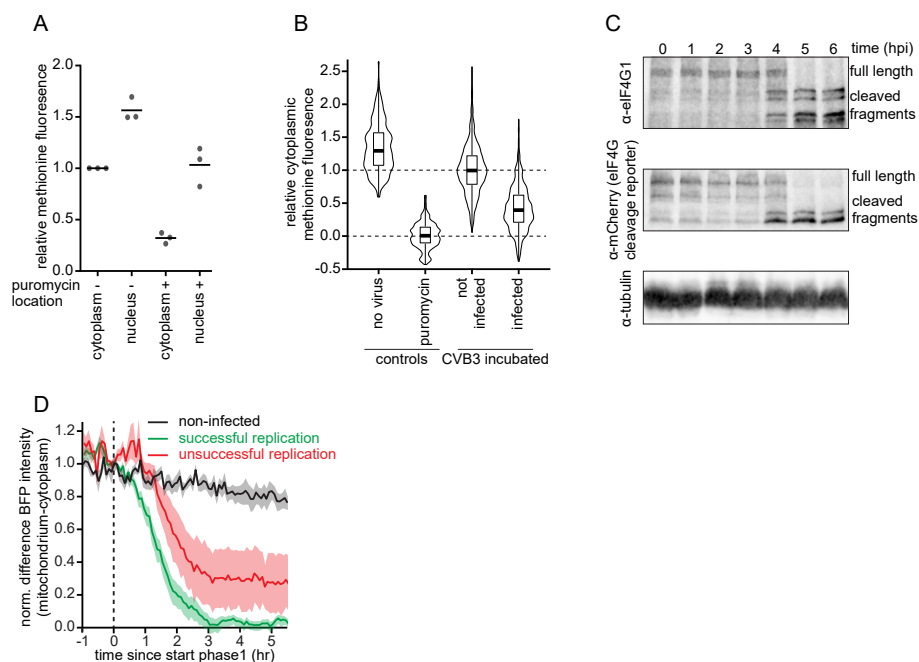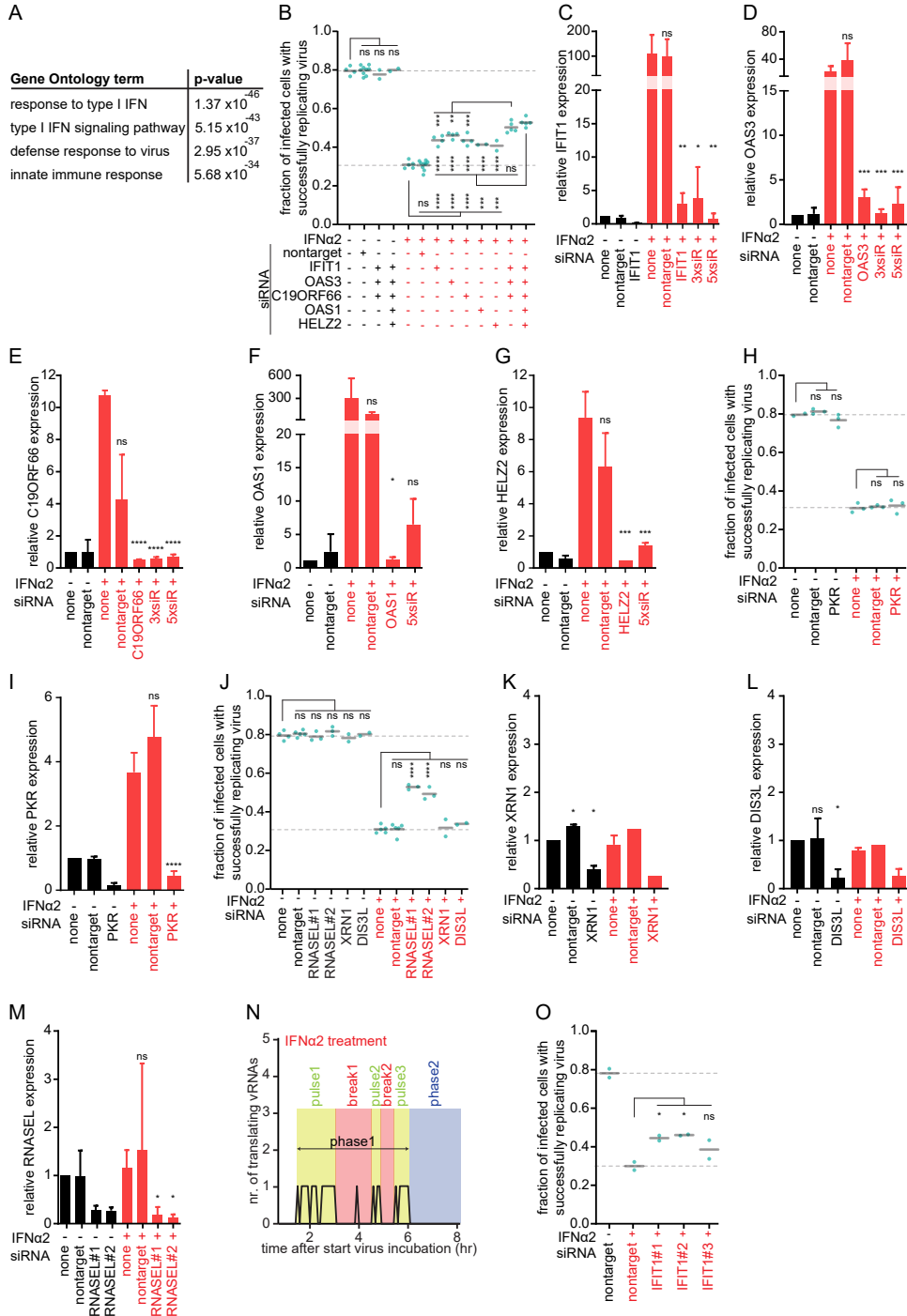
Supplemental Figure 4 - related to Figure 3. Timing of translation-to-replication transition.

(A) Frequency of multiple infections or phase1 pulses in infected cells. (B) Kaplan Meier graphs showing durations of infection phase1 (sum of all pulses and breaks; indicated with arrow in Fig. 3B, S3C). (C) Quantification of a representative example cell of the number of translating vRNAs over time after 10 µM Rupintrivir treatment. (D) Kaplan Meier graphs of the fraction of uninfected cells remaining after incubation with SunTag-CVB3. Data is corrected for the fraction of cells that were infected before the start of the movies, as indicated by the gap at the start of each graph. (E) Frequency at which indicated number of consecutive time-

Figure S4 continued:

points without a translating vRNA (gaps) occurred in the analysis of time-lapse movies of STAb cells infected with SunTag-CVB3. (F-K) Kaplan Meier graphs showing the duration of individual phase1 pulses (F-H) or individual phase1 breaks (I-K). The data plotted in grey (F-K) represents the combination of all data. Datasets with less than six cells were not plotted, but are included in the combined datasets (grey). Data plotted in orange (F-H) is replotted from Fig. 3D, for comparison. (L, M) Kaplan Meier graphs showing the duration of phase3 (J) and phase4 (K). Data plotted in black is replotted from Fig. 2J, K, for comparison. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.

Supplemental Figure 5 – related to Figure 4. Inhibition of host cell translation.

(A) Fluorescence intensity of the labeled methionine analog either in the cytoplasm or in the nucleus after indicated treated measured in single cells. Data are normalized to the mean cytoplasmic intensity. Every dot indicates an experimental repeat; lines indicate mean values. (B) Violin and boxplots of the cytoplasmic fluorescence intensity of the methionine analog in single cells, after indicated treatments. All datasets were normalized to the mean of uninfected (set to 1) and puromycin treated (set to 0) cells. For comparison, the data on uninfected cells in A is replotted from Fig. 3B, C (in blue). (C) Western blots of STAb cells expressing the eIF4G cleavage reporter after indicated incubation times with CVB3. (D) STAb cells expressing the eIF4G cleavage reporter were imaged after infection with SunTag-CVB3. BFP fluorescence intensity differences between mitochondria and cytoplasm over time for infected cells with either successful or unsuccessful replication of the incoming vRNA. The data in black is replotted from Fig. 4F. Shaded areas indicate SEM. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.

Supplemental Figure 6 – related to Figure 5. IFN-induced antiviral state represses replication in phase2.

(A) Selection of the top gene ontology terms and the corresponding p-values for genes with increased

Figure S6 continued:

expression upon IFNα2 treatment. (B) Fraction of infected cells with successfully replicating virus after transfection with the indicated siRNAs, and/or treatment with IFNα2. (C-G) qPCR analysis of expression levels of indicated genes, relative to GADPH expression levels, normalized to untreated controls. (H) Fraction of infected cells with successfully replicating virus after transfection with the indicated siRNA, and/or treatment with IFNα2. (I) qPCR analysis of expression levels of PKR, relative to GADPH expression levels, normalized to untreated controls. (J) Fraction of infected cells with successfully replicating virus after transfection with the indicated siRNA, and/or treatment with IFNα2. (K-M) qPCR analysis of expression levels of indicated genes, relative to GADPH expression levels, normalized to untreated controls. (N) An example single-cell trace of pulses and breaks in an IFNα2-treated cell. (O) Fraction of infected cells with successfully replicating virus after transfection with the indicated siRNAs, and/or treatment with IFNα2. Every dot (B, H, J, O) represents an independent experiment, black lines indicate the mean, and dashed lines indicate the mean of the untreated controls. Error bars indicate SD. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.01$; $p < 0.001$; ****, $p < 0.0001$; ns, not significant, based on Dunnett's multiple comparisons tests (B, H, J, O) or two-tailed unpaired Student's t-test compared to no siRNA conditions (C-G, I, K-M). The number of experimental repeats and cells analyzed per experiment are listed in Table S1.
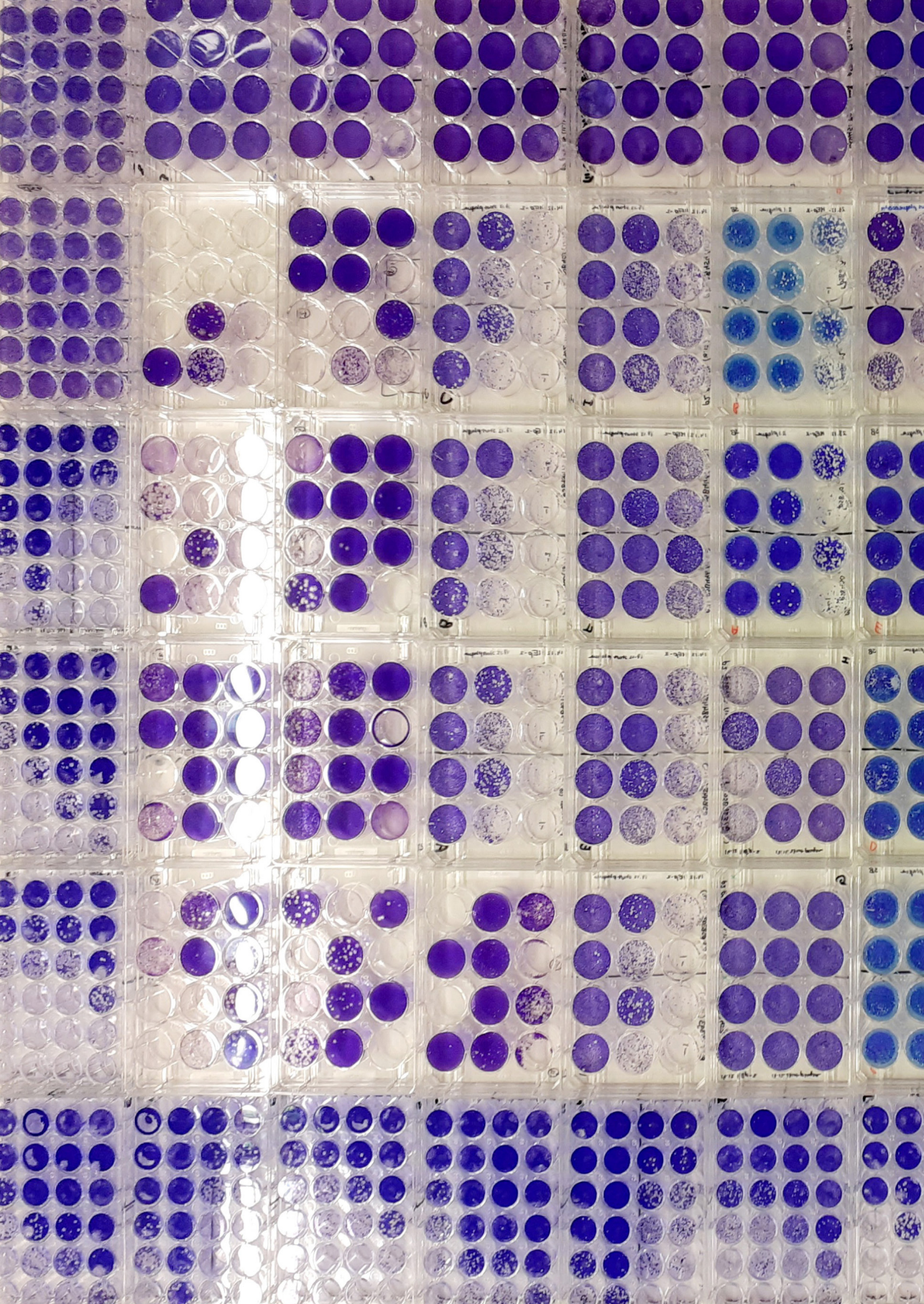
Chapter 5

Viral Gene Expression Dynamics and Antiviral Response During Early RSV Infection

Sanne Boersma[1], Dhanushika Ratnayake[1], Stijn Sonneveld[1], Iris Bally[1], Rupa Banjeree[1], Huib R. Rabouw[1], Lucas J. M. Bruurs[1], Louis J. Bont[2], Jean-François Eléouët[3], Marie-Anne Rameix-Welti[4,5], and Marvin E. Tanenbaum[1,]

[1] Oncode Institute, Hubrecht Institute – KNAW and University Medical Center Utrecht, Utrecht, The Netherlands

[2] Wilhelmina Children's Hospital and UMC Utrecht, Utrecht, The Netherlands

[3] Université Paris-Saclay, INRAE, UVSQ, VIM, Jouy-en-Josas, France

[4] Université Paris-Saclay, INSERM, Université de Versailles St. Quentin, UMR 1173 (2I), Versailles, France

[5] AP-HP Université Paris Saclay, Hôpital Ambroise Paré, Laboratoire de Microbiologie, Boulogne-Billancourt, France

Background illustration: Collection of plaques assay plates used to determine viral titers
Related to: Chapter5-Fig. S5A

SUMMARY

Respiratory syncytial virus (RSV) is a leading cause of respiratory infections, particularly among young children and elderly. The RSV genome is a single-stranded negative-sensed RNA molecule comprising of 10 genes that are transcribed upon entry into a host cell by the viral polymerase. Expression of RSV genes is proposed to follow a gradient with upstream genes transcribed at higher rates compared to downstream genes. However, currently available assays lack sensitivity to study expression dynamics early in infection; the time window during which the infection is initiated, and the antiviral defenses may start to counteract the infection. Here, we applied single-molecule imaging to study early viral gene expression dynamics and coordination of expression between genes. Our data challenges the basic gradient model of RSV gene expression, but instead suggests that additional mechanisms, like bidirectional scanning, contribute to well-balanced RSV gene expression. Furthermore, there is an opposing trend between potent early viral gene expression and launching of an antiviral response in single cells. Collectively, we established a collection of single-molecule imaging tools to decipher the dynamics of early viral gene expression which may help to understand how and when the antiviral response can inhibit an RSV infection.

INTRODUCTION

Respiratory syncytial virus (RSV) is an enveloped RNA virus from the *Mononegavirales* order, causing respiratory infections including pneumonia and bronchiolitis. RSV is one of the leading causes of lower respiratory tract infections and can be pathogenic and even lethal particularly in infants, elderly and immunocompromised patients (Htar et al., 2020; Hu et al., 2020; Shi et al., 2017). To date, only limited treatment options are available for RSV patients.

An RSV infection typically starts with entry and release of the viral genome of one virion into a single cell. Next, the virus synthesizes new virions by genome expression to produce new viral proteins and by replication of the genome. At the same time, host cell viral sensors can detect the presence of a virus, for example by sensing the viral replication intermediate double-stranded RNA (dsRNA) or the presence of viral RNA or surface proteins in endosomes (Sun and López, 2017). Sensing of a virus induces rapid activation of antiviral pathways, including synthesis of interferons (IFNs), which results in the expression of IFN induced genes (ISGs). By activating an antiviral state through ISG expression, host cells can inhibit the infection by blocking various virus or host cell processes (Schneider et al., 2014). To evade the antiviral response, viruses have evolved a plethora of mechanisms that interfere with the host cell's ability to inhibit an infection. Consequently, the outcome of an infection is probably dependent on a complex competition between the virus and the antiviral response. RSV is particularly well able to escape an antiviral response, as IFN levels are typically low in nasal swabs from RSV patients (Hijano et al., 2019; Sedeyn et al., 2019; Spann et al., 2004). In contrast, RSV strains that have impaired expression of several RSV genes skews the virus-host competition in favor of the host (Spann et al., 2004), indicating that adequate viral gene expression is essential for successful evasion of an antiviral response and to establish a successful infection.

Like other *Mononegavirales*, such as mumps virus, measles virus, or Ebola virus,

I

II

III

IV

V

VI

&

the RSV genome is a single-stranded negative-sensed RNA molecule (-ssRNA) and the RSV genome contains 10 genes encoding 11 proteins. Both replication and transcription are performed by the large RNA dependent RNA polymerase (L), which is an essential component of the virion. Loading of L and initiation of transcription takes place at the 3' end of the genome, from where the L scans the genome towards the first gene (Noton and Fearns, 2015). Each viral gene is surrounded by a Gene Start (GS) and Gene End (GE) sequence, analogous to eukaryotic transcription promoters and termination sequences. Upon recognition of a GS, L transcribes the gene and produces a fully functional mRNA with a cap and a polyadenylated tail, which can be translated into viral proteins by host cell's ribosomes (Cao et al., 2021). At a GE, L releases the mRNA molecule and can then either dissociate from the genome or continue scanning till the next GS. Consequently and similar to other *Mononegavirales*, L transcribes the RSV genes sequentially from 3' to 5'; the transcription-rate of each RSV gene is thought to decrease per gene along the genome resulting in a gradual decrease in viral gene expression from the 3' to 5'end of the genome (Barr et al., 2008; Brauburger et al., 2016; Kolakofsky et al., 2004; Noton and Fearns, 2015).

In agreement with this gradient-model of viral gene expression, mRNAs encoded by upstream RSV genes are typically more abundant compared to mRNAs generated from downstream genes (Aljabr et al., 2016; Barik, 1992; Piedra et al., 2020). By employing gradient-based gene expression, RSV may establish appropriate expression levels of each gene. For example, non-structural protein 1 and 2 (NS1 + NS2) are encoded by the two most upstream RSV genes and are responsible for preventing an antiviral response (Sedeyn et al., 2019; Sun and López, 2017; Valarcher et al., 2003). In contrast, L is encoded by the most downstream RSV gene. Although L is essential in the viral life cycle, low expression of L compared to NS1 and NS2 is sufficient for a successful infection. Adequate expression of RSV genes may thus be required for successful production of new virions and for potent evasion of an antiviral response.

Several indicators suggest that RSV gene expression may not fully follow the gradient model. For example, the GS of the gene encoding L is located upstream of the GE of the preceding gene. Based on Vesicular stomatitis virus (VSV), another *Mononegavirales*, transcription initiation on a GS upstream of a GE may depend on backwards scanning of L to an upstream GS (Barr et al., 2008). Moreover, analyses of relative transcript abundance by qPCR or RNA-sequencing suggest that mRNAs from the 7th gene (encoding glycoprotein (G)) are the most abundant viral transcripts ≥24 hr after infection (Aljabr et al., 2016; Piedra et al., 2020). Therefore, transcription coordination between individual RSV genes may entail unknown factors.

Currently employed approaches lack sensitivity to investigate RSV gene expression dynamics or the coordination between viral gene expression and the antiviral response particularly at the start of an infection when viral gene expression and the virus-host competition are initiated. For instance, analysis of virus infection often requires a large collection of cells to generate a robust read-out, which is incompatible with detailed analysis of cell-to-cell variation in the onset and progression of an infection and the variation in the success-rate of launching an antiviral response in an individual cell. Moreover, many assays require fixation or lysis of the infected cells, complicating examination of temporal dynamics. Even assays that enable live analysis of virus infections in individual cells, e.g., using a fluorophore-expressing

virus, fail to zoom-in on the first events of an infection, as it may take several hours after infection until sufficient fluorescence proteins have been generated and matured to be detected above background (Balleza et al., 2018). Likewise, to study the relation between viral gene expression and the antiviral response, both aspects should be measured in the same sample or specifically in the same individual cell, rather than generating two distinct read-outs from parallel samples.

To investigate RSV gene expression dynamics and the virus-host competition, new tools are required that overcome these limitations. Recently, we established a live-cell single-molecule imaging assay (called Virus Infection Real-time Imaging or VIRIM) to study translation and replication dynamics of single-stranded positive-sensed RNA (+ssRNA) viruses (Boersma et al., 2020). Using VIRIM, we uncovered extensive heterogeneity in the early stages of picornavirus infection and identified replication of the first genome as a major bottleneck for successful infection. VIRIM bypasses most of the constraints that limited analysis of early +ssRNA virus infection and illustrates how high-resolution assays can aid in deciphering the complex virus-host competition. Here, we generated a collection of single-molecule imaging tools to examine early -ssRNA virus infection in individual cells, specifically the dynamics of early RSV gene expression and the coordination between RSV gene expression and launching of an antiviral response.

RESULTS
smFISH tools to explore early RSV gene expression

To examine RSV gene expression dynamics with single-molecule resolution we generated a single-molecule fluorescence in situ hybridization (smFISH) toolkit, based on Atto565- and Atto633-labeled smFISH probesets targeting the mRNAs of each RSV gene (see Methods; Fig. 1A). In contrast to other assays, such as expression of a fluorescent protein or RNA sequencing, smFISH is highly sensitive and can be used to detect single mRNA molecules directly upon its synthesis (Femino et al., 1998; Lyubimova et al., 2013). Moreover, smFISH has been used successfully to study the absolute number of viral RNAs and the spatial distribution of viral RNAs of various RNA viruses (Boersma et al., 2020; Garcia-Moreno et al., 2019; Lee et al., 2021; Ramanan et al., 2016; Singer et al., 2021). To validate the RSV smFISH probesets, we infected airway epithelium cells (A549-adenocarcinomic human alveolar basal epithelial cells), reflecting the natural tissue infected by RVS, with human RSV strain A2 and performed smFISH 4 hours post inoculation (hpi) (Fig. 1B). For all probesets, bright smFISH foci could be readily detected. These foci were only observed in a subset of cells, in accordance with the low MOI (0.25) that was used to infect the cells. Hardly any foci were detected in mock-infected cells (0.07-0.63 foci/cell) (Fig. 1C), confirming the specificity of the smFISH probesets. To determine the efficiency of the smFISH probesets, smFISH was performed using cells stably expressing fusion mRNAs containing the coding sequence of both the nucleocapsid (N) gene and the phosphoprotein (P) gene (Fig. S1A). More than 97% of the N-P fusion transcripts were dual labeled with N and P probesets, whereas dual labeling was abolished (<2%) if the image of the P smFISH was flipped, demonstrating that the detection efficiency of the smFISH probes is high (Fig. 1D, S1A).

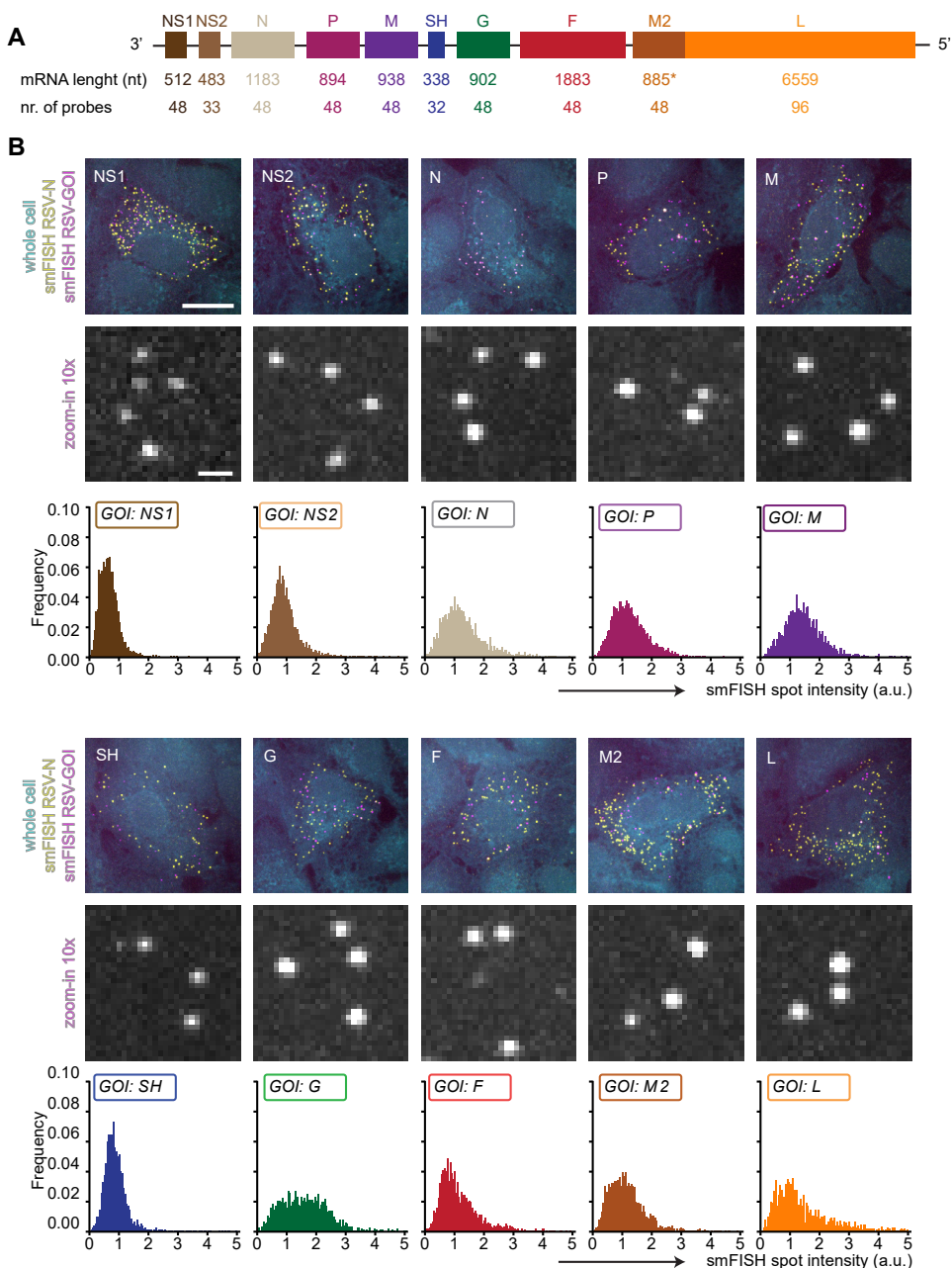As we aimed to use transcript abundance based on smFISH as a read-out of RSV

gene expression, information on the RSV mRNA stability is required; if an mRNA is degraded rapidly, the transcript count may not accurately represent gene expression. The stability of all RSV transcripts was calculated based on mRNA abundance after treatment with an L inhibitor (ALS-8112) (Fig. S1B). The stability of mRNAs from all RSV genes was similar (6.1 ± 0.7 hr half-lives; Fig. 1E) and as such, their stability does not affect viral gene expression analysis based on the relative mRNA abundance using smFISH. Collectively, we conclude that the smFISH probesets are sensitive, specific, and efficient and therefore ideally suited to study early RSV gene expression.

smFISH analysis of early RSV gene expression challenges basic gradient model

To analyze expression of RSV genes early in infection, we infected A549 cells and performed smFISH 2, 4, or 6hpi using probesets targeting two RSV targets: N transcripts (labeled with NAtto565) and each of the RSV transcripts (labeled with gene of interest (GOI)Atto633) (Fig. 1B, 2A). Although there was widespread heterogeneity in the number of N transcripts between single cells, the N transcript abundance correlated well with each of the GOI transcripts (Fig. 2A, S3A). Moreover, the correlations of N vs. GOIs were similar for each time-point (Fig. 2A, B). The relative abundance of mRNAs generally followed a declining trend from the most upstream (NS1) to the most downstream (L) gene of the RSV genome, as predicted by the gradient model (Fig. 2A). Note that in the NAtto565 vs. N Atto633 smFISH experiments (called N vs. N) the two N probesets competed for target binding, leading to reduction of the smFISH detection efficiencies (Fig. 1B, S2) and, the N vs. N experiments were therefore excluded from further analysis.

Based on the mean ratio between each GOI and N and assuming the gradient model, we calculated the transcription reinitiation chance (μ) at each gene transition (Fig. 2C; see Methods). Interestingly, the μ is variable at different gene transitions (Fig. 2C), even though the GSs and GEs of all genes are very similar and no additional sequence motives controlling reinitiation are known (Collins et al., 1986; Cowton et al., 2006; Kuo et al., 1996; Moudy et al., 2003). For example, the transcription reinitiation chance between the NS1 and NS2 genes is 100%, whereas transcription reinitiation between G and F genes takes place only ~69%. The intergenic length (i.e. the distance between GS_{n+1} and GE_n) has been proposed to affect the reinitiation efficiency, based on experiments with the *Mononegavirales* VSV (Barr et al., 2008). However, there is no apparent relationship between μ and the intergenic length for RSV, suggesting that transcription reinitiation occurs independently of the intergenic length and that unknown aspects are likely to contribute to gene-to-gene variations in RSV transcription reinitiation efficiencies (Fig. 2D).

Even though the average transcript ratios between N and each of the GOIs followed a declining trend from the first to the last gene of the genome, several aspects from the GOI vs. N smFISH experiments seem to contradict the gradient model. Firstly, G mRNAs are relatively more abundant compared to SH mRNAs (~1.5-fold), although the G gene is located downstream of the small hydrophobic protein (SH) gene (Fig. 2B). Previous experiments using RNA-sequencing and qPCR that focused primarily on relatively late time-point in infection (≥24 hpi) also suggested that G mRNAs are highly abundant and that G expression may deviate from the gradient model (Aljabr et al., 2016; Piedra et al., 2020). Expression of the G gene thus diverges from the gradient-based gene expression model during early and late infection. Secondly,
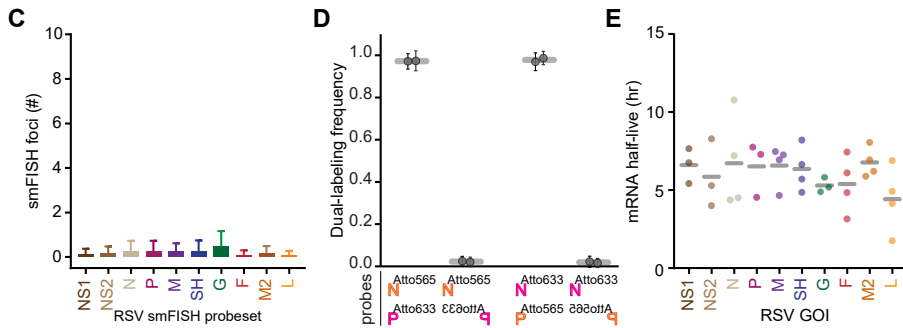
Figure 1. smFISH to analyze early RSV gene expression.

(A) Schematic illustrating the position and relative size of viral genes along the genome, including details on the length of the mRNAs targeted by the smFISH probesets and the number of probes per probeset. M2 and L partially overlap, and the overlapping sequences were excluded during probe design, as indicated with an asterisk. (B) Representative smFISH images of A549 cells infected with RSV using N-Atto565 (yellow) and indicated gene of interest (GOI) Atto633 probesets, including a zoom-in of the GOI smFISH in the second row. The bottom row shows histogram of GOI smFISH spot intensity distributions. Scale bars in top and middle row, 15 μm and 1 μm respectively.

Figure 1 continued:

(C) Quantification of false-positive smFISH spot calling in mock infected cells. Error bars indicate SD. (D) Fraction of smFISH foci dual labeled by indicated probesets in cells expressing N-P fusion mRNAs. Control analyses with flipped P smFISH images is indicated with mirrored probeset annotations. (E) Table indicating mean ± SD half-lives of RSV mRNAs. See also Figure S1. The number of experimental repeats, cells, and foci analyzed per experiment are listed in Table S1.

there is widespread variation in the relative abundance of mRNAs; although the average transcript ratios between N and each GOI declined along the genome, the standard deviations of each average GOI vs. N combination were substantial (Fig. 2A, B, S3). Thirdly, detailed inspection of individual cells revealed that for each GOI vs. N combination cells are present that seem to contradict the gradient model. For example, NS1 and NS2 genes are located upstream of N and conversely NS1 or NS2 mRNAs should always be more abundant than N mRNAs, according to the gradient model. However, N mRNAs exceeded NS1 or NS2 mRNAs in many cells. Similarly, mRNAs of genes downstream of N should have a lower concentration compared to N, whereas cells in which the GOI:N ratio >1 were present for each GOI (Fig. S3). Cells in apparent contradiction to the gradient model are particularly prevalent among cells with few N mRNAs (Fig. S3).

To investigate quantitatively how well the observed data aligns with the gradient model, we performed computational simulations (see Methods). First, we generated a 'basic model', based on the gradient model and the transcription reinitiation rates at each gene transition as determined using the relative GOI expressions (Fig. 2C). Comparison between simulated data and the experimentally obtained data revealed that the simulated data was less variable (i.e. the SD of the GOI vs. N ratio of the simulated data was smaller compared to the experimentally-obtained data) (Fig. 3A and S3 vs. S4A). Inspection of simulated cells with a low number of N mRNAs showed that these cells are the main source of variation in the GOI vs. N ratio between single cells, similar to the experimentally obtained data (Fig. S3 vs S4A). However, no simulated cell passed the GOI:N=1 threshold; the simulations did not result in cells with less NS1 or NS2 mRNAs compared to N or cells with more mRNAs from downstream GOIs compared to N. As described above, these cells are abundantly present in the experimentally obtained data. We therefore concluded that the basic model does not satisfactorily recapitulate the experimentally obtained data.

We reasoned that additional factors needed to be included into the basic model to improve the comparison between experimentally obtained data and simulated data.
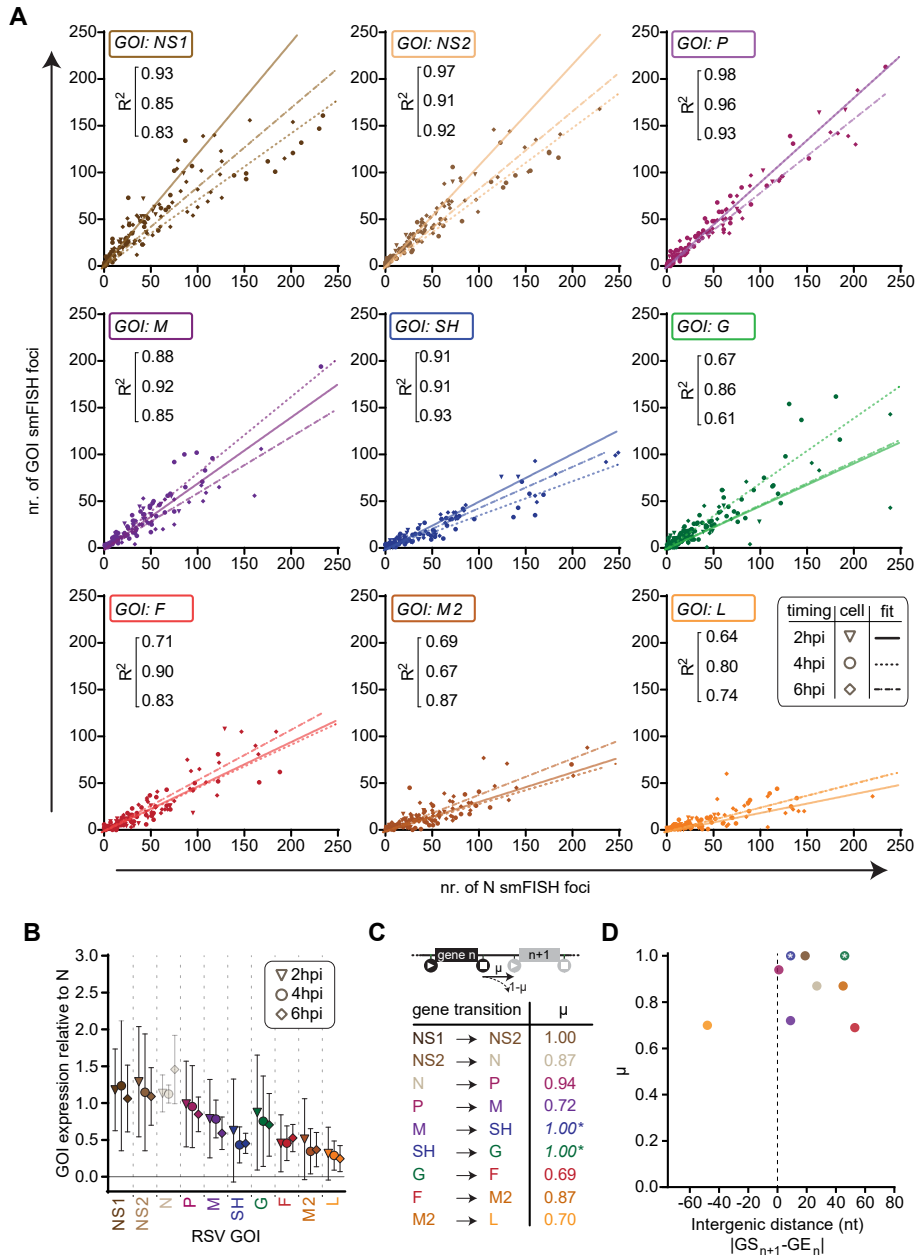
A

B

C

D

Figure 2. Gradient of gene expression along RSV genome.

(A) Scatterplots of correlations between number of N and GOI smFISH foci in single cells. Data is subdivided into nine panels, based on the GOI, as indicate in the top left corners. The x- and y-axis labeling are equal for each panel and are indicated with arrows. The L panel includes an additional legend to clarify symbols: each triangle, circle, or diamond indicates a single cell and solid, dashed, or interrupted lines indicate linear correlations from 2, 4, or 6 hpi experiments respectively. The Pearson R2 of the correlations (2, 4, 6 hpi top-to-bottom) is indicated underneath the GOI in the top left corner. Note that same data is replotted in Figure S3B. (B) Mean ratio between GOI and N

Figure 2 continued:

mRNAs from 2, 4, or 6 hpi experiments, indicated by triangles, circles, or diamonds respectively. Error bars indicate SD. (C) Schematic (top) of transcription reinitiation (μ) at a gene-gene transition and table (bottom) of μs at each transition based on mean ratios in B. Note that G mRNAs are relatively more abundant compared to SH mRNAs, which is inconsistent with gradient model. As correction, μs are artificially set to 1, as indicated with an asterisk. (D) Scatterplot of relation between intergenic distance and μ. Note that consistent color labeling is applied for GOIs in all figure panels. See also Figure S2, 3. The number of experimental repeats, and cells analyzed per experiment are listed in Table S1.

Such an additional factor compared with the basic model should result in an increase in single cell variation and the presence of cells passing the GOI:N=1 threshold among simulated cells. We envisioned three mechanisms that could introduce heterogeneity and whose combination with the basic model could thereby improve the model: backtracking, skipping, or loss of detection sensitivity. According to the backtracking module (Fig. 3B), L could scan to identify the next GS bidirectionally, as was suggested previously (Barr et al., 2008; Piedra et al., 2020); instead of only scanning the genome for the next downstream GS, L could also scan backwards to the GS of the same gene and transcribe the same gene again. In the skipping module (Fig. 3C), L may fail to recognize a GS; rather than reinitiating transcription on the next GS, L may skip a gene and instead continue scanning until the second next GS. Finally, the loss of detection sensitivity module is based on technical noise in quantifying the number of smFISH foci per cell; although smFISH-based quantification of mRNAs is highly accurate (Fig. 1), we cannot exclude some technical noise.

To test the basic model combined with the new modules, we ran series of simulated cells with varying probabilities of backtracking, skipping, or loss of detection sensitivity and compared the standard deviation between single cells from the simulations and smFISH data. In this comparison, the simulations and smFISH data match well, if the standard deviations are similar. The simulations and smFISH data only matched at very high probabilities for the skipping module or loss of detection sensitivity module (Fig. 3C). In contrast, the simulations with the backtracking module matched the experimental data well if the backtracking probability was approximately 0.5. Additionally, cells passing the GOI:N=1 threshold were abundantly present among the simulated cells using backtracking probability 0.5 (Fig. S4C). In conclusion, the basic gradient model supplemented with the backtracking module at probability 0.5 explains the experimentally obtained data well.

The absolute number of mRNAs of a GOI synthesized in an infected cell depends on the location of the gene along the RSV genome and the number of Ls transcribing the genome. In addition to the relative abundance of GOIs, we also used the absolute number of N or GOI mRNAs as input for the simulations and determined the L recruitment rate most in agreement with the experimentally obtained data (see Methods). According to the basic model, the L recruitment rate was $\sim 0.33\,s^{-1}$, whereas the rate was $0.2\,s^{-1}$ according to the basic model supplemented with the backtracking 0.5 module, suggesting that approximately every 5 seconds an L molecule initiates transcription of a genome (Fig. 3D).
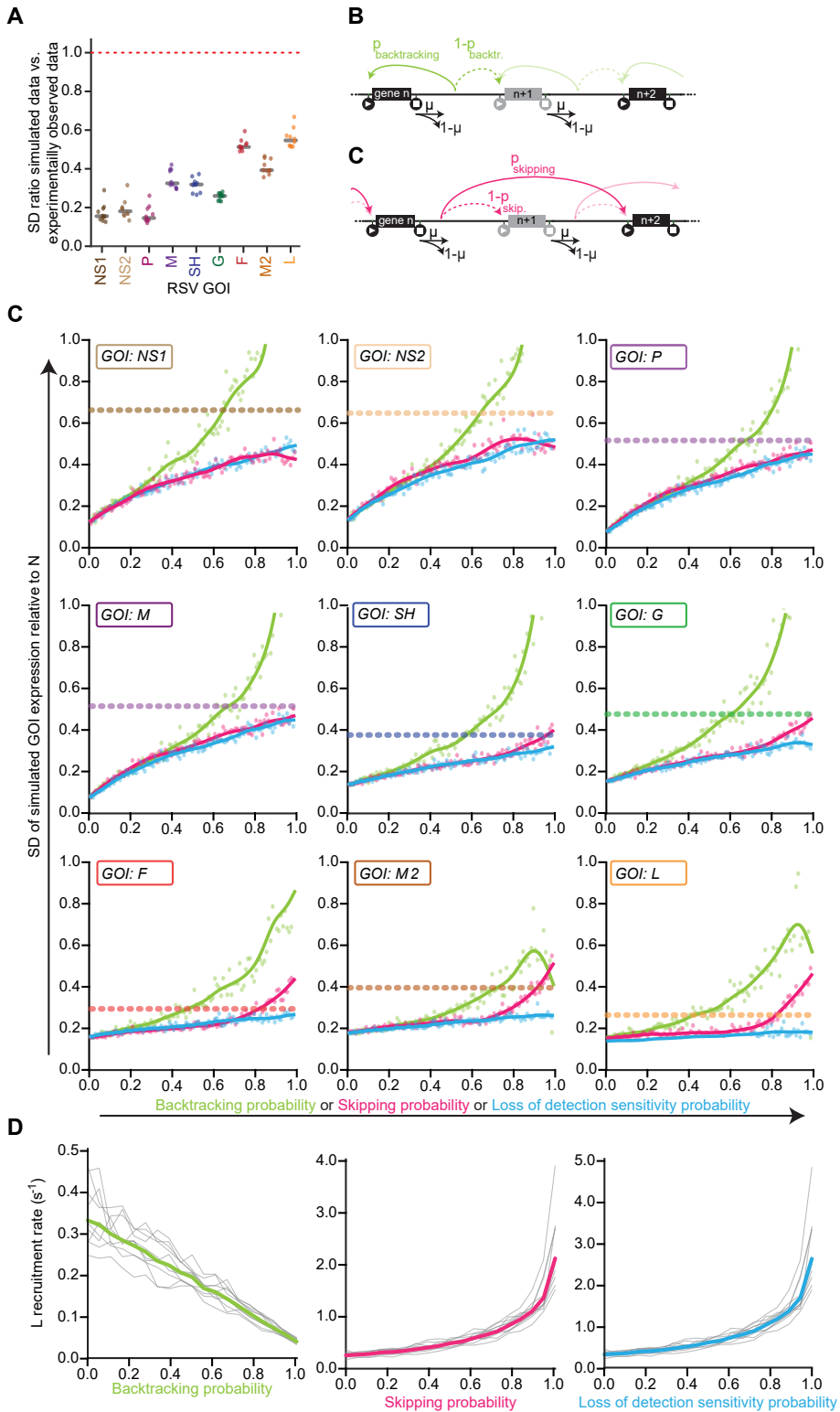
Figure 3. Computational modeling challenges basic viral gene expression gradient model.

(A) Comparison of noise in GOI/N ratio between simulated data (Fig. S3B) and experimentally obtained data (Fig. S2) using SD of each dataset. Each dot indicates a simulation run of ~2500 cells; solid lines indicate mediates; dashed red line indicates 1.0, which would mean similar noise in simulated and experimental data. (B, C) Cartoons of alternative transcription models. Upon transcription termination at a GE, L may continue transcription (μ) or dissociate from the genome (1-μ). In the backtracking model (B) continued transcription may be bidirectionally; L scans the genome backwards to the GS of the gene that was just transcribed (solid green line) or scan towards the GS of the next gene (dashed green line). In the skipping model (C) L continues scanning towards the next GS, which may lead to transcription of the next gene (pink dashed line) or L may fail to recognize the first GS and therefore fails to transcribe the next gene (solid pink line). P means probability. (D) Effect of different noise-introducing models on noise in simulated data for each GOI (indicated in top left corner of each panel). Data in green, pink, or blue originate from simulations with backtracking, skipping, or loss of detection sensitivity models respectively. Each dot indicates the mean of ~2500 simulated cells at indicated noise-inducing probability; solid lines indicate moving average; colored dashed line represents SD of experimentally obtained data. (E) Effect of increasing noise-inducing probabilities on L recruitment rate. Grey lines represent data from individual GOIs; solid line indicates mean. See also Figure S4 and Methods for more details on simulations. The number of simulations per experiment are listed in Table S1.

Development of VIRIM$^{2.0}$: a live-cell early infection viral expression assay

Although smFISH experiments provided insight into coordination between RSV genes, these experiments do not enable live-cell analysis of viral expression. To study temporal dynamics of RSV gene expression, we set-out to developed a live-cell single-molecule viral gene expression assay, inspired by the SunTag-based VIRIM assay (Boersma et al., 2020). The SunTag translation imaging system is based on stable expression of an GFP-coupled intrabody (scFv-sfGFP, called STAb) and fusion of a SunTag peptide array to a protein of interest to which STAbs can bind (Tanenbaum et al., 2014). During translation of an mRNA encoding this protein of interest, SunTag peptides are synthesized and readily bound by STAbs, resulting in GFP recruitment to the nascent peptide chain. As mRNAs are typically translated by multiple ribosomes, many GFP molecules accumulate on the mRNA, which can be detected as a bright fluorescence spot, thereby enabling live-cell analysis of translation on single mRNAs (Morisaki et al., 2016; Pichon et al., 2016; Wang et al., 2016; Wu et al., 2016; Yan et al., 2016). Furthermore, SunTag can also be used to detect translation of single viral RNAs and the number of translating viral RNAs can be used to investigate replication dynamics of +ssRNA viruses (Boersma et al., 2020). SunTag translation imaging can also contribute to studies on viral gene expression, for example by analyzing IRES-mediated translation initiation (Koch et al., 2020). We reasoned that SunTag translation imaging could be additionally used to analyze expression of -ssRNA viruses, such as RSV.

We engineered an RSV strain in which an additional gene was knocked-in between the P and matrix (M) gene: the viral expression reporter gene that encodes an array of SunTag peptides (see Methods; Fig. 4A). During viral transcription, RSV mRNAs including the viral expression reporter mRNAs are produced and each single reporter mRNA can be detected based on the STAb accumulation that occurs during translation of the mRNA. Introduction of the viral expression reporter led
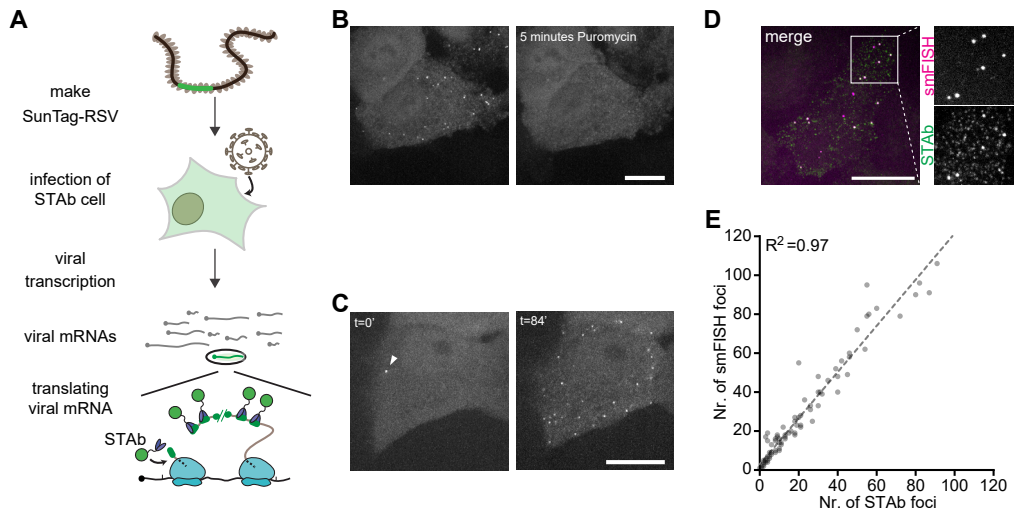
Figure 4. VIRIM$^{2.0}$: a single-molecule virus infection real-time imaging assay to measure viral gene expression in live cells.

(A) Cartoon of VIRIM$^{2.0}$, see main text for details. (B-C) Representative images of timelapse movies of STAb A549 2 – 5 hr after infection with SunTag-RSV. Time in (C) indicates time in minutes since first detection of a STAb spot (arrow). (D) Representative smFISH and STAb staining images of A549 STAb cells infected with SunTag-RSV. (E) Scatter plot of number of smFISH and STAb foci in the same cells. Each dot indicates a single cell and dashed line indicates linear correlation with Pearson R2 (top left). Scale bars, 15 µm. See also Figure S5. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.

to a slight reduction in long-term RSV growth, similar to another engineered strain with an additional gene (RSV-mCherry) (Fig. S5A, B), indicating that introduction of the viral expression reporter genes does not strongly affect the viral life cycle. To test the SunTag-RSV strain, we infected A549 cells expressing STAb (STAb-A549) at a low MOI ($\sim$0.25; Fig. S5C). In a fraction of cells, we observed STAb foci that rapidly disappeared upon treatment with translation inhibitor puromycin, confirming that the foci represent translating viral reporter mRNAs (Fig. 4B). In most infected cells, the number of translating viral reporter mRNAs increased over time (Fig. 4C), suggesting that RSV gene expression was occurring. To examine how well the observed translating viral mRNAs represented the abundance of the reporter mRNAs, we combined smFISH and STAb foci analysis in the same cells. The STAb foci and smFISH foci co-localized well and the number of foci correlated strongly (R2 0.97; Fig. 4D, E), indicating that the number of STAb foci accurately reflects the number of mRNAs. Therefore, the SunTag-RSV strain is a great tool to zoom-in on viral transcription dynamics in live cells during early infection. Because this live-cell single-molecule imaging assay to investigate early viral gene expression of -ssRNA viruses is inspired by VIRIM, which we named the virus gene expression assay VIRIM$^{2.0}$ (Virus Infection Real-time Imaging - version 2).

Live-cell imaging reveals dynamics and heterogeneity of viral gene expression along the RSV genome

To examine RSV transcription dynamics early in infection, we performed VIRIM[2.0] during the first hours of an infection. We observed widespread heterogeneity in viral transcription rates between individual infected cells, in agreement with the widespread cell-to-cell variation in the number of transcripts observed by smFISH (Fig. 5A, S3A). In some infected cells, the number of translating viral reporter mRNAs increased rapidly, indicating a high transcription rate, whereas the transcription rate is low in some other cells (for example Fig. 5A – cell 1 vs. cell 7). Additionally, we observed temporal transcription heterogeneity in single cells; transcription often followed a burst-like ON-OFF pattern (see Methods; Fig. 5B). During the ON time, transcription takes place at a constant rate, whereas during the OFF time the number of translating viral reporter mRNAs plateaus, indicative of a pause in viral transcription. The frequency of the transcription bursts was variable between cells (Fig. 5C, D); no plateaus were observed in ~35% of infected cells, and in the remaining cells the bursting frequency ranged from 0.2- to 1.5 hr^{-1} (for example, Fig. 5A, C – cell 7 and cell 1). Additionally, the timing of the onset of the plateaus was variable; in some cells the first plateau was observed several minutes after the first detection of translating reporter mRNAs, whereas in some other cells viral transcription occurred for a while before a plateau was observed (Fig. 5C, for example Fig 5A - cell 4 vs. cell 2). Moreover, the duration of the plateau was highly variable ranging from 20 minutes to > 2 hr (Fig. 5C). Finally, in approximately 25% of infected cells transcription seemed to have aborted; after a transcription burst and a plateau, we never observed a new burst, implying that the virus failed to continue transcription (for example, Fig. 5 – cell 8 & 9).

The RSV genome from the infecting virion serves as a template for transcription and replication, resulting in mRNAs and more genomes (Noton and Fearns, 2015). We wondered whether the transcription plateaus may reflect replication of the incoming genome; during the plateau there is no transcription, instead the genome would be used for replication. In that case, at the end of a plateau there would be an increase of genomes, which may result in an increased transcription rate, as more genomes became available for transcription. Comparison of the transcription rates prior to and after plateaus revealed no change in transcription rates, and the plateaus therefore unlikely reflect replication (Fig. 5E).

Because viral gene expression depends on the position of the viral gene (Fig. 2), the interpretation of the live-cell expression may be influenced by the knock-in site of the viral expression reporter gene. To enable VIRIM[2.0] based on different locations of the reporter gene, we expanded the collection of RSV strains. In addition to inserting the viral expression reporter gene between P & M (Fig. 4), we also generated strains with inserts between the RSV genes G & F, and F & M2, respectively referred to as upstream, middle, or downstream strains (Fig. 6A, see Methods). To confirm that introduction of the viral expression reporter gene does not hamper viral growth, we performed plaque assays. Although the plaques formed by the middle or downstream strains were slightly smaller compared to plaques from wildtype, mCherry, or upstream strains, the middle and downstream strains produced yields comparable to the upstream strain (Fig. S5A, B). All three reporter strains can thus be used to examine viral gene expression dynamics.
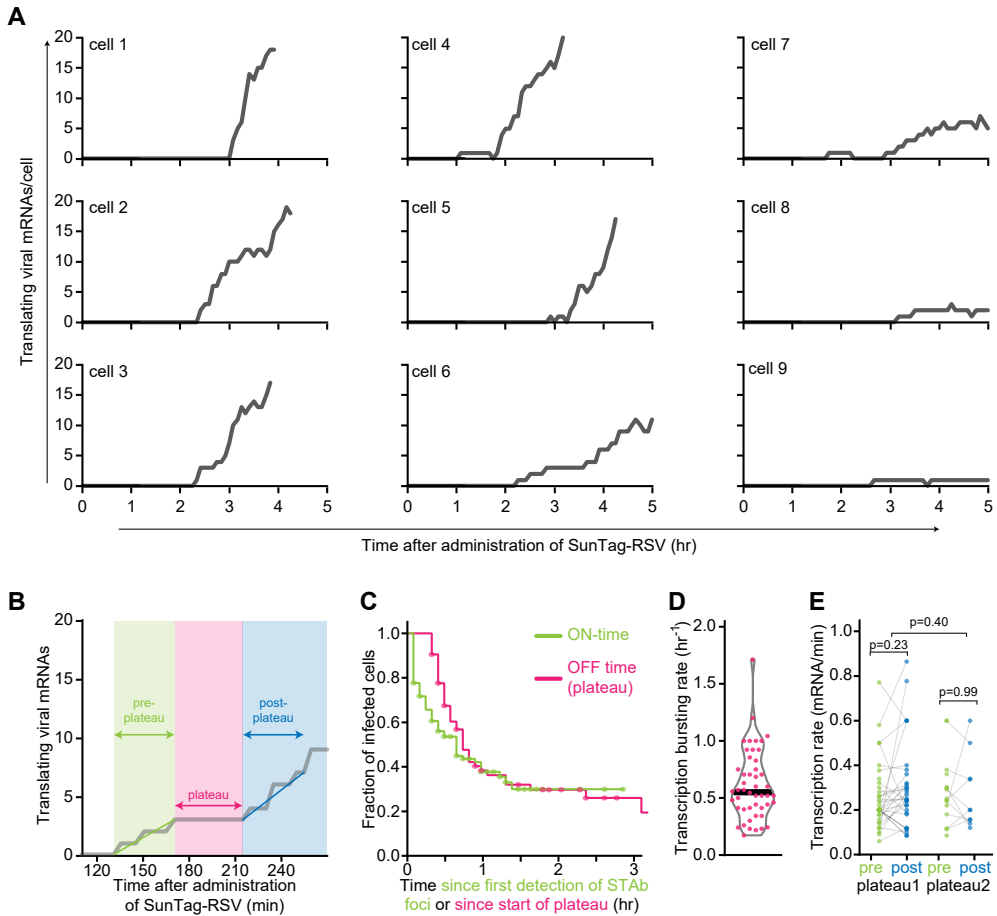
Figure 5. Single-cell and temporal heterogeneity in viral gene expression.

(A) Representative example quantifications of the number of translating viral reporter mRNAs over time in single STAb A549 cells after infecting with SunTag-RSV. (B) Zoom-in on example cell 6 to illustrate the transcription burst-like pattern. (C) Kaplan-Meier graph showing distribution of plateau duration (OFF time) and of the timing between first detection of a translating viral reporter mRNA and start of a plateau (ON time). (D) Violin plot of transcription bursting frequency. Each dot indicates a single cell. Cells without plateaus are excluded from analysis. (E) Comparison of transcription rate before (green) or after a plateau (blue). Statistics based on paired student's t-test. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.

Upon infection of A549 cells, translating viral reporter mRNAs were detected for all three reporter strains. On average, the transcription rate declined between the upstream, middle, and downstream strain (Fig. 6B), similarly to the average gene expression decline along the RSV genome based on smFISH experiment (Fig. 2). However, inspection of transcription in single cells revealed temporal variations in viral transcription; burst-like transcription and abortive transcription were observed for all three strains (Fig. 6C, D). Moreover, the duration between RSV administration and first detection of a translating viral reporter mRNA was variable
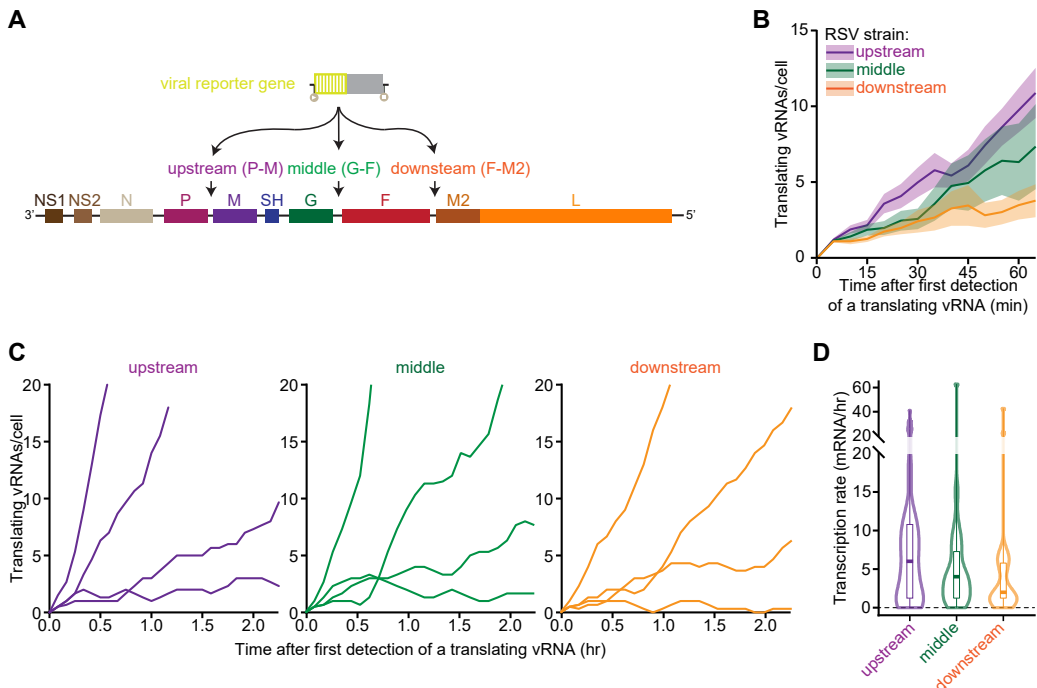
Figure 6. Viral gene expression dynamics and heterogeneity along the RSV genome.

(A) Schematic illustrating the insert location of the viral expression reporter in the RSV genome. (B) Increase in number of translating viral reporter mRNAs based on VIRIM[2.0] using the upstream, middle, or downstream strain. Solid line indicates mean and shaded areas indicate SEM. (C) Representative example quantifications of the number of translating viral reporter mRNAs over time for the upstream (purple), middle (green), and downstream (orange) strains. (D) Boxplots of the transcription rates in single cells upon infection with the indicated strains. See also Figure S5. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.

ranging from 15 minutes to > 12 hr (Fig. S5C), independently of the transcription rates (Fig. S5D).

Taken together, live-cell single-molecule imaging of viral gene expression revealed extensive heterogeneity between cells, over time in single cells, and at multiple locations along the RSV genome, illustrating the power of VIRIM[2.0].

Early viral gene expression vs. activation of antiviral response

Launching an antiviral response is key to preventing a virus from multiplying and spreading to other cells. One of the main pathways leading to activation of an antiviral response depends on the IFN pathways. Upon sensing a virus, TANK-binding kinase1 (TBK) is activated and induces translocation of IFN regulatory transcription factor 3 or 7 (IRF) into the nucleus. Upon nuclear entry, IRF induces IFN gene expression, leading to secretion of IFN and resulting in the expression of ISGs in neighboring cells that can inhibit various viral processes (Schneider et al., 2014; Sun and López, 2017). RSV can prevent IFN induction and IFN-induced antiviral responses well, as
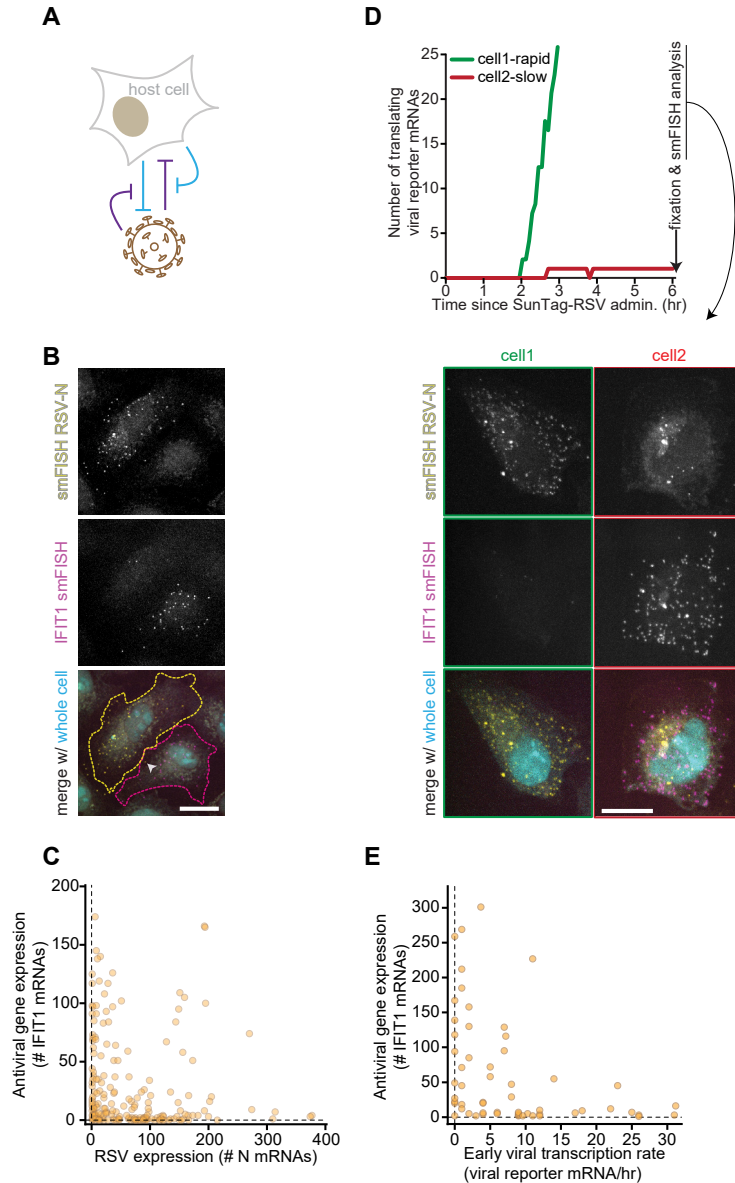
Figure 7. Antiviral response vs. early viral gene expression.

(A) Schematic illustrating virus-host competition. (B) Representative smFISH images of A549 cells infected with RSV illustrating heterogeneity between cells with high (yellow outline) or low (pink outline) number of N mRNAs. Arrowhead indicates presence of a single N mRNA in pink cell. (C) Scatter plot of IFIT1 and RSV-N smFISH analysis. Each circle indicates a single cell with ≥1 N mRNA. (D) Example traces and smFISH images of combined live-cell expression analysis (left) and smFISH analysis (right) of the same cells to compare of early viral expression based on VIRIM 2.0 and activation of antiviral response based on IFIT1. (E) Scatter plot of combined analysis live-cell and smFISH analysis. Each dot indicates a single cell. Scale bare, 15 μm. See also Figure S6. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.

exemplified by the low levels of IFN in RSV patients (Hijano et al., 2019; Sedeyn et al., 2019; Spann et al., 2004).

To study the virus-host competition (Fig. 7A) and specifically how heterogeneity in early viral gene expression relates to activation of the antiviral responses, we aimed to perform combined analysis in the same cells of both early viral gene expression and activation of an antiviral resonse. To score whether a host cell senses a virus infection, we sought to use a marker of IRF transcriptional activity. TBK-induced IRF translocation into the nucleus upon sensing of a virus, leads to expression of IFN as well as IFN-induced protein with tetratricopeptide repeats 1 (IFIT1) (Fig. S6A (DeFilippis et al., 2006; Grandvaux et al., 2002)). In contrast to IRF-dependent expression of IFN, IFIT1 is one of the earliest IRF3/7-dependent genes expressed upon an infection (3 hpi vs. 9 hpi; (Doğanay et al., 2017)). Moreover, basal expression of IFIT1 is low (0.9 ± 1.3 smFISH IFIT1 foci per mock infected A549 cell (mean±SD); n=120 cells from 4 repeats) and an increase in IFIT1 expression is thus readily detectable. Indeed, IFIT1 expression was detected using smFISH after infecting A549 cells with RSV at 6hpi (Fig. 7B). In addition to IRF-dependent expression, IFIT1 is also expressed upon JAK-induced nuclear translocation of signal transducer and activator of transcription 1 or 2 (STAT), which occurs upon IFN exposure to a cell (Fensterl and Sen, 2011). TBK-IRF-dependent IFIT1 expression therefore reflects sensing of a viral infection in a cell, whereas JAK-STAT-dependent IFIT1 expression is an indication of IFN production by any cell in proximity of the IFIT1 expressing cell (Fig. S6A). TBK inhibition, but not JAK inhibition, diminished IFIT1 expression during early RSV infection, confirming that IFIT1 expression during early RSV infection exclusively reports on cell intrinsic virus-sensing and antiviral activation (Fig. S6B). Collectively, IFIT1 expression is a sensitive and specific marker for the host cell's ability to sense an RSV infection and initiate an antiviral response.

IFIT1 expression is variable between individual infected cells, ranging from absent (i.e., similar to mock infected cells) to hundreds of mRNAs per cell (Fig. 7B, C, S6C). Specifically, IFIT1 mRNAs were abundantly present in infected cells with a low number of N mRNAs, whereas IFIT1 expression was scarce in infected cells with high levels of N mRNAs. Only few cells (10% of infected cells) had high numbers of both N and IFIT1 mRNAs (Fig. 7B, C, S6C). Next, we combined live-cell analysis of early viral gene expression and fixed-cell analysis of antiviral response activation by combining VIRIM[2.0] and smFISH against IFIT1 in the same cells (Fig. 7D). In agreement with the N vs. IFIT1 smFISH experiments, early viral expression and potent activation of the antiviral response seemed mutual exclusive in a single cell (Fig. 7E). The early viral transcription rate based on VIRIM[2.0] is calculated using only the first hour of an infection (see Methods). This opposing relation between early viral expression and IFIT1 expression thus implies that viral expression during the first hour is already predictive for the host cell's ability to launch an antiviral response.

DISCUSSION

In this study, we established a collection of tools for single-molecule analysis of viral gene expression dynamics during early RSV infection. Specifically, we generated a smFISH-based analysis pipeline to examine the coordination of expression between individual RSV genes and we developed a live-cell single-molecule imaging assay, VIRIM[2.0], to examine viral transcription in single cells over time. Using

these approaches, we identify polymerase backtracking as a likely mechanism contributing to coordinated viral gene expression and we uncover single cell variations in dynamics of viral gene expression. Additionally, we combined analysis of viral gene expression with analysis of antiviral activation in the same cells and revealed a mutually exclusive relation between early viral gene expression and the host cell's ability to launch an antiviral response, exemplifying how single-molecule imaging of early RSV infection may aid in understanding the competition between RSV infection and an infected host cell.

RSV gene expression dynamics

Based on smFISH analysis of expression of viral genes at early time-points during an RSV infection, we find that the proposed viral gene expression gradient along the genome is recapitulated reasonably well if all single cell measurements are averaged. At a single cell level, there is variation in gene expression dynamics, and, in a subset of cells, the viral gene expression deviates from the gradient model. Previous studies focusing on the expression gradient along the RSV genome used 'bulk' measurements; i.e. generated a single read-out by combining many single cells (Aljabr et al., 2016; Barik, 1992; Piedra et al., 2020) and therefore did not have the sensitivity to detect single cell heterogeneity. Moreover, these previous studied primarily relied on RNA-sequencing or qPCR to measure the relative abundance of RSV mRNAs at relatively late time-points during infection (mostly ≥24 hpi). These approaches may have failed to reliably test the early viral gene expression dynamics, as qPCR-based or RNA-sequencing based analysis of different mRNA species may be influenced by the relatively low GC-content of some RSV genes (ranging from 29 to 44%) and variation in the GC-content may affect the detection sensitivity by qPCR or sequencing (Dohm et al., 2008). For example, the RSV gene with the highest GC-content is G (44%) and incidentally, the previous studies reported that G mRNAs are among the most abundant RSV mRNAs. In contrast, analysis of viral gene expression using the number of viral mRNAs based on smFISH is not affected by the variable GC-content of the RSV genes and smFISH can be used to determine the number of viral mRNAs well for each gene (Fig. 1). Although our smFISH data also suggest that G mRNAs are overrepresented compared to mRNAs from the surrounding genes, the magnitude of the G over-abundance is different (~5-fold vs. ~1.5-fold), illustrating that single-molecule measurements of the absolute number of mRNAs are superior to sequencing in analyzing viral gene expression dynamics.

The basic gradient model does not fully explain observed variations in relative RSV gene expression. Instead, simulation with the basic model supplemented with 50% backtracking efficiency leads to an accurate match with the experimentally obtained data. A backtracking efficiency of 50% suggests bidirectionality of scanning after termination of a gene; L may either scan in 3' (backwards) or 5' (forwards) direction, but there is no dominant scanning direction. Bidirectional scanning by L has consequences for the interpretation of early viral gene expression. For example, early in an infection the concentration of L in the host cell may be low. Bidirectional scanning by L may be a mechanism for the virus to prolong L dissociation from a genome and reduce the requirement to recruit a new L molecule. Indeed, we estimate that recruitment of L is reduced from 0.33 s^{-1} to 0.2 s^{-1}. Moreover, bidirectional scanning by L may have consequences for transcription dynamics of different L

molecules on a single genome. For example, sequentially initiated L molecules may collide and induce release from the genome if the L molecules process the same genome molecule simultaneously in different directions. Future experiments on L dynamics are required to understand the exact consequences of bidirectional scanning.

Applying VIRIM[2.0], we confirm an overall gradient of viral gene expression along the RVS genome, yet we observe substantial cell-to-cell variation in viral transcription rates. Additionally, we identify that RSV transcription is not uniform over time, but instead is highly variable over time in single cells, further highlighting the heterogenous nature of RSV gene expression.

RSV vs. host cell

Rapid and potent activation of an antiviral response by an infected host cell is crucial for successful inhibition of virus infection. At the same time, RSV produces several proteins that efficiently prevent host cell sensing of an infection and launching of an antiviral response (Sun and López, 2017). Our study uncovers that how well an RSV infection progresses through the first hour of an infection is a strong predictor for the host cell's ability to launch an antiviral response. We recently reported that the virus-host competition upon infection by picornaviruses is also initiated rapidly upon infection (Boersma et al., 2020). Initiation and determining the winner of the virus-host competition during the earliest events of an infection may therefore be a general phenomenon among RNA virus infections. Traditional assays lack the sensitivity to study the initiation and outcome of a virus-host competition during early infection. Live-cell single-molecule imaging approaches such as VIRIM and VIRIM[2.0] are therefore crucial tools to decipher dynamics of viral and host cell processes.

The current study identifies an opposing trend between early viral gene expression dynamics and activation of an antiviral response. However, the exact mechanism and dynamics are yet unknown. For instance, it is unclear whether low viral transcription rates are the consequence of an activated antiviral response, or whether low viral transcription rates enable the launching of an antiviral response. Better understanding of the coordination between early viral transcription dynamics and host cell dynamics may help to resolve this question and to identify targets for potential antiviral therapies.

Future directions

This study provides a promising starting point for many follow-up studies and further development of single-molecule imaging tools.

Currently, VIRIM[2.0] provides a single read-out of transcription dynamics from one location along the RSV genome. Based on smFISH analysis, there is some cell-to-cell variation in expression of individual RSV genes and VIRIM[2.0] has revealed widespread temporal variations. Furthermore, we identify polymerase backtracking as a likely mechanism of transcription start site scanning. An interesting follow-up on the current study may combine these observations and try to address how temporal variation in transcription rates relate to the coordination of expression between

genes. For example, how does transcription bursting or a high transcription rate of one gene relate to transcription of another gene and how does a high transcription rate relate to the scanning orientation of the polymerases? Expanding VIRIM[2.0] into a multi-color imaging system by using an RSV strain with multiple reporter genes may shed light on these dynamics. Additionally, single-molecule analysis of L recruitment, transcription initiation rates, and scanning direction may further contribute to understanding viral gene expression coordination.

In this study we addressed early viral gene expression dynamics and the relation between viral transcription and the antiviral response. During an infection, yet another essential viral process takes place: genome replication. Replication of the RSV genome requires the same viral proteins and uses the same template as viral transcription (Noton et al., 2019). Moreover, genome replication may result in the presence of dsRNA in the infected cell (Groskreutz et al., 2006), which is a potent activator of an antiviral response. The dynamics of viral replication may therefore impact viral expression dynamics and launching of an antiviral response. Future studies should combine analysis of replication dynamics with viral gene expression analysis and host cell response analysis.

Many viral processes, like coordination of transcription and replication, are thought to be highly conserved between different *Mononegavirales*. In particular, many conclusions about various *Mononegavirales* are based on extrapolating findings from VSV (Fearns and Plemper, 2017; Noton and Fearns, 2015). We reason that the tools developed here may be extended to also study other *Mononegavirales*. Alternatively, some of our conclusions about early RSV infection may be extrapolated onto other *Mononegavirales* and could therefore help in resolving how other pathogens, such as measles virus or Ebola virus, interact with the antiviral response.

Finally, by expanding the live-cell single-molecule imaging toolset to interrogate early viral infection dynamics, we can now start to compare different RNA virus families to identify unique or common features of viruses and antiviral responses. As mentioned above, both VIRIM and VIRIM[2.0] reveal that early viral infection dynamics may predict the outcome of an infection. VIRIM and VIRIM[2.0] could thereby help to identify the bottleneck for a successful infection of vast range of RNA viruses.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS
SB: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Visualization, and Writing-Original draft; DR: Conceptualization, Methodology; SS: Software, Formal analysis; IB & RB: Investigation; HHR: Conceptualization; LJMB & LJB & JFE & MAR: Resources; MET: Supervision and Funding acquisition.

MATERIAL & METHODS
RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
scFv-sfGFP-StrepII (STAb-sfGFP)	(Boersma et al., 2020)	N/A
Virus and bacteria strains		
hRSV-wt	(Rameix-Welti et al., 2014)	hRSV-62 or vSB401
hRSV-mCherry	(Rameix-Welti et al., 2014)	hRSV-mCherry3 or vSB402
hRSV-GFP	Bont lab	N/A
hRSV-upstream expression reporter	This study	vSB393
hRSV-middle expression reporter	This study	vSB453
hRSV-downstream expression reporter	This study	vSB445
MegaX DH10B T1R Electrocompetent cells	ThermoFisher Scientific	Cat# C640003
Chemicals, Peptides, and Recombinant Proteins		
DMEM	Gibco	Cat# 31966021
Leibovitz's L15 medium	Gibco	Cat# 21083-027
Opti-MEM	Sigma-Aldrich	Cat# 11058-021
Opti-MEM Reduced Serum Medium, GlutaMAX Supplement	ThermoFisher Scientific	Cat# 51985034
MEM	ThermoFisher Scientific	Cat# 42360032
10xMEM	ThermoFisher Scientific	Cat# 11430030
Fetal Bovine Serum (FBS)	Sigma-Aldrich	Cat# F7524
Penicillin-Streptomycin	Gibco	Cat# 15140-122
Lipofectamine 3000	ThermoFisher Scientific	Cat# L3000008
Fugene 6	Promega	Cat# E231A
Polyethylenimine	Polysciences Inc	Cat# 23966
Puromycin	ThermoFisher Scientific	Cat# 12122530
TMC353121 (F inhibitor)	MedChem Express	Cat# 16427117
ALS-8122 (L inhibitor)	MedChem Express	Cat# 16481319
MRT67307 (TBK1 inhibitor)	Sigma-Aldrich	Cat# SML0702-5mg
Tofacitinib (JAK inhibitor)	Sigma-Aldrich	Cat# PZ0017-5mg
MluI-HF	NEB	Cat# R3198
StuI-HF	NEB	Cat# R0187S
BstEII-HF	NEB	Cat# R4162S
Proteinase K	Sigma-Aldrich	Cat# P6556-100MG
Glucose oxidase	Sigma-Aldrich	Cat# G2133-10KU
Catalase	Sigma-Aldrich	Cat# C3515-10MG
Bovine serum albumin (BSA)	Sigma-Aldrich	Cat# A9647-100G
Ribonucleoside vanadyl complex (RVC)	NEB	Cat# S1402S
tRNA from E. coli MRE600	Sigma-Aldrich	Cat# 10109541001
Dextran sulfate	Sigma-Aldrich	Cat# D8906-50G
Trizol	ThermoFisher Scientific	Cat# 15596018
Atto633-NHS	Atto-Tec	Cat# AD 633-31
Atto565-NHS	Atto-Tec	Cat# AD 565-31

I

II

III

IV

V

VI

&

Pacific blue succinimidyl ester	ThermoFisher Scientific	Cat# P10163
Amino-11-ddUTP	Lumiprobe	Cat# 15040
Paraformaldehyde	Aurion	Cat# 15710
Formamide	ThermoFischer Scientific	Cat# AM9342
Avicel RC-591 NF MCC carboxymethylcellulose sodium	Dupont	Cat# NFBA500
Crystal violet solution	Sigma-Aldrich	Cat# HT90132
PEG6000	Sigma-Aldrich	Cat# 804911000
Critical commercial assays		
Terminal deoxynucleotidyl Transferase	ThermoFisher	Cat# EP0162
Tetro reverse transcriptase	Bioline	Cat# BIO-65050
NucleoBond Xtra Maxi kit for transfection-grade plasmid DNA	Marchery Nagel	Cat# 740414.50
Experimental Models: Cell Lines		
U2OS-STAb cells	(Yan et al., 2016)	C5
A549 wildtype cells	Rameix-Welti lab	N/A
A549 STAb cells - STAb-sfGFP + NLS-BFP - STAb-saGFP - STAb-saGFP + BFP-CAAX	(Boersma et al., 2020) This study This study	#1378 #1385 #1383
A549 NbAlfa-sfGFP	This study	#1382
HEp-2 wildtype cells	Rameix-Welti lab	N/A
BHK T7/5 (BsrT7/5) cells	(Buchholz et al., 1999)	N/A
HEK293T cells	Tanenbaum lab	N/A
Oligonucleotides		
See Table S2 for smFISH probes and qPCR primers	This study & (Ruijtenberg et al., 2020)	Table S2 is available upon request
Recombinant DNA		
pCITE-N	(Rameix-Welti et al., 2014)	N/A
pCITE-P	(Rameix-Welti et al., 2014)	N/A
pCITE-M2.1	(Rameix-Welti et al., 2014)	N/A
pCITE-L	(Rameix-Welti et al., 2014)	N/A
pacNR-hRSV (hRSV62)	(Rameix-Welti et al., 2014)	pSB401
pacNR-hRSV-mCherry (hRSV-mCh3)	(Rameix-Welti et al., 2014)	pSB402
pHR BFP-NLS-P2A-N	This study	Rupa
pHR BFP-NLS-P2A-P	This study	Rupa
pHR BFP-NLS-P2A-N-P2A-P	This study	Rupa
pacNR-hRSV-[24xSunTag-BFP] @MluI	This study	pSB393
pacNR-hRSV-[18xAlfaTag-kif18b] @StuI	This study	pSB453
pacNR-hRSV-[12xSunPP7-kif18b] @BstEII	This study	pSB445
Software and Algorithms		
ImageJ	NIH	https://imagej.nih.gov/ij/
NIS-Elements Imaging software	Nikon	https://www.microscope.healthcare.nikon.com/en_EU/products/software
Graphpad Prism 8	GraphPad Software Inc	http://www.graphpad.com/scientific-software/prism/
MATLAB R2012b	The Mathworks, Inc.	https://nl.mathworks.com/products/matlab.html
Other		
96-well glass bottom imaging plates (Matri-plates)	Brooks Life Science Systems	Cat# MGB096-1-2-LG-L

EXPERIMENTAL MODEL AND SUBJECT DETAILS
Cell lines

Human U2OS and HEK293T cells used for imaging and lentivirus production were grown in DMEM (4.5 g/L glucose, Gibco) supplemented with 5% fetal bovine serum (FBS) (Sigma-Aldrich) and 1% penicillin/streptomycin (Pen/Strep) (Gibco) (hereafter abbreviated to DMEM; 5% FBS; 1% Pen/Strep). Unless noted otherwise, human A549 and Hamster BHK-21 for imaging and RSV production were cultured in in DMEM; 10% FBS; Pen/Strep. Human HEp-2 RSV production were cultured in in MEM; 10 % FBS; 1% Pen/Strep, unless noted otherwise. All cells were grown with 5% CO2 at 37°C. Cells were confirmed to be mycoplasma negative.

METHOD DETAILS
Plasmids

The sequences of plasmids generated in this study to produce the various RSV strains are available upon request.

Cell line generation

Cell lines with stable expressing of transgenes were generated using lentiviral transduction. To produce lentivirus, HEK293T cells were transfected using Polyethylenimine (PEI) with the lentiviral plasmid of interest and packaging vectors psPax and pMD2. The cell culture medium was refreshed 1 day after transfection and the supernatant containing the lentivirus was collected 3 days after transfection. To make stable cell lines, cells were seeded 1 day before infection at 40% confluency. Cells were spin-infected with the lentiviral supernatant of interest and Polybrene (10 mg/ml) (Santa Cruz Biotechnology Inc) for 90-120 minutes at 2000 rpm at 25°C. Next, the spin-infection medium was replaced with fresh medium and cultured for at least 4 more days before further analysis. To generate monoclonal cell lines with uniform expression of the transgenes, single cells were sorted into 96-wells plates.

Design and production of RSV

Design

To enable single-molecule live-cell analysis of RSV transcription, we designed RSV strains with a reporter gene whose expression can be visualized (Fig. 6A). We used a previously established recombinant human RSV reverse genetics system, to enable generation of engineered RSV strains with various inserts in different locations along the viral genome. In this reverse genetics system, unique restriction sites were introduced between RSV genes to facilitate cloning with the RSV genome (Rameix-Welti et al., 2014). Specifically, we designed inserts at the following restriction sites: 1) MluI between the P and M gene, referred to as the upstream insert location, 2) StuI between the G and F gene, referred to as the middle insert location, and 3) BstEII between F and M2, referred to as the downstream insert location.

The introduced reporter genes were designed to contain the same gene regulatory sequences as the N gene (gene start (GS), gene end (GE), and 5' and 3' UTRs).

Furthermore, the reporter genes were introduced without affecting the gene regulatory sequences of the gene upstream or downstream of the insert location. The coding sequence of the reporter genes contained a start codon in optimal Kozak sequence (GCCACCATGG), followed by a sequence encoding a SunTag array (Tanenbaum et al., 2014) or AlfaTag array (Götzke et al., 2019) and a downstream gene of interest. For practical reasons, only engineered genomes with the following inserts were used:

- Upstream: 24xSunTag-BFP (pSB393)
- Middle: 18xAlfaTag-kif18b (pSB453)
- Downstream: 12xSunPP7-kif18b (pSB445)

Generation of plasmids encoding recombinant RSV strains

The pacNR plasmids containing the RSV genome with the reporter genes were generated by Gibson assembly of the opened backbone and PCR-amplified inserts. To clear and concentrate the Gibson assembly products, the products were ethanol precipitated. Next, the cloning products were transformed by electroporation into MegaX DH10B T1R Electrocompetent cells (ThermoFisher). After electroporation, the bacteria were recovered in Recovery Medium (ThermoFisher) at 28°C for 2 hr, plated on LB plates, and incubated at 28°C. At least 30 hr after transformation ~25 colonies per plasmid were inoculated in 1.5 ml LB medium and grown at 28°C at 100 rpm. Of note, only relatively small colonies were selected as the large colonies typically contained small side-products of the cloning. Despite the size selection, many colonies (65-100%) turned out to contain truncated versions of the intended cloning product. Therefore, colony-PCRs were performed one day after inoculation to screen for colonies that contained the desired insert and retained the entire RSV genome. For each plasmid, 2-3 positive colonies were then further grown to 400 ml LB at 28°C at 150 rpm for 28-32 hr. The plasmids were prepped with the NucleoBond Maxiprep kit according to the manufacturer's low-yield protocol (Machery-Nagel). To confirm that the products contained the reporter gene inserts and the full RSV genome, the products were test digested. Finally, the correct insertion of the reporter genes was verified by sequencing the insert and the gene upstream and downstream of the insert location.

Generation and rescue of P0 stock

To generate the various RSV strain, BsrT7/5 cells (Buchholz et al., 1999) were transfected as described previously (Bouillier et al., 2019). In brief, for each RSV strain BsrT7/5 cells were plated at ~50% confluency in a 6 wellsplate one day before transfection. Just before transfection, the medium on the cells was replaced with 1.5 ml medium (MEM; 10% FBS; -Pen/Strep). The transfection mix was prepared in 500 µl Optimem Reduced serum (ThermoFisher) containing 10 µl Lipofectamine 3000 (ThermoFisher) and 4 µg DNA of the pCITE plasmids encoding the RSV genes N, P, M2.1, and L, and the pacNR plasmid containing the full RSV genome (ratio N : P : M2.1 : L : pacNR = 4 : 4 : 1 : 2 : 5). The transfection mix was added to the BsrT7/5 cells and incubated for 3 days at 37°C with 5% CO_2. To rescue the virus from the BsrT7/5 cells, the cells were detached from the plastic by vigorous scratching and the cell-virus

suspension was collected. To further release the virus, the suspension was vortexed for > 30 seconds. The resulting suspension was used as the first passage (P0) of the virus strains.

Generation of P1-P3 stock

The rescued P0 stocks or later passages of the various RSV strains were amplified using HEp-2 cells, as described previously (Bouillier et al., 2019). In brief, for each RSV strain HEp-2 cells were plated at ~50% confluency in a 75 cm2 flask one day before amplification. To infect the cells, the medium on the cells was replaced with 3-5 ml P0 virus suspension in medium (MEM; -FBS; -Pen/Strep) and incubated for 2 hr at 37°C with 5% CO2. To prevent production of defective particles, a very low amount of P0 virus was used, corresponding to MOI 0.01 pfu/ml. At the end of the incubation, the virus-containing medium was replaced with low-serum medium (MEM; 2% FBS; 1% Pen/Strep). Depending on the onset of cytopathic effect (CPE), the P1, P2, or P3 was harvested 2-4 days after infection as described at 'Generation and rescue of P0 stock'.

Generation of concentrated and clean working stocks

The virus suspensions of P0-P3 contained increasing titers of virus and could be used successfully in most experiments. However, smFISH and single-molecule antibody experiments revealed that extracellular viral mRNAs and large fragments of RSV-positive cell debris was present in the P0-P3 suspension, probably due to the scratching-mediated harvesting of the virus suspensions. To generated RSV stocks with little cell debris and cleared from the extracellular viral mRNAs, we generated P4 RSV stocks by performing multiple centrifugation steps and PEG-precipitation.

For each RSV strain, HEp-2 cells were plated at 40% confluency in 20 162 cm2 flasks 1 day before infection. To infect the cells, the medium on the cells was replaced with 10 ml of infection medium per flask (MEM; 1% FBS; 1% Pen/Strep; 0.01 pfu/ml from the P2 or P3 stock). The infection medium was incubated for 45-60 minutes at 37°C with 5% CO2 and then 10 ml low-serum medium per flask was added (MEM; 1% FBS; 1% Pen/Strep). Depending on the onset of CPE, the P4 was harvested 3-4 days after infection by collecting the medium from all flasks per RSV strain. To pellet cell debris, the collected virus-containing medium was centrifuged for 5 minutes at 1350 rpm and the pellet was discarded. To precipitate the virus from the supernatant, a cold PEG-solution was added until the ratio between virus-supernatant and PEG-solution was 4:1. The PEG-solution consisted of 50% (m/v) PEG6000 (Sigma-Aldrich), 150 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl (pH 7.4). To ensure mixing of the virus-supernatant and PEG-solution, the PEG-solution was added slowly (~10 minutes) under constant stirring and thereafter stirred for 3-4 hr, under ice-cold conditions. To pellet the precipitated virus, the virus-PEG solution was centrifuged 30 minutes at 3000 rpm and the supernatant was discarded. The concentrated virus was resuspended in 10% (m/v) sucrose in PBS, aliquoted into single-use aliquots, and stored at -80°C.

In comparison to the P2-P3 stocks, the PEG-precipitated P4 stocks were 10-500-fold more concentrated. Very little cell debris or extracellular viral transcripts could be detected in the cleaned stocks.

Plaque titration assay

Plaque assays were performed as described previously (Bouillier et al., 2019) to determine the titers of newly produced RSV strains (P0-P4). In brief, for each stock 6 wells of a 12 wellsplate were prepared by plating HEp-2 cells at 40% confluency one day before infection. A 10-fold dilution series of each stock was prepared in MEM (Gibco) supplemented with 1% penicillin/streptomycin (Gibco). To infect the cells, the medium on each well was replaced with 400 µl of a dilution from the series and the virus was incubated for 2 hr at 37°C with 5% CO2. Then, 2-3 ml of the overlay solution was added to the wells and incubated at 37°C with 5% CO2. The overlay solution consisted of 0.24% microcrystalline cellulose suspension (DuPont) in low-serum medium (MEM; 2% FBS; 1% Pen/Strep).

The plaque titration assay was finalized by staining the cells 6 days after infection. First, the virus-overlay suspension was detached by gentle moving of the plated, the virus-overlay suspension was removed, and the cells were washed twice with PBS. Next, the cells were fixed and stained using the crystal violet solution (8% crystal violet (v/v) (Sigma-Aldrich), 2% formaldehyde (v/v), and 20% ethanol (v/v) in water). The crystal violet solution was removed after 10-20 minutes of incubation at room temperature.

To calculate the titer of each stock, the number of plaques in all wells with distinguishable plaques was counted. The titer was then calculated using the titer formula:

Titer is the concentration of virus in the tested stock in pfu/ml. The dilution depends on the dilution that was used to count plaques. The inoculum volume in all plaque assays was 0.4 ml. The concentration of each stock for each RSV strain was calculated as the mean titer based on the 2 dilutions for which the plaques could be counted readily.

Viral growth assay

To access whether introduction of the reporter genes in the RSV genome affected the viral fitness, viral growth assays were performed on A549 cells. On day before the start of the first timepoint, A549 cells were plated at 10-15% confluency in 6 wellsplates. To start the infection, the medium on the cells was replaced with virus-containing medium (DMEM; 10% FBS; 1% Pen/Strep and virus corresponding to 3 pfu/cell). 1 Hr after the inoculation, the virus-containing medium was washed away with PBS and replaced with 1.5 ml cell culture medium. To determine the production of new viruses, the cells were scratched, collected with the supernatant, and vortexed at the end of the incubation time (1, 12, 24, 36, 48, 72, 96, or 120 hr after inoculation - corresponding to 6 days after plating for all time-points). The viral titers of the suspensions were determined based on plaque assays, as described in 'Plaque titration assay'.

smFISH

Design of probesets

To analyze the abundance of RSV transcripts and RSV genome in infected cells, we designed single-molecule Fluorescence In Situ Hybridization (smFISH) probes targeting the positive-sensed RSV mRNAs or the negative-sensed RSV genome. Custom-made oligonucleotides probesets were designed through two steps. Firstly, we used the webtool www.biosearchtech.com to design probes. To obtain probesets with good sensitivity, we aimed to include 48 oligonucleotides per probeset. As the RSV mRNA target sequences of some RSV genes are quite short and relatively adenine/uracil-rich, the webtool did not design enough probes for most targets. To include more oligonucleotides in the probesets, we secondly designed oligonucleotides manually with maximally 80% adenine/uracil-content. Additionally, we allowed the manually designed oligonucleotides to overlap with the other probes. The overlap allowance was limited to maximally 2 probes targeting any nucleotide in the target (See Table S2 for sequence of the probes and Fig. 1A for number of oligonucleotides per probeset).

As the RSV genome is covered by many nucleoproteins and other viral proteins, we anticipated that the accessibility of a genome molecule for smFISH probes might be limited. To achieve maximal smFISH potential, we included all oligonucleotides designed through the webtool in the probeset against the RSV genome (336 in total).

Labeling of probesets

All RSV smFISH probesets were labeled with Atto565 and Atto633 (Atto-Tec), as described previously (Gaspar et al., 2018). In brief, the NHS-conjugated dyes were dissolved in DMSO and mixed with NaHCO3 (final concentration 0.05 M; pH 8.4) and Amino-11-ddUTP (5 mM; Lumiprobe). For each probeset, a probe-solution was prepared containing 200 µM of each oligonucleotide. Next, the probes-sets were labeled by incubating the probe-solution with 1x Terminal deoxynucleotidyl Transferase (TdT) buffer, 10 mM dye solution, and TdT (ThermoFischer) at 37°C for 16 hr. To obtain the labeled probes and remove unbound dye, the probesets were precipitated with 100% ethanol, washed thrice with 80% ethanol, and resuspended in nuclease-free water to a final concentration of 20-30 µM.

smFISH sample preparation and probe hybridization

Single-molecule Fluorescence In Situ Hybridization (smFISH) was performed, based on previously described protocols (Lyubimova et al., 2013; Raj et al., 2008). At least one day before fixation, A549 cells were plated on glass at 20-50% confluency. To infect cells with RSV, the RSV stain of interest was diluted in medium (either cell culture medium (DMEM; 10% FBS; 1% Pen/Strep) or imaging medium (L15; 5% FBS; 1% Pen/Strep)) and the medium on the cells was replaced with the virus-containing medium. Unless noted otherwise, the virus-containing medium was kept on the cells until the moment of fixation. The time-point of adding the virus-containing medium to the cells was used the moment of inoculation. 5 minutes – 7 hr after inoculation, the virus-containing medium was removed, the cells were washed twice with PBS, and fixed with 3.7% paraformaldehyde (Aurion) in PBS for

5-10 minutes at room temperature. After fixation, 2 PBS-washes were performed, the cells were permeabilized with 100% ice-cold ethanol at 4°C for 30 minutes, and washed twice with smFISH wash buffer (2x SSC, 10 % formamide (ThermoFisher) in nuclease-free water). To hybridize the smFISH probes to their targets, the smFISH probesets of interest (10 nM each) were incubated in hybridization buffer (1% dextran sulfate (Sigma-Aldrich), 2x SSC, 10% formamide (ThermoFisher), 1 mg/ml tRNA (Sigma-Aldrich), 2 mM Ribonucleoside vanadyl complex (NEB), 200 µg/ml BSA (Sigma-Aldrich), in nuclease-free water). The hybridization was incubated at 37°C for 16-24 hr. Next, the cells were washed twice with smFISH wash buffer at 37°C for 1 hr. A whole cell staining was included by incubating the cells in Succs-wash buffer (200 ng/ml Pacific blue succinimidyl ester (ThermoFisher) in smFISH wash buffer) for 15 minutes at room temperature and washing twice with smFISH wash buffer. The wash buffer was replaced with imaging buffer (10 mM Tris, pH 8; 2x SCC; 0.63 % glucose, supplemented with glucose oxidase (Sigma) and catalase (Sigma)) and cells were stored at 4°C until imaging.

Variations in the smFISH protocol

To examine the contribution of proteins interacting with the RSV genome to the heterogeneity in smFISH RSV genome spot intensity, we included proteinase K treatment in the smFISH protocol. The proteinase K treatment was performed after the ethanol-permeabilization and the 2 PBS washes by incubating the cells with proteinase K solution (5 µg/ml proteinase K (Sigma-Aldrich) in 2x SCC)) for 5 minutes at room temperature. Next, the cells were washed twice with 2xSCC and the smFISH protocol was continued from the hybridization step onwards.

As the STAb fluorescence signal is not optimally conserved during the smFISH protocol, we included a STAb incubation step in the smFISH protocol. Before adding the imaging buffer to the cells, purified STAb-sfGFP (1:100) (Boersma et al., 2020) was incubated for 5 minutes in imaging buffer (without glucose oxidase and catalase). After washing away unbound purified STAb with imaging buffer (without glucose oxidase and catalase), imaging buffer was added, and cells were stored at 4°C until imaging.

qPCR

The stability of RSV mRNAs was determined using qPCR. One day before the start of the experiment, A549 cells were plated at 20% confluency in 12 wellsplates. To start the infection, the medium on the cells was replaced with virus-containing medium (DMEM; 10% FBS; 1% Pen/Strep and virus corresponding to 1.5 pfu/cell). 1 Hr after the inoculation, the virus-containing medium was washed away trice with PBS and replaced with cell culture medium containing inhibitors: fusion inhibitor (225nM TMC353121(MedChemExpress)) and L inhibitor (2 µM ALS-8112 (MedChemExpress)). The inhibitor-containing medium was refreshed every 12 hr. To measure transcript abundance, cells were collected by trypsinization at 36, 24, 12, 8, and 0 hr after inoculation (all corresponding to 2.5 days after plating cells). RNA was isolated using Trizol (ThermoFisher) and cDNA was synthesized using random hexamers and Tetro Reverse Transcriptase (Bioline). mRNA levels were determined by qPCRs using SYBR-Green Supermix (Bio-Rad) on a on a Bio-Rad Real-time PCR machines (CFX Connect Real-Time PCR Detection System). See Table S2 for

sequence details of oligonucleotides used for qPCR. All RNA levels were normalized to GAPDH mRNA levels.

Microscopy

Microscope equipment

Imaging experiments were performed with either 1) a Nikon TI inverted microscope with NIS Element Software (Nikon), equipped with a perfect focus system, a Yokagawa CSU-X1 spinning disc, an iXon Ultra 897 EM-CCD camera (Andor), or 2) a Nikon TI2 inverted microscope with NIS Element Software, equipped with a perfect focus system, a Yokagawa CSU-X1 spinning disc and an Prime 95B sCMOS camera (Photometrics). Both microscopes were equipped with a temperature-controlled hood. For experiments with live-cell time-lapse analysis a 60x 1.40 NA oil-immersion objective was used, while a 100x 1.49 NA oil-immersion objective was used for experiments of fixed samples. For experiments in which live-cell imaging and fixed-cell imaging of the same cells was combined, the 60x 1.40 NA oil-immersion objective was used for all imaging.

Cell culture before imaging

To perform live-cell imaging, A549 cells expressing the relevant transgenes were plated 1 day before imaging in a 96-well glass-bottom plate (Matriplates, Brooks Life Science Systems). An MOI of 0.2-0.25 was used to infect the cells, by diluting the RSV strain of interest into imaging medium (pre-warmed CO2-independent Leibovitz's-15 medium (Gibco); 10% FBS; 1% Pen/Step) and replacing the cell culture medium with 200 µl RSV-L15. The MOI was based on the fraction of cells in which translating vRNAs became visible within the first 10 hr after virus administration.

To perform fixed-cell imaging, a similar set-up was used as for live-cell imaging. However, RSV was diluted into cell culture medium instead of imaging medium.

The following combinations of cells and viral strains were used:

RSV strain (generated using plasmid #)	A549 cell line (cell line nr. in Tanenbaum lab library)
Wildtype (pSB401)	wildtype
mCherry (pSB402)	wildtype
Upstream expression reporter (pSB393)	STAb-sfGFP + NLS-BFP (#1378) STAb-saGFP + BFP-CAAX (#1383) STAb-saGFP (#1385)
Middle expression reporter (pSB453)	Nb^Alfa-sfGFP (#1382)
Downstream expression reporter (pSB445)	STAb-sfGFP + NLS-BFP (#1378)

Drug treatment

Translation inhibitor puromycin (0.1 mg/ml; ThermoFischer) was added to the imaging medium at the indicated time-point to investigate viral translation and to confirm that translating viral mRNA foci are sensitive to inhibition of translation.

In experiments in which the pathways contributing to IFIT1 expression were analyzed, cells were treated with the TBK inhibitor MRT67307 (1 µM; Sigma-Aldrich)

or JAK inhibitor Tofacitinib (1 µM; Sigma-Aldrich) 12-16 hr before addition of virus and treatment with these drugs was continued during virus inoculation until the moment of fixation.

Live-cell image acquisition

Time-lapse imaging to analyze the number of translating viral reporter mRNAs per cell was performed using randomly selected x,y-positions. Images were acquired every 2 or 5 minutes for 2-14 hr, using 35-70 ms exposure times. Time-lapse image acquisition was started 10-30 minutes after administration of virus-containing imaging medium. Multiple Z-slices (10-15 planes with 0.8 µm steps) were acquired for GFP to image the entire cell. If relevant, a single BFP image was acquired for BFP, using a Z-slice of the middle of the cells. Time-lapse image acquisition was started 10-30 minutes after administration of virus-containing imaging medium.

Fixed cell image acquisition

For experiments to analyze the relative abundance of viral transcripts (by comparing the number of N and GOI smFISH foci) x,y-positions were selected based on the presence of ≥ 1 N smFISH spot. To acquire images of cells that were fixed and stained after live-imaging, the same x,y-positions were selected as those used for live-imaging. In all other experiments, imaging of fixed cells was performed using randomly selected x,y positions. For Atto565, Atto633, and – if relevant – GFP, multiple Z-slices were acquired to image the entire cell (15-22 planes with 0.5 µm steps; 50-100 ms exposure times each). To detect the outline of the cells, BFP or Pacific Blue images were acquired of the same Z-slices, using 15-50 ms exposure times.

QUANTIFICATION AND STATISTICAL ANALYSIS
Post-acquisition processing of microscopy data

For experiments in which images of multiple Z-slices were acquired per x,y-position, maximal intensity projections were generated using NIS elements software. All downstream analyses and quantifications were performed on the maximal intensity projections. If relevant, bleach correction was performed on images from time-lapses using the ImageJ plugin 'Bleach correction'. In brief, the fluorescence intensity of the total field of view was determined for each time-point. The bleaching rate is calculated by fitting the fluorescence over time with an exponential decay function. The bleaching rate is then used to correct the fluorescence of each time-point.

smFISH quantification

Spot count and intensity

To analyze the relative expression levels of RSV genes, cells were selected based on the presence of ≥ 1 N smFISH spot. For cells on which combined live-cell and fixed-cell analysis was performed, cells were selected based on the live-cell analysis. In the control experiments (without virus infection or to test origin of IFIT1 expression),

cells were selected randomly. A manual ROI around the selected cells was drawn and spots were identified with the ImageJ 'Spot Counter' plugin applying its Gaussian filter and using manually set intensity and size thresholds. Identical parameters were used for the same day controls. The fluorescence intensity of single smFISH foci was calculated based on the 'Spot Counter' spot intensity. For each spot, a background correction was performed based on the total cell fluorescence intensity.

Co-localization

The co-localization frequency of two types of smFISH spots in a multicolor acquisition was determined using the ImageJ plugin 'Spot intensity in all channels' after manually drawing an ROI around the cell of interest. First, an additional channel (selection channel) was generated by overlaying the two channels of interest and scaling both so that smFISH spot of the two colors of interest were equally bright. This selection channel was then added to the multicolor image, resulting in a four-channel image: 1) the total cell (Pacific blue succinimidyl ester) image, 2) smFISH of colorA, 3) smFISH of colorB, and 4) the selection channel. Second, the selection channel was used to select the locations of spots independently of their color, using generous selection (i.e. allowing ~10% false-positives) parameters. Note that to account for minor decentering of co-localizing spots no Gaussian filter was applied. Third, the fluorescence intensity in channel 2 and 3 were measured for each of the locations selected on the selection channel. Additionally, the local background of each spot and channel was measured. Note that the channel 2 and 3 intensities were measured in relatively large ROIs, to ensure that partially decentered channel 2 and channel 3 foci were completely included in the ROIs. Consequently, the measured fluorescence intensities from the co-localization analysis does not accurately reflect the single spot intensity and is only suitable for co-localization analysis. Fourth, it was determined whether each location was positive for either of the two channels based on a manually determined threshold. If a location is positive in both channels, it is called a co-localizing spot. The exact thresholds and parameters were optimized manually for each repeat; identical settings were used for same day controls.

IFIT1 threshold

In the analyses to determine the fraction of cells that is positive for IFIT1 expression, a cut-off of 10 IFIT1 smFISH foci was used. A threshold of 10 foci was selected based on comparable previous experiments (Lucas) and as cells with >10 IFIT1 foci per cell were not detected in mock-infected cells.

mRNA half-lives

The stability of RSV transcripts was calculated using qPCR data. The relative abundance of a transcript over time was fitted with an exponential decay function and this function was used to determine the mRNA half-lives.

Stochastic simulations of RSV transcription to test gradient model

Basic model

To simulate RSV transcription in MATLAB, we started with a simple model, referred to as the 'basic model'. In the basic model, the RNA polymerase is loaded at the 3' end of the genome. Next, each gene is transcribed in a sequential manner as the RNA polymerase engages in repeated cycles of initiation at each GS and termination at each GE. In the basic model, we assumed that initiation occurs in all of the cases the RNA polymerase encounters a GS. Note that, we ignored the fact that the GS of the L gene is upstream of the GE of the upstream gene M2 (i.e. in this case the polymerase would need to scan in the opposite direction before reaching the gene start, but we treated this gene start as if it were downstream of the previous gene). After termination, one of two things occurs: (1) the RNA polymerase continues and reinitiates transcription at the next gene start or (2) the RNA polymerase is dissociates from the genome. We define a 'reinitiation probability' for each gene-to-gene transition (referred to as $\mu_{n->n+1}$), which is the probability that a polymerase remains associated and continues scanning towards the next GS to reinitiate and transcribe the next gene. Taken together, in the basic model, transcription starts with an RNA polymerase transcribing the first gene (i.e. at the 3' end of the RSV genome) an then the polymerase then either continues to the next gene or dissociates from the genome. This process is repeated for consecutive genes until the end of the genome is reached or until the polymerase dissociates from the genome.

To calculate the reinitiation probabilities at ach gene transition, we used the mean transcript ratios that were determined by smFISH (Fig. 2A, B). For example, if the transcript ratio of two consecutive genes is 0.50 (i.e. there are twice as many transcripts of the upstream gene compared to the downstream gene), the reinitiation probability would be 0.50 as well (i.e. only half of the RNA polymerases will continue transcribing the downstream gene). The reinitiation probabilities for each gene-to-gene transition was computed based on the smFISH experiments in which the number of transcripts of each gene was measured simultaneously with the number of transcripts of the N gene. As such, we can compute the transcript ratio of each gene to the N gene (referred to as the GOI/N ratio), but not from each gene to the next one. Therefore, the reinitiation probabilities were calculate using the ratio between different GOI/N ratios. For example, if the GOI/N ratios of two consecutive genes is 0.80 and 0.60, respectively, the reinitiation probability between these two genes is 0.75 (i.e. 0.60/0.80). Note, that the GOI/N ratio of the G gene downstream genes was larger than the GOI/N ratio of the gene directly upstream gene of the G gene (i.e. the SH gene). However, in the basic model downstream genes are per definition expressed at equal or at lower levels than upstream genes, as the initiation probability at a gene start is 1.00 and polymerases can only start at the 3' end of the genome. Therefore, a reinitiation probability of 1.00 was used for the M-SH and SH-G gene transitions (as indicated with asterisks in Fig. 2C and 2D).

Supplementation of basic model

Next to the basic model, we supplemented the basic model with three other modules. As we currently have no experimental data to test the relative contribution of the

different modules, we only supplemented the basic module with the modules individually and we did not combine modules simultaneously. The modules are referred to as the 'backtracking', the 'skipping', and the 'loss of detection sensitivity' modules:

1. Backtracking module

In the backtracking module, we assumed that the polymerase can scan in two directions (Fig. 3B). After transcription of a gene, and if the polymerase remains associated with the genome, the polymerase scans towards the subsequent GS: (1) the GS of the next (i.e. downstream) gene (i.e. GS_{n+1}) or (2) the GS of the same gene that was just finished transcribed (i.e. GS_n). The chance that the polymerase scans backwards or forward, depends on the backtracking probability (pbacktracking): the chance that a polymerase returns to the gene start of the current gene. For simplicity and as we currently have no approach to measure the pbacktracking for each gene individually, we assumed that the pbacktracking is equal at each GS.

In simulations with high pbacktracking , a polymerase can return often to the same GS (i.e. due backwards scanning) instead of continuing to the next gene (foward scanning). During each transcription round (i.e. synthesis of one transcript including the initiation and termination steps) dissociation of the polymerase from the genome can occur. As such, polymerases will have a much higher chance of being removed from the genome at high pbacktracking and fewer polymerases will reach downstream genes. This means that the GOI/N transcript ratios are dependent on the pbacktracking and to accurately reflect the GOI/N ratios observed in the smFISH data we have to update the reinitiation probabilities based on the pbacktracking. We first expressed the reinitiation probabilities as termination probabilities (i.e. the termination probability is complementary to the reinitiation probability; $1 - \mu_{n\to n+1}$). Then, the updated termination probability ($1 - \mu_{n\to n+1, +}$) is described by Equation 1:

$$1 - \mu_{n\to n+1} = \left(1 - \mu_{n\to n+1,+}\right) + \mu_{n\to n+1,+} \cdot p_{backtracking}\left(1 - \mu_{n\to n+1,+}\right) + \mu_{n\to n+1,+}^2 \cdot p_{backtracking}^2\left(1 - \mu_{n\to n+1,+}\right) + \cdots$$
Equation 1

In the backtracking module, a polymerase has multiple opportunities to terminate at a single GE as long as the polymerase keeps backtracking after each transcription round. In the first transcription round, each polymerase has a chance to terminate, yielding the first term the right side of the equation ($1 - \mu_{n\to n+1, +}$). In the second transcription round, only the fraction of polymerases that have not terminated in the first round ($\mu_{n\to n+1, +}$) and have not continued to the next gene (depending on pbacktracking) have a new chance to terminate on the same GE ($1 - \mu_{n\to n+1, +}$). To compute the fraction of polymerases that terminates in the second round, we all the previous three terms are multiplied. For the third transcription round, the same procedure is applied, only now we have to calculate the fraction of polymerases that have not terminated in the first and second round ($\mu_{2n\to n+1, +}$) multiplied by the fraction of polymerases that have backtracked both in the first and second round (p2backtracking). All transcription rounds combined (i.e. from first until infinite) are used to compute the total fraction of polymerases terminating after a particular gene and this equals the termination probability in the basic model ($1 - \mu_{n\to n+1}$; referred to as the standard termination probability). To prevent that the simulations would get

stuck in infinite loops of backtracking, a maximal number of transcription rounds was imposed onto the backtracking module: a polymerase can maximally backtrack 20 times to the same GS. If pbacktracking is low, higher order terms (i.e. frequent transcription rounds) contribute negligibly to transcription of a GOI, as backtracking is rare. In contrast, backtracking is prevalent and higher order terms contribute substantially, if pbacktracking is high. The upper limit of 20 terms does not affect the simulations if pbacktracking <0.90. However, if pbacktracking>0.90, the upper limit of terms interferes with the simulations. Consequently, future studies on very high pbacktracking, may require expansion of the upper limit of terms. Using maximally 20 terms, we can rewrite Equation 2 as a polynomial equation:

$$-\mu_{n\to n+1} + \left(1 - p_{backtracking}\right) \times \mu_{n\to n+1,+} + \left(p_{backtracking} - p_{backtracking}^2\right) \times$$
$$\mu_{n\to n+1,+}^2 + \left(p_{backtracking}^2 - p_{backtracking}^3\right) \times \mu_{n\to n+1,+}^3 + \cdots + \left(p_{backtracking}^{18} - \right.$$
$$\left. p_{backtracking}^{19}\right) \times \mu_{n\to n+1,+}^{19} + \left(p_{backtracking}^{19} - p_{backtracking}^{20}\right) \times \mu_{n\to n+1,+}^{20} = 0$$

Equation 2

We solved Equation 2 at each gene transition and for each simulated backtracking probability and used the real solution that was between 0.00 and 1.00 as the corrected reinitiation probability (see Fig. S4B).

2. Skipping module

In the skipping module, we assumed that an RNA polymerase may skip a GS and, instead, continue scanning to the next GS (Fig. 3C). In the simulations with the basic model supplemented with the skipping module, an extra simulation step is included after the polymerase has transcribed a gene and only if the polymerase remained associated with the genome. During this extra step, the polymerase scans to the next GS (GS_{n+1}), whereupon, (1) the RNA polymerase initiates transcription or (2) the RNA polymerase fails to initiate transcription, but instead continues scanning to the next GS (GS_{n+2}). The chance of transcription initiation on GS_{n+1} or failure to initiate and scanning to GS_{n+2}, depends on the 'skipping probability' (pskipping). The pskipping is defined as the chance that a polymerase fails to initiate transcription upon encountering a GS. For simplicity and as we currently have no approach to measure the pskipping for each gene individually, we assumed that the pskipping is equal at each GS.

In simulations with high pskipping, the polymerase often fails to initiate transcription on a GS, and scans to the next GS. In this scenario, the frequency of termination on some GEs is reduced; polymerases have a much lower chance to dissociate on some GEs and more polymerases will reach downstream genes. Consequently, the GOI/N transcript ratios are dependent on the pskipping and, to accurately reflect the GOI/N ratios observed in the smFISH data, we corrected the reinitiation probabilities based on the skipping probability used in the simulations. We first expressed the reinitiation probabilities as termination probabilities (i.e. the termination probability is complementary to the reinitiation probability; $1 - \mu_{n\to n+1}$). The updated termination probability ($1 - \mu_{n\to n+1}$) is described by Equation 3:

$$1 - \mu_{n\to n+1} = \left(1 - p_{skipping}\right) \times \left(1 - \mu_{n\to n+1,+}\right)$$ **Equation 3**

Here, the termination probability in the basic model ($1 - \mu_{n\to n+1}$, referred to as the

standard termination probability) equals the updated termination probability in the skipping modules multiplied by the fraction of polymerases that can terminate. The fraction of polymerases that can terminate depends on the pskipping (i.e. if more polymerases skip initiation, then more polymerases also skip the termination step) and is described by 1 - pskipping. We rewrote Equation 3 into Equation 4 to express the updated reinitiation probability in terms of the standard reinitiation probability and the skipping probability (note that we can easily switch here between reinitiation and termination probability as they are complementary).

$$\mu_{n \to n+1,+} = 1 - \frac{1 - \mu_{n \to n+1}}{1 - p_{skipping}}$$
 Equation 4

By using the updated reinitiation probabilities in the skipping module, the simulated GOI/N ratios reflect the GOI/N ratios observed in the smFISH data. However, at high skipping probabilities the correction on the reinitiation probabilities does not suffice anymore(>0.70, see Fig. S4B) ; at high skipping probabilities all polymerases that initiate transcription of a gene are removed from the genome after termination. As a result, at even higher skipping probabilities the reinitiation probability cannot be further corrected. Therefore, the skipping module can only be used until pskipping reaches ~0.70; at higher skipping probabilities the GOI/N ratio no longer matches the GOI/N ratios observed in the smFISH data.

3. Loss of detection sensitivity module

In the loss of detection sensitivity module we assumed that some of the transcripts that are synthesized are 'lost' and not detected in the smFISH data, for example due to decay of the transcript or false-negative detection of the transcript. Note that the smFISH validation experiments indicated that our smFISH probesets are well-suited to study early RSV gene expression, but we cannot exclude some variations in detection of RSV mRNAs. To include loss of transcripts in the model we added a new parameter describing the 'loss probability' (ploss), which is the probability that a transcript is not detected even though the transcript was generated. In the simulations with the basic model supplemented with the 'loss of detection sensitivity module', an extra simulation step is included after synthesis of a transcript, resulting in: (1) failure to detect the transcript or (2) transcript detection . The chance of detecting or failure to detect the transcript is 1-ploss or ploss respectively. For simplicity and as we currently have no approach to distinguish the loss of probability for each gene individually, we assumed that the loss probability is equal for each mRNA.

Simulations

We simulated RSV transcription using the 4 models described in the previous section (i.e. the basic model or the basic model supplemented with either of the three modules). We performed two types of simulations: (1) investigating the simulated GOI/N ratios vs. the GOI/N ratios from the smFISH data and (2) investigating the polymerase recruitment rates based on the absolute number of transcripts and by including a time component in the simulations.

Note that for all simulations, we used uncorrected smFISH data; no correction for

detection sensitivity, probeset aspecificity, or transcript half-lives were performed, as we reasoned that potential corrections would not strongly improve the simulations (see Fig. 1). Moreover, we did not correct for changes over the course of an infection; i.e. the µ per gene, ploss, pskipping, pbacktracking, polymerase recruitment rate, etc. were constant values in each simulation round and not subjected to changes over the course of the infection.

In the first type of simulation, we aimed to compare the variation in the simulated GOI/N ratios to the variation in the experimentally-obtained GOI/N ratio based on the smFISH experiments.

To this end, we first divided the smFISH data in bins based on the number of N transcripts (i.e. the first bin contains cells that have 1-10 N transcripts, the second bin contains cells that have 11-20 N transcripts, etc.), and computed how many cells are present in each bin. We performed this binning procedure because we observed that the variation in GOI/N ratio depended on the amount of N transcripts (see Fig. S3B; i.e. at lower number of N transcripts there is more variation in the GOI/N ratio of individual cells). Next, we simulated for each bin the same amount of cells as were present in the data. To simulate a single cells, we determined the number of N transcripts at which the simulation is stopped (i.e. a random number taken from that bin, referred to as the threshold number) and we performed the simulation until the threshold number of N transcripts was reached. For each cell, we performed above described procedure 10 times (i.e. we simulated all data points 10 times). Next, we computed the standard deviation of the simulated data (taking together all bins) for each of the 10 repeats and averaged the standard deviation observed in the 10 repeats. This comparison was made for each GOI individually. For the basic model with the modules, we performed these simulations at increasing ploss, pskipping, and pbacktracking (0.01 - 0.99), respectively. The comparison of the basic model with the smFISH data is plotted in Fig. 3A and the comparison of the smFISH model with the various modules is plotted in Fig. 3C.

In the second type of simulation, we aimed to determine the number of polymerases recruited to initiate transcription at the 3' end of a genome. To this end, a new component was included to all 4 models: the amount of polymerases recruited to the 3' end of the genome during the simulation time. Moreover, we performed simulations for a certain amount of 'simulation time' and assessed the number of transcripts generated in this time (instead of stopping the simulations once a threshold number of N transcripts is reached). We performed simulations for 3 different simulation times, namely 2 hpi, 4 hpi, and 6 hpi, as the smFISH data was also obtained at these 3 time-points. In each simulation, we simulated 1000 cells and for each cell the simulation is comprised of two parts: (1) determine how many polymerases are recruited to the 3' end of the genome and (2) how many transcripts are made by each polymerase. The second part was performed as described above (simulation-type 1). Importantly, we assumed that transcription is instantaneous: if a polymerase is recruited to the 3' end of the genome before the simulation time is finished, it will continue transcription until it dissociates from the genome or until it reaches the end of the genome (i.e. transcription itself doesn't take any simulation time).

To determine how many polymerases are recruited to the 3' end of the genome in the simulation time, we determined (1) the time until a cell is infected and (2) the recruitment times of consecutive polymerases. The time from the start of the simulation (i.e. administration of the virus) until infection is referred to as the 'infection time' and we determined the infection time by drawing a random number between 0 and 6 from an exponential distribution, which was fitted to the experimentally observed infection time distribution (see Fig. S5C). Next, we determine how many polymerases can be recruited to the 3' end during the remaining simulation time (i.e. the total simulation time - the infection time). To this end, we determine the recruitment time of individual polymerases by drawing random numbers from an exponential distribution (referred to as polymerase recruitment time distribution) and compute how many polymerases are recruited within the remaining simulation time (the recruitment time of each consecutive polymerase is subtracted from the remaining simulation time). The simulations were performed using multiple polymerase recruitment time distributions with distinct mean values and the simulation that resulted in the best fit with the data was selected and the mean of the distribution was recorded.

To assess which polymerase recruitment time distribution resulted in the best fit with the data, we defined a cost function by calculating the sum of squared errors (SSETotal) between the simulation and the smFISH data. The total SSE is calculated with Equation 5 and is comprised of the SSE for the 2 hpi, 4 hpi, and 6 hpi time-points.

$$SSE_{Total} = SSE_{2\ hpi} + SSE_{4\ hpi} + SSE_{6\ hpi} \qquad \text{\textbf{\textit{Equation 5}}}$$

To calculate the SSE for each time-point, we randomly selected from the simulated cells as many simulation points as there are data points for each specific time-point. Next, we paired each simulation point to a data point in such a way that the total sum of distances between paired simulation and data points is minimized (Equation 6).

$$SSE_{n\ hpi} = \min \left[\sum_{i=1}^{N} |data_i - simulation_i| \right] \qquad \text{\textbf{\textit{Equation 6}}}$$

We repeated this procedure 10 times and computed for each time-point an average SSE score and added the average SSE scores of the three time-points together to calculate the total SSE score (equation 5). Finally, by employing a minimization function in MATLAB (fminbnd), we searched in each simulation (i.e. at different ploss, pskipping, or pbacktracking) which polymerase recruitment time distribution minimizes the total SSE score (Equation 5), record the mean polymerase recruitment time of this distribution, and convert the mean polymerase recruitment time to the polymerase recruitment rate (inverse of the mean polymerase recruitment time, see Fig. 3D). Note that the resulting polymerase recruitment rate reflects the mean number of polymerases recruited to any genome molecule during the simulation time; the simulations did not correct for the total number of genomes per cell or an increase in genomes per cell during the simulation time.

Counting the number of translating viral reporter mRNAs per cell

Quantification of the number of translating viral reporter mRNAs per cell was performed, based on previously described guidelines (Boersma et al., 2020). In

brief, cells with spots were selected randomly to be included in the quantification. However, only cells that were in the FOV for the duration of analysis were selected, maximally 5 cells per x,y position were selected, and cells that underwent mitosis during the time-lapse were excluded from analysis.

For each selected cell, the number of GFP spots was determined for each time-point either until the end of the time-lapse or until individual spots could no longer be distinguished and counted. Loss of spot-calling confidence can be caused by 1) an accumulation of mature (i.e. released from ribosomes) proteins, which decreases the homogeneity of the background; 2) an increase in crowdedness of the spots, which complicates distinguishing individual spots from each other; or 3) a reduction in freely available STAbs to bind to nascent chains due to increasing accumulation of mature proteins (point1) and translating viral mRNAs (point2), which reduces the brightness of translating viral reporter mRNAs. The exact time-point after which the number of spots in a cell could no longer be counted confidently, differed between cells, and was determined manually for each cell.

The number of translating viral reporter mRNAs per time-point per selected cells was determined by manually counting all bright GFP spots. However, a few additional criteria were considered in the counting of the GFP spots: 1) in some cells (both infected and non-infected), we observed large and relatively immobile GFP foci that represent background foci. Based on the morphology and mobility of these foci, these foci were not included in counting the number of GFP spots; 2) in many cells expressing STAb, we have observed one or two GFP foci next to the nucleus, probably indicating STAb's weak affinity for centrosomes. As these foci were observed independently of infection in uninfected cells, we excluded these foci from analysis; and 3) as translating mRNAs are highly mobile, a single translating viral mRNA spot can often be observed in >1 slice of a Z-stack at slightly different x,y-coordinates in the different Z-slices. In the resulting maximal intensity projection, two very close together spots can be observed. To circumvent double counting of the same translating viral mRNA, we quantified only one spot, if we observed two spots together.

Analysis of transcription rates and burst-like behavior

To calculate transcription rates, the first hour of each trace was fit with a linear function and the slope of this function was used as transcription rate. Analysis of transcription burst-like patterns was performed based by manually annotate plateaus in the traces of each infected cell. Plateaus were defined as periods of >20 minutes during which the number of foci per cells does not increase. No restrictions regarding pre-, inter-, or, post-plateau timings were enforced. To compare transcription rates before and after a plateau, the transcription rate was determined by linear fitting. As the pre- and post-plateau timing distribution is wide (Fig. 5C) and as some of these timings are short, the linear fitting was performed over 30 minutes, instead of an hour. Moreover, for some traces an even shorter pre-/post-plateau period was used to determine the transcription rate, for example if the plateau is initiated rapidly after the start of a trace or a preceding plateau. In those cases, only relevant time-points (i.e. part of trace and not part of a plateau) were included in transcription rate calculations.

SUPPLEMENTAL ITEMS

Table S1 – related to all figures. Number of experimental repeats, cells, and spots analyzed per experiment.

Overview of the number of experimental repeats, cells, and mRNAs that were used to generate all the graphs. If relevant, (re)plotting of the same datasets is indicated in the last column.

Fig.	panel	Description	color of dataset	repeats	datapoints	units of datapoints	notes
1	1B	smFISH intensity distribution NS1	darkbrown	4	3740	foci	
		smFISH intensity distribution NS2	lightbrown	4	4390	foci	
		smFISH intensity distribution N	grey	2	2379	foci	
		smFISH intensity distribution P	pink	4	4687	foci	
		smFISH intensity distribution M	purple	4	2378	foci	
		smFISH intensity distribution SH	blue	4	2375	foci	
		smFISH intensity distribution G	green	4	2096	foci	
		smFISH intensity distribution F	red	4	2130	foci	
		smFISH intensity distribution M2	darkorange	4	1919	foci	
		smFISH intensity distribution L	orange	4	1427	foci	
	1C	smFISH spots mock NS1	darkbrown	3	95	cells	
		smFISH spots mock NS2	lightbrown	3	67	cells	
		smFISH spots mock N	grey	3	64	cells	
		smFISH spots mock P	pink	3	72	cells	
		smFISH spots mock M	purple	3	71	cells	
		smFISH spots mock SH	blue	3	85	cells	
		smFISH spots mock G	green	3	98	cells	
		smFISH spots mock F	red	3	72	cells	
		smFISH spots mock M2	darkorange	3	74	cells	
		smFISH spots mock L	orange	3	75	cells	
		N-Atto565 + P-Atto633		2	151 ; 8202	cells ; foci	
		N-Atto565 + 336ottA-P		2	105 ; 12136	positions ; foci	
		N-Atto633 + P-Atto565		2	154 ; 6366	cells ; foci	
		N-Atto565 + 565ottA-P		2	101 ; 9324	positions ; foci	
	1E	qPCR-based mRNA half-life		4	-	-	based on S1C
2	2A	N vs NS1 2hpi	darkbrown	1	28	cells	
		N vs NS1 4hpi	darkbrown	2	49	cells	
		N vs NS1 6hpi	darkbrown	2	29	cells	
		N vs NS2 2hpi	lightbrown	1	29	cells	
		N vs NS2 4hpi	lightbrown	2	48	cells	
		N vs NS2 6hpi	lightbrown	2	30	cells	

		N vs P 2hpi	pink	2	51	cells	
		N vs P 4hpi	pink	2	57	cells	
		N vs P 6hpi	pink	2	32	cells	
		N vs M 2hpi	purple	2	46	cells	
		N vs M 4hpi	purple	2	54	cells	
		N vs M 6hpi	purple	2	27	cells	
		N vs SH 2hpi	blue	2	54	cells	
		N vs SH 4hpi	blue	2	55	cells	
		N vs SH 6hpi	blue	2	27	cells	
		N vs G 2hpi	green	2	55	cells	
		N vs G 4hpi	green	3	100	cells	
		N vs G 6hpi	green	2	14	cells	
		N vs F 2hpi	red	2	52	cells	
		N vs F 4hpi	red	2	52	cells	
		N vs F 6hpi	red	2	30	cells	
		N vs M2 2hpi	darkorange	2	56	cells	
		N vs M2 4hpi	darkorange	2	52	cells	
		N vs M2 6hpi	darkorange	2	30	cells	
		N vs L 2hpi	orange	2	52	cells	
		N vs L 4hpi	orange	2	56	cells	
		N vs L 6hpi	orange	2	34	cells	
	B	N vs GOI 2, 4, 6 hpi	variety	2	1199	cells	based on Fig. 2A
3	3A	comparision in variation	variety	10		simulations	
	3C	simulated SD	green, pink, blue	10		simulations	
4	4E	STAb vs smFISH foci	grey	2	86	cells	
5	5C	on-time	green	6	103; 82	on-times ; cells	
		off-time	pink	6	82; 63	plateaus; cells	
	5D	bursting frequency	pink	6	48	cells	
	5E	pre-plateau1	green	6	48	cells	
		post-plateau 1	blue	6	32	cells	
		pre-plateau2	green	6	15	cells	
		post-plateau2	blue	6	10	cells	
6	6B	upstream	purple	6	82	cells	
		middle	green	1	25	cells	
		downstream	orange	1	25	cells	
	6D	upstream	purple	6	82	cells	based on Fig. 6B

		middle	green	1	25	cells	based on Fig. 6B
		downstream	orange	1	25	cells	based on Fig. 6B
7	7C	N vs IFIT1	yellow	3	194	cells	
	7E	RSV Txn rate vs IFIT1	yellow	4	60	cells	
S1	S1C	qPCR-based mRNA half-life	variety	4	-	-	used to generate fig. 1E
S2	S2A	smFISH intensity N from N vs N	pink	4	3977	foci	
		smFISH intensity N from N vs GOI	grey	4	26362	foci	
	S2B	N dual-labeling	grey	5	132 ; 5114	cells; foci	
	S2C	N vs N 2hpi	grey	2	56	cells; foci	
		N vs N 4hpi	grey	3	74	cells; foci	
		N vs N 6hpi	grey	1	25	cells; foci	
S3	S3B	N vs NS1 2hpi	darkbrown	1	28	cells	based on Fig. 2A
		N vs NS1 4hpi	darkbrown	2	49	cells	based on Fig. 2A
		N vs NS1 6hpi	darkbrown	2	29	cells	based on Fig. 2A
		N vs NS2 2hpi	lightbrown	1	29	cells	based on Fig. 2A
		N vs NS2 4hpi	lightbrown	2	48	cells	based on Fig. 2A
		N vs NS2 6hpi	lightbrown	2	30	cells	based on Fig. 2A
		N vs P 2hpi	pink	2	51	cells	based on Fig. 2A
		N vs P 4hpi	pink	2	57	cells	based on Fig. 2A
		N vs P 6hpi	pink	2	32	cells	based on Fig. 2A
		N vs M 2hpi	purple	2	46	cells	based on Fig. 2A
		N vs M 4hpi	purple	2	54	cells	based on Fig. 2A
		N vs M 6hpi	purple	2	27	cells	based on Fig. 2A
		N vs SH 2hpi	blue	2	54	cells	based on Fig. 2A
		N vs SH 4hpi	blue	2	55	cells	based on Fig. 2A
		N vs SH 6hpi	blue	2	27	cells	based on Fig. 2A
		N vs G 2hpi	green	2	55	cells	based on Fig. 2A
		N vs G 4hpi	green	3	100	cells	based on Fig. 2A
		N vs G 6hpi	green	2	14	cells	based on Fig. 2A
		N vs F 2hpi	red	2	52	cells	based on Fig. 2A
		N vs F 4hpi	red	2	52	cells	based on Fig. 2A
		N vs F 6hpi	red	2	30	cells	based on Fig. 2A
		N vs M2 2hpi	darkorange	2	56	cells	based on Fig. 2A
		N vs M2 4hpi	darkorange	2	52	cells	based on Fig. 2A
		N vs M2 6hpi	darkorange	2	30	cells	based on Fig. 2A
		N vs L 2hpi	orange	2	52	cells	based on Fig. 2A

		N vs L 4hpi	orange	2	56	cells	based on Fig. 2A
		N vs L 6hpi	orange	2	34	cells	based on Fig. 2A
	S3A	2hpi	blue	2	482	cells	based on Fig. 2A
		4hpi	pink	2	518	cells	based on Fig. 2A
		6hpi	purple	2	290	cells	based on Fig. 2A
S4	S4A	N vs GOI	variety	10		simulations	
	S4C	N vs GOI	variety	10		simulations	
S5	S5B	viral growth	variety	2	-	-	
	S5C	upstream	purple	2	163	cells	
		middle	green	2	193	cells	
		downstream	orange	3	226	cells	
	S5D	upstream	purple	6	82	cells	based on Fig. 6D
		middle	green	1	25	cells	based on Fig. 6D
		downstream	orange	1	25	cells	based on Fig. 6D
S6	S6B	no virus ; no drug	black	4	120	cells	
		virus inc.; no drugs; no viral smFISH foci	black	4	249	cells	
		virus inc.; no drugs; viral smFISH foci	red	4	195	cells	
		virus inc.; TBKi; no viral smFISH foci	black	3	198	cells	
		virus inc.; TBKi; viral smFISH foci	red	3	116	cells	
		virus inc.; JAKi; no viral smFISH foci	black	3	78	cells	
		virus inc.; JAKi; viral smFISH foci	red	3	68	cells	

Table S2 – related to Methods. Overview of the oligonucleotides used for smFISH and qPCRs.

Available upon request.
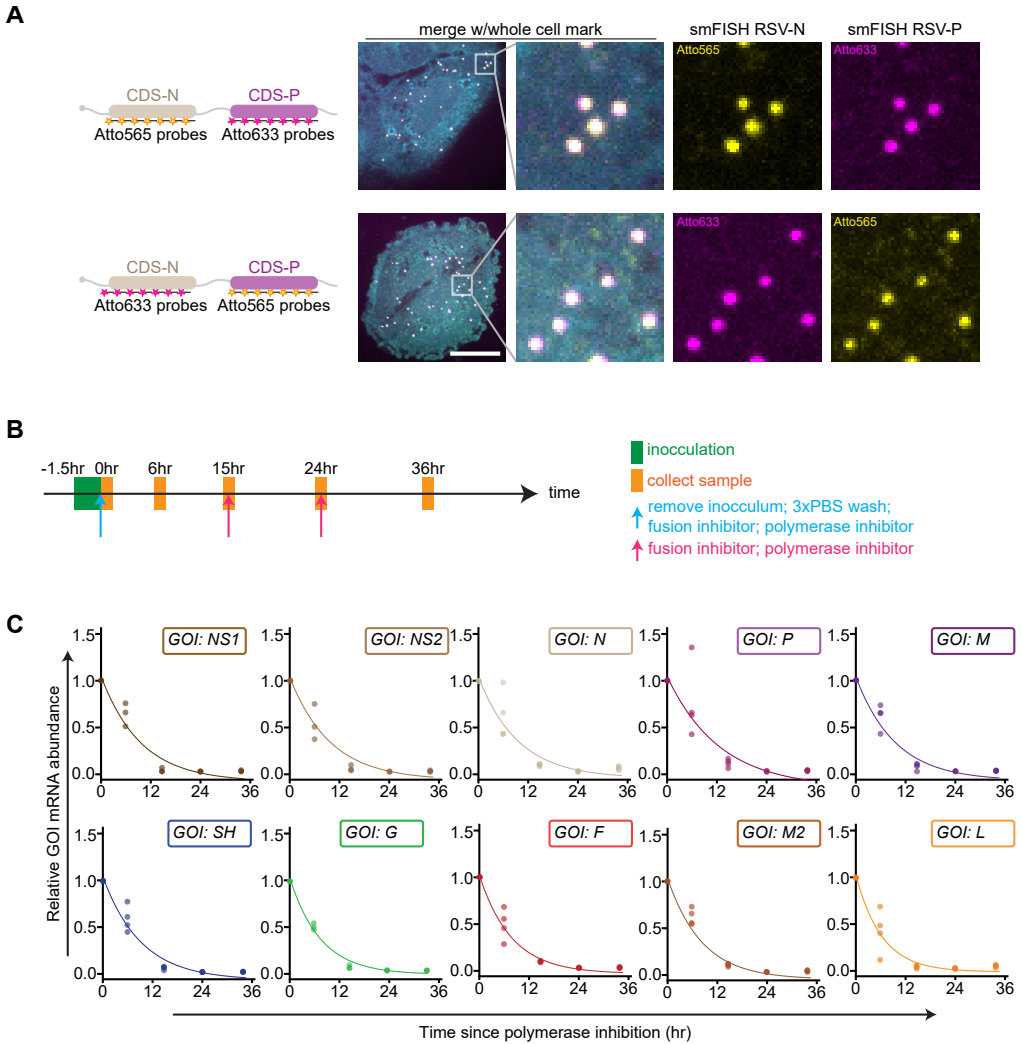
Supplemental figure 1 - related to figure 1. smFISH probesets are suited to study early RSV gene expression dynamics.

(A) Schematics (left) illustrating N-P fusion mRNAs and smFISH probesets to test smFISH detection efficiency. Representative smFISH images (right) of A549 cells stably expressing N-P fusion mRNAs. Scale bare, 15 μm. (B) Schematic timeline of experiment to determine RSV mRNA half-lives. (C) qPCR analysis of RSV GOI transcript abundance, relative to GAPDH, normalized to drug-untreated control. Each dot indicates a repeat; lines represent the best one-component exponential fit to the mean of all repeats. The x- and y-axis labeling are equal for each panel and are indicated with arrows. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.
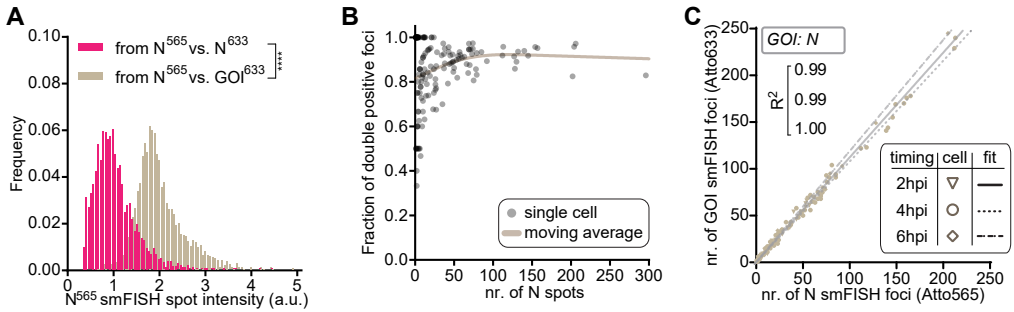
Supplemental figure 2 – related to figure 2. Exclusion of N vs. N from smFISH gene expression analysis.

*(A) Distribution of N-Atto565 smFISH spot intensity if N-Atto565 is combined with N-Atto633 or with any other GOI. ****p < 0.0001 based on unpaired Mann-Whitney test. (B) Dual-labeling efficiency of N-Atto565 and N-Atto633. (C) Correlations between number of N-Atto565 and N-Atto633 smFISH foci in infected cells. The Pearson R2 of the correlations (2, 4, 6 hpi top-to-bottom) is indicated underneath the GOI in the top left corner. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.*

Figure S3 continued:

top left corners indicates frequency of cells with unexpected GOI:N ratio (i.e. <1 for NS1 and NS2 or >1 for all other GOIs). The number of experimental repeats and cells analyzed per experiment are listed in Table S1.
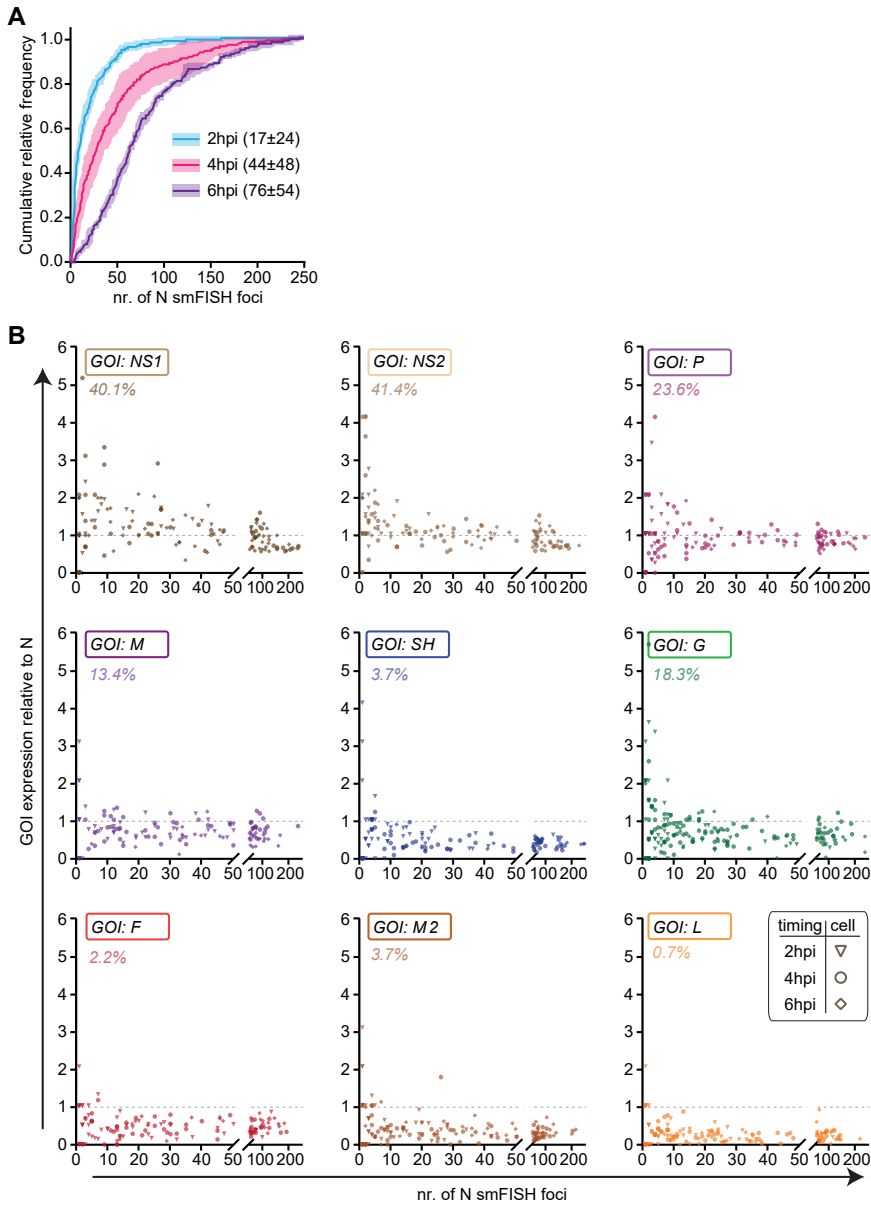
Supplemental figure 3 - related to figure 2. Single-cell heterogeneity in RSV mRNA abundance.

(A) Cumulative frequency distributions of the number of N transcripts per cell. Solid line indicates mean, shaded area indicates SEM of repeats. Values in brackets in legend refer to mean ±SD of single cells. (B) Scatterplots of number of N and relative number of GOI smFISH foci in single cells. Data is subdivided into nine panels, based on the GOI, as indicate in the top left corners. The x- and y-axis labeling are equal for each panel and are indicated with arrows. The L panel includes an additional legend to clarify symbols: each triangle, circle, or diamond indicates a single cell from 2, 4, or 6 hpi experiments respectively. Dashed line indicates GOI vs. N ratio of 1.0. Percentages in
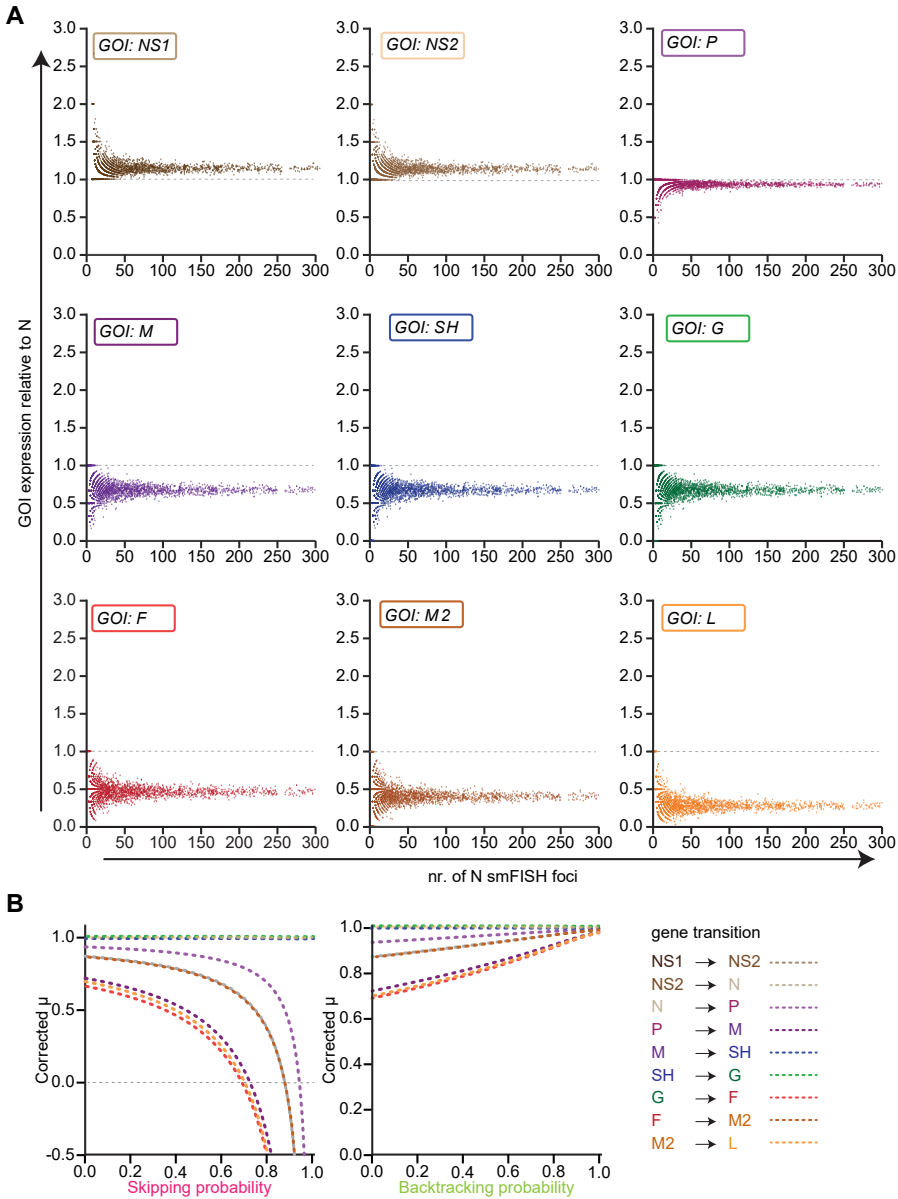
Supplemental figure 4 - related to figure 3. Computational simulations to test gradient model.

(A, C) Scatterplots of simulations using the basic model (A) or basic model supplemented with backtracking model (probability 0.5) on the number of N and relative number of GOI smFISH foci in single cells. Data is subdivided into nine panels, based on the GOI, as indicate in the top left corners. The x- and y-axis labeling are equal for each panel and are indicated with arrows. The L panel includes an additional legend to clarify symbols: each triangle, circle, or diamond indicates a single cell from 2, 4, or 6 hpi experiments respectively. Dashed line indicates GOI:N ratio of 1.0. Percentage in top right corners (C) indicates frequency of cells with unexpected GOI:N ratio (i.e. <1

Figure S4 continued:

C
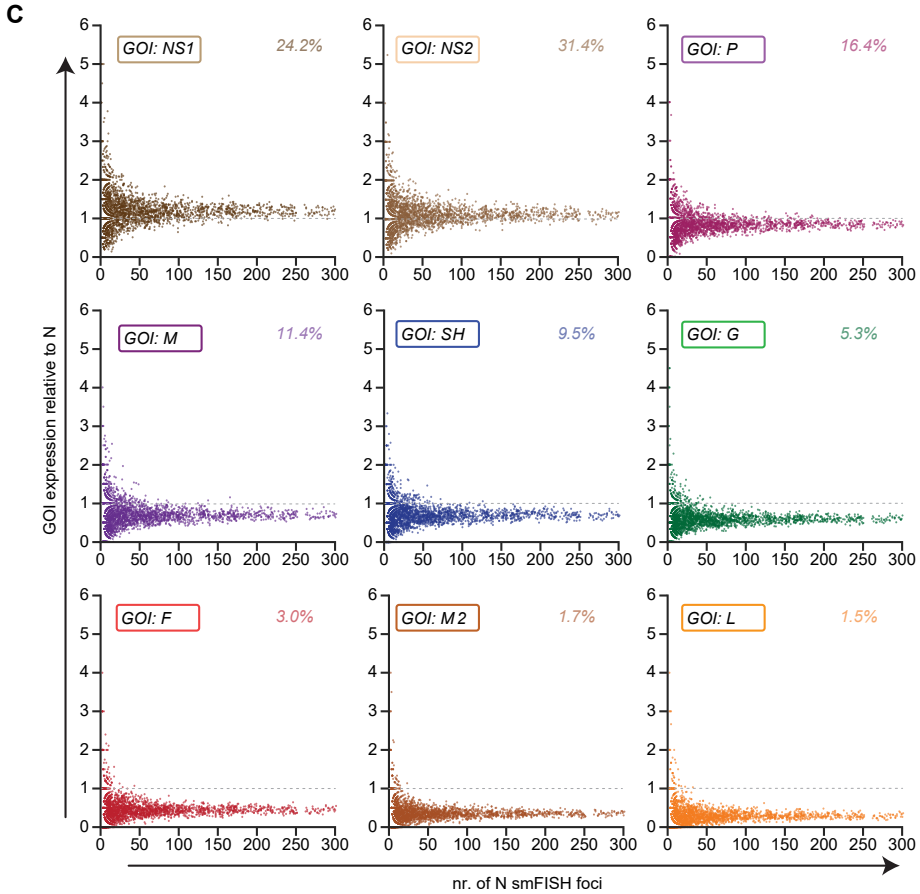
for NS1 and NS2 or >1 for all other GOIs). *(B) Corrected μ of each gene transition that were used in simulations of the skipping or backtracking models, so that indicated probability and corrected μ resulted in mean relative GOI abundance in accordance with experimentally obtained data. The number of simulations per experiment are listed in Table S1.*
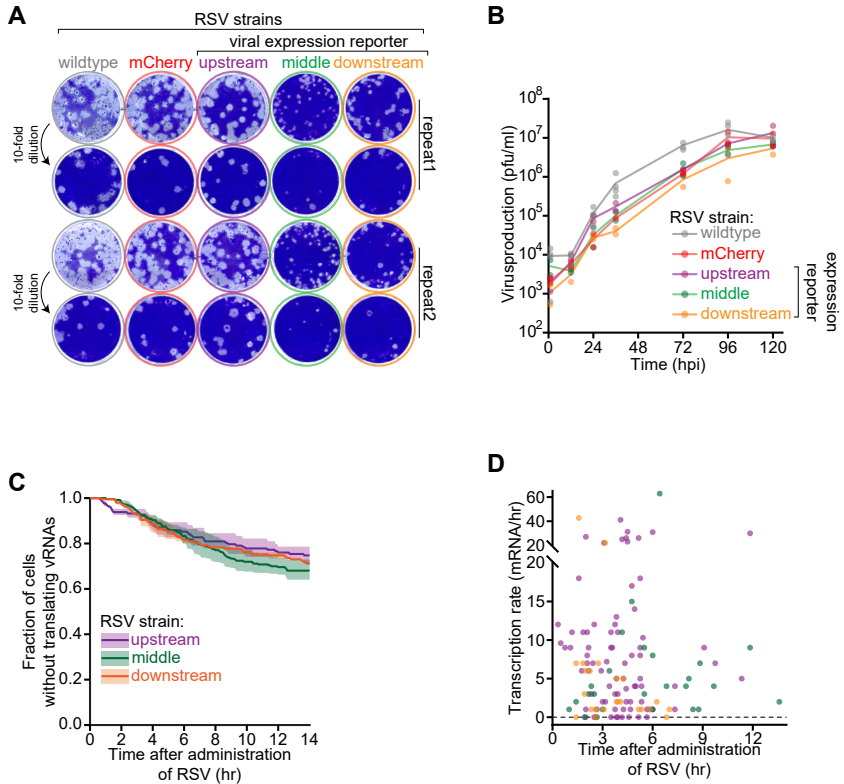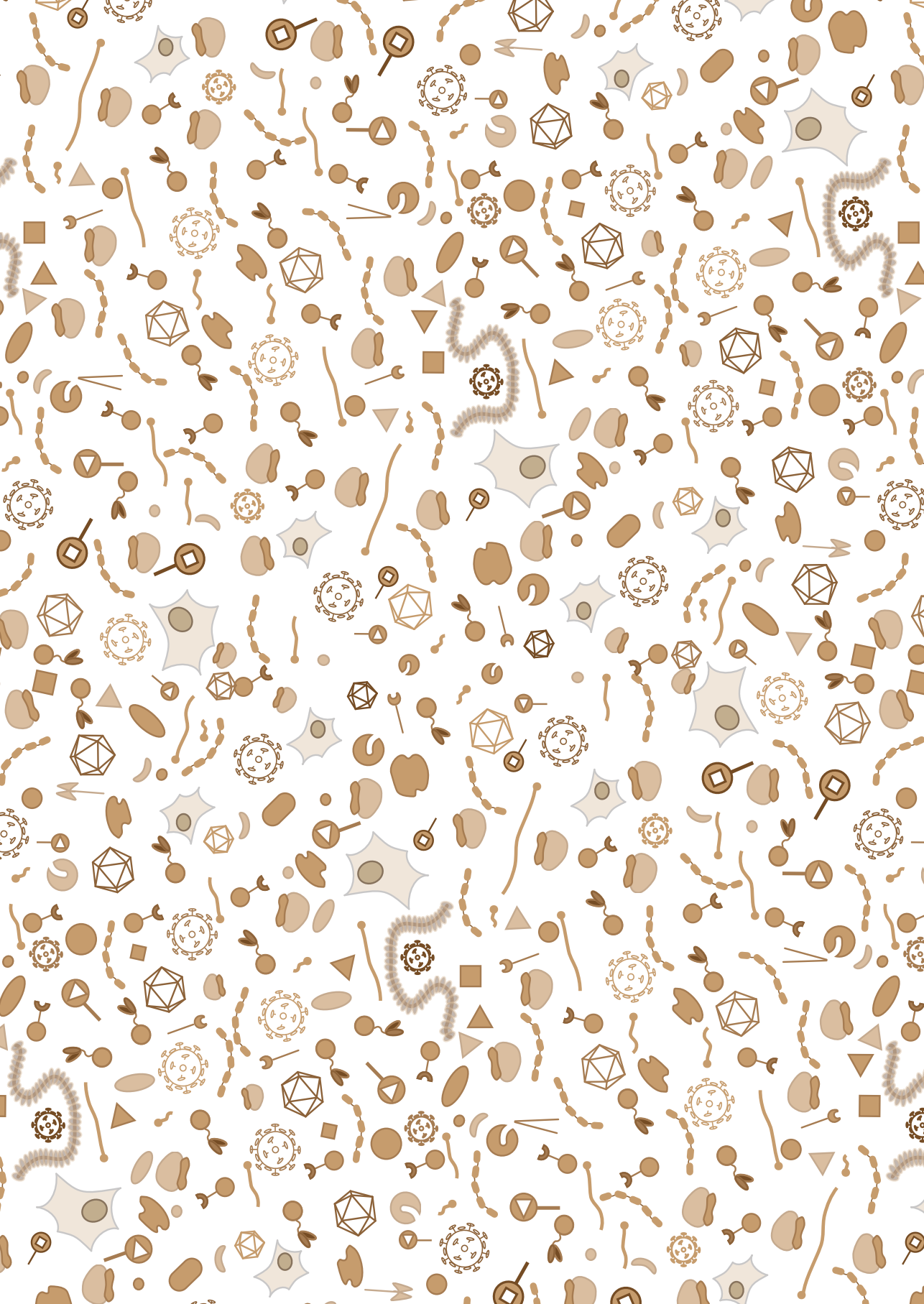
Supplemental figure 5 - related to figure 4 and 6. Live-cell expression analysis using engineered RSV strains.

(A, B) Viral growth analysis based on plaque assays to determine fitness of engineered RSV strains compared to wildtype strain. (A) Images of plaques 6 days after inoculation. For each RSV strain, 2 repeats and per repeat 2 dilutions are shown. (B) Each dot indicates an experimental repeat and lines connect the mean at each time-point. (C) Kaplan-Meier curve indicating fraction of viral translation site-positive cells since RSV administration to cells. Lines and shaded areas indicate mean and SD respectively. (D) Scatterplot of relation between first detection of a translating viral mRNA and transcription rate based on first hour of live-cell trace. Each dot represents a single cell; color refers to RSV strain, as indicated in panel A-C. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.
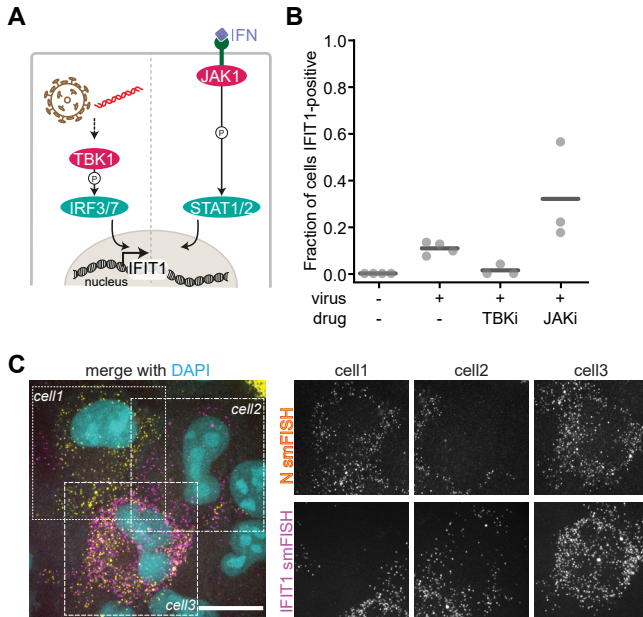
Supplemental figure 6 - related to figure 7. smFISH analysis of RSV gene expression and antiviral activation.

(A) Cartoon illustrating two main pathways that can result in IFIT1 expression. (B) Fraction of A549 cells with > 10 IFIT1 transcripts 6 hr after administration of RSV. Drug treatments: TBKi 1µM MRT67307; JAKi 1µM Tofacitinib. (C) Example image of A549 cells infected with RSV (16hpi) to illustrate cells with high N & low IFIT1 expression (cell1), low N & high IFIT1 expression (cell2) and high N & high IFIT expression (cell3). Scale bare, 15 µm. The number of experimental repeats and cells analyzed per experiment are listed in Table S1.

Chapter 6

Discussion

SUMMARY

In this thesis, live-cell single-molecule imaging assays to study dynamic processes have been developed to address some of the questions raised in **chapter 1**. In **chapter 2**, I described the development of the MoonTag and MashTag single-molecule translation imaging assays to analyze canonical and non-canonical translation on single mRNAs. Using these assays, we revealed extensive heterogeneity in translation start site selection between individual mRNAs from the same gene and over time on single mRNAs. In **chapter 3**, I described how we used live-cell, single-molecule translation imaging to investigate the contribution of eIFs to the dynamics of translation initiation. I then switched topics and described the development of single-molecule imaging assays to investigate virus infections. In **chapter 4**, we established VIRIM to study translation and replication of +ssRNA viruses and identified the bottleneck in the picornavirus life cycle that is targeted by the antiviral response. In **chapter 5**, we studied gene expression of an -ssRNA virus and identify a relation between early viral gene expression variations and the launching of an antiviral response. In **this chapter**, I discuss the advantages and limitations of the new assays. I discuss the implications of the finding and suggest future applications of the assays.

DEVELOPMENT OF NEW ASSAYS
Live-cell imaging

One of the main improvements of these assays in comparison to alternative assays relates to the live nature of the MoonTag & MashTag and VIRIM & VIRIM$^{2.0}$ assays. Many alternative assays to measure mRNA translation or viral replication and gene expression require fixation or lysis of a sample and can therefore only generate read-out of a single time-point. For instance, mRNA translation can be studied by ribosomal profiling: a powerful technique to assess the ribosomal occupancy of mRNAs to determine the translation efficiency and to uncover non-canonical translation (Ingolia et al., 2009, 2011). Gene expression, including viral gene expression, can be studied using single-cell RNA sequencing assays to measure the abundance of mRNAs of each gene (Hashimshony et al., 2016; Tang et al., 2009). However, these assays require destruction of the sample by lysis, and it is therefore complicated to infer dynamics, as only a single time-point at the time can be studied.

It is possible to obtain insights into the dynamics of a process by performing ribosomal profiling or RNA-sequencing on samples obtained at different time-points or by performing computation processing to infer pseudo-timing (Trapnell et al., 2014), although the level of insight may be limited. First, the timescale of obtaining sample may not correspond with the timescale of the dynamics that one wants to study. If samples are, for instance, obtained at a one-hour interval, dynamics at the minute scale cannot be studied. Indeed, variation in translation start site selection on single mRNAs occurs within minutes, as revealed using the MashTag assay (Chapter 2). Second, dynamics can only be studied in samples representing different time-points, if the dynamic process is synchronized (i.e. occurring at the same rate and initiated from the same moment). For example, progression of an infection is

often measured by analyzing samples from different time-point after inoculation of a virus, assuming that the progression of an infection is similar in each sample. To ensure that the moment of infection is similar for each sample, those experiments are performed with high viral titers, complicating analysis of single infections. Moreover, such an approach of collecting samples at different time-point after inoculation to measure progress of an infection only works if all infections progress at a similar pace. If the progress of infection is variable between cells, the samples obtained at different time-point after inoculation contain cells in different stages of an infection. Using VIRIM, we indeed discovered, that early infection dynamics are highly variable between single cells (Chapter 4). In some cells, it takes multiple hours to complete translation and replication of the first incoming genome, whereas the infection progresses much faster in other cells. Both types of cells are present at any time-point after inoculation. Third, it can be challenging to couple observations from different time-points, especially if there are multiple different phenotypes. For instance, if cells behave heterogeneously, it is complex to link a phenotype of one cell at one type-point to the phenotype of another cell at another type-point. However, live tracking of a single cell over multiple time-points enables comparison and coupling of different observations. For example, using VIRIM we observed that there were multiple replication attempts in some infected cells (referred to as 'pulse-phenotype'). By using live-cell imaging, we could follow these cells and reveal that the success-rate of infection was strongly reduced in cells with multiple pulses, exemplifying the power of live-cell imaging.

In sum, we can now follow the same processes over time in the same cell/on the same molecule and therefore, we can couple phenotypes at different time-points to obtain detailed insights into the dynamics of cellular and viral gene expression.

Single-cell and single-molecule imaging

In addition to the live nature, the single-molecule sensitivity is another main advantage of the MoonTag & MashTag and VIRIM & VIRIM[2.0] assays in comparison to other assays. With the new assays, we can follow mRNA translation on single-molecules (MoonTag & MashTag) and we can follow virus infections in single cells by quantifying the number of viral replication or transcription products with single-molecule sensitivity (VIRIM & VIRIM[2.0]). Many alternative assays (such as ribosomal profiling, Western blotting, qPCRs, etc.) depend on combining many cells at once to study mRNA translation or virus infections. By analyzing many cells and thereby many single molecules at once, rare phenotypes may be masked by the dominant phenotype. For example, using VIRIM we identified that infections are unsuccessful in a small fraction of cells. Similarly, using the MashTag assay we uncovered that canonical and non-canonical start site selection sometimes alternate on single mRNAs. If many cells or molecules would be analyzed at once, such rare phenotypes are likely undetected.

Moreover, many alternative assays do not have single-molecule sensitivity. For example, cellular gene expression is often studied by using a fluorescent reporter that reflects protein production. Similarly, virus infection dynamics can be studied using a virus strain that expresses a fluorescent protein and then use the fluorescence intensity of individual cells to examine the progression of a virus infection (Guo et al., 2017; Singer et al., 2021). However, there is a lag time between

transcription and/or translation of a reporter gene and accumulation and maturation of the fluorescent protein (Balleza et al., 2018). Consequently, the fluorescent protein reporting on cellular gene expression or virus infection is no longer located on the same position as the mRNA encoding the protein. Moreover, it may be challenging to observe and trace a single fluorescent protein in a cell. In contrast, the Sun-, Moon -, and MashTag assays enable direct observation of a translation product on single mRNAs by recruiting multiple copies of a stably expressed and fluorescently labeled intrabody or nanobody resulting in readily detectable fluorescent foci. As a result, there is no lag time between translation of an mRNA and detection of the translation product and mRNA translation can be studied on single mRNAs. Moreover, direct analysis of mRNA translation by visualizing the product co-translationally prevents complications due to variable protein stability to the read-out. Especially non-canonical protein products may be subjected to rapid protein degradation (Ruiz Cuevas et al., 2021). Using a fusion between a non-canonical translation product and a fluorescent protein may therefore not accurately reflect the frequency of non-canonical protein synthesis. Similarly, the lag time between viral replication and detection of the replication products or the time between viral transcription and detection of the transcription products is reduced in the VIRIM and VIRIM[2.0] assays respectively. Therefore, VIRIM and VIRIM[2.0] enable direct quantification of the number of viral replication or transcription products readily after their production and rapidly from the start on infection onwards.

Future technical developments

Although the MoonTag & MashTag and VIRIM & VIRIM[2.0] assays allow analysis of viral and cellular gene expression with unprecedented sensitivity, the assays have several limitations and could be further improved by some the adaptations listed here.

MoonTag & MashTag - limitations and proposed improvements

Single-molecule analysis of translation on single mRNAs based on Sun-, Moon -, or MashTag requires introduction of a relatively large peptide array at the N-terminus of the protein of interest. While designing an experiment, it is important to consider the effect of this large insert. For example, the function of the protein of interest is likely disturbed by the large tag at the N terminus. Loss of protein function may be mitigated by introducing additional elements between the peptide array and the protein of interest to separate the peptide array and protein. For example, 2A self-cleaving peptides could be introduced to separate the peptide array from the protein of interest (Kim et al., 2011). However, co-translational separation of the peptide array affects the sensitivity of the single-molecule translation imaging assay as the length of sequence encoding the protein of interest determines how long a translating ribosome can be detected (Chapter 2). Alternatively, post-translational separation of the peptide array could be achieved using inducible cleavage systems such as the Tobacco etch virus (TEV) protease (Nam et al., 2020).

Post-translational separation and decay of the peptide array from the protein of interest may also be beneficial to solve another limitation of the Sun-, Moon -, or MashTag assays, related to the sequestration of antibodies or nanobodies by mature

proteins of interest (i.e. proteins released from the ribosomes upon synthesis). The single-molecule translation imaging assays depend on the readily available presence of antibodies or nanobodies. However, if many peptide arrays have been formed in a cell, the number of available antibodies of nanobodies may be limited. For example, if the protein of interest is highly expressed, proteins of interest fused to peptide arrays accumulate rapidly in a cell and each of the peptide arrays will sequester antibodies of nanobodies. As a result, the pool of available antibodies or nanobodies gets depleted leading to reduced accumulation of the antibody or nanobody on translating mRNAs. This so called antibody depletion heavily depends on the expression of the protein of interest (Khuperkar et al., 2020). In a typically Sun-, Moon -, or MashTag experiments the antibody depletion impairs analysis > 2 hours after the start of the experiment. Post-translational degradation of the peptide array may help to circumvent antibody depletion, as the already produced proteins of interest will no longer sequester antibodies or nanobodies. Inducible degradation of the protein of interest or using a short-lived protein of interest can be used in single-molecule translation imaging assays (Wang et al., 2016; Wu et al., 2016).

Furthermore, we introduce PP7 stemloop structures in the 3'UTR of the reporter mRNA to facilitate labeling and tethering of mRNAs. Although these structures are not in the coding sequence, introduction of the exogenous sequences may interfere with regulatory elements in the 3'UTR. Additionally, immobilizing mRNAs to the plasma membrane is not advised for all mRNAs. Some mRNAs reside in specific subcellular locations and their localization depends on translation, implying a contribution of the subcellular compartment to their translation (Chouaib et al., 2020). To study the role of the subcellular environment in mRNA translation of these mRNAs, it is recommended to bypass the PP7-mediated tethering of mRNAs to the plasma membrane. Indeed, the single-molecule translation imaging assays can also be performed without tethering the mRNAs (Morisaki et al., 2016; Pichon et al., 2016; Wang et al., 2016; Wu et al., 2016).

The MashTag system is based on combining the SunTag and MoonTag arrays to enable analysis of two ORFs on a single mRNA. However, one may want to analyze translation of more than two ORFs. For example, simultaneous analysis of all three ORFs may be very informative. Currently, the MashTag does not directly report on translation of the third ORF. Instead, the number of ribosomes translating the third ORF are inferred based on estimations on the total number of ribosomes translating either of the three ORFs and the sum of ribosomes translating the main (MoonTag) or an alternative (SunTag) ORFs. Introducing a third, orthogonal peptide-antibody pair might enable direct visualization of translation of the third ORF. To further complicate single-molecule mRNA translation analysis, one may want to observe translation of more than three ORFs. For example, comparison of uORF translation and translation of the three ORFs in the coding sequence may be of interest. For such complex studies, four or more orthogonal peptide-antibody pairs would be required. Fortunately, several promising candidates have already been developed, like the Frankenbody, the Spaghetti monster, and the AlfaTag (Götzke et al., 2019; Viswanathan et al., 2015; Zhao et al., 2018). Moreover, little cross-reactivity was observed between these tags (Alessandra-Vigano et al., 2021) and the next generation MashTags could therefore be developed by combining these tags with the SunTag and/or MoonTag.

Summarizing, there are various possible improvements to the Sun-, Moon -,

I

II

III

IV

V

VI

&

or MashTag assays. Depending on the exact research question, some of these adaptations could be considered to expand the range of applications of the single-molecule translation imaging assays.

VIRIM - limitations and proposed improvements

VIRIM is based on the SunTag assay and could therefore benefit from some of the improvements of the single-molecule translation imaging assays (see section 'MoonTag & MashTag - limitations and proposed improvements').

VIRIM is based on introducing a SunTag peptide array at the N terminus of the viral genome so that translation of single vRNAs can be observed and the number of vRNAs can be quantified using the SunTag assay. We speculate that VIRIM can be used to study many different types of +RNA viruses by introducing SunTag peptides into the viral coding sequence. Similar to designing a single-molecule translation imaging assay, the effect of fusing a peptide array to a viral protein should be considered while designing a VIRIM experiment. The fusion of the peptide array likely disrupts the function of the viral protein, and it is therefore recommended to check whether the insert affects viral growth. Additionally, steps could be taken to induce separation of the peptide array from the viral protein. For example, we introduced a viral 3C protease motive between the 5x SunTag peptide array and the CVB3 polypeptide to ensure removal of the peptide array from the polypeptide (Chapter 4). Alternatively, the ORF expressing the peptide array and the ORF expressing viral proteins could be separated by introducing an IRES downstream of the ORF expressing the peptide array to drive translation of the ORF expressing viral proteins. In such a set-up, the peptide array would not interfere with viral proteins, but VIRIM could no longer be used to examine viral translation dynamics, as the translation of the peptide array would no longer reflect the synthesis of viral proteins.

Another potential complication relates to the site in the viral coding sequence where the SunTag peptides are introduced. The genome of many +RNA viruses, like picornaviruses or flaviviruses, encodes a single polyprotein (Baggen et al., 2018; Barrows et al., 2018). Inserting the SunTag peptide array the N terminus of the viral polyprotein thus enables visualization of translating vRNAs. However, synthesis of viral proteins of many other +RNA viruses, such as alphaviruses or coronaviruses, depends on translation of multiple ORFs (Singer et al., 2021; de Wit et al., 2016). Introducing SunTag peptides in one ORF may not be sufficient to detect all translating vRNAs; vRNAs can only be observed if the ORF encoding SunTag peptides is translated, but not if another ORF is translated. Selection of the ORF to insert the SunTag peptides impacts the interpretation of the results. If early events during infection are of interest, it is recommended to insert the peptide array into the ORF that is translated first after infection. Alternatively, different types of peptide arrays could be introduced into each ORFs. For example, SunTag and MoonTag peptides could be used to enable visualization of translation of two different ORFs of a viral genome. In this way, translating vRNAs can be observed regardless of the ORF that is translated. Moreover, simultaneous analysis of translation of multiple viral ORFs may yield insights into the dynamics and coordination of translation of the ORFs.

Analysis of viral replication using VIRIM is based on quantifying the number of translating vRNAs over time. However, the process of replication itself is not

visualized. Instead, the rate of replication is inferred based on an increase in translating vRNAs. Consequently, untranslating vRNAs are not included in the number of vRNAs. We use the first detection of a translating vRNA as the start of an infection. However, prior to observing the first translating vRNA, a virion needs to bind to a host cell and release its genome into the host cell. These steps cannot be visualized using VIRIM and are therefore not taken into account in the current analyses. To incorporate entry steps into the analysis of early virus infection, RNA labeling strategies could be used. For example, metabolic labeling of the viral genomes has been used to stud entry and release of the viral genome (Brandenburg et al., 2007). Combined genome labeling and VIRIM may yield more insights into the entry of a virus into the host cell and the first initiation of translation.

Moreover, combined analysis of translating vRNAs and the number of vRNAs by smFISH revealed that 40% of the CVB3 vRNAs did not undergo translation (Chapter 4). These untranslating vRNAs may be defective in translation and therefore never get translated or vRNAs may cycle between translation 'on' and translation 'off' state. During the translation-to-replication-switch, translation is turned-off and the vRNA is used as a template for replication. If replication is unsuccessful, translation on the vRNA may be reinitiation, illustrating that vRNAs can cycle between translation 'on' and translation 'off' states. To examine the dynamics of translational cycling or the abundance of translation-defective vRNAs, direct observation of the vRNAs in addition to observing the translation state would be required. To visualize vRNAs directly, RNA labeling strategies could be employed, like PP7 stemloops or RhoBAST aptamers (Braselmann et al., 2020; Sunbul et al., 2021). Notably, introduction of RNA structures into a vRNA may interfere with regulatory elements in the vRNA. Therefore, additional controls are required to exclude any interference from the elements introduced to visualize vRNAs.

By combing VIRIM with other live-cell biosensors, we studied the dynamics of an infection and dynamics of virus-induced loss of nuclear pore integrity and cleavage of eIF4G. It may be of interest to expand the collection of live-cell biosensors to be combined with VIRIM. For instance, many host cell proteins are degraded by viral proteases, although the timing and efficiency of degradation differs between proteins (Saeed et al., 2020). New biosensors, based on the eIF4G cleavage sensor, could be developed to examine the difference in timing and dynamics of protein degradation during early infections. Additionally, the cell cycle state of the host cell has been implicated in the efficiency of CVB3 replication (Feuer et al., 2002), although the exact mechanism by which the cell cycle affects CVB3 replication remains unclear. The fluorescent ubiquitination-based cell cycle indicator (FUCCI) biosensor system would be an ideal candidate to be combined with VIRIM to study the interplay between the host cell's cell cycle and virus infection (Sakaue-Sawano et al., 2008)

In sum, there are several potential improvements to further expand the application options of VIRIM by including more parts of the viral life cycle in the analysis or by expanding the collection of live-cell biosensors reporting on host cell processes.

VIRIM$^{2.0}$ - limitations and proposed improvements

Inspired by VIRIM, VIRIM$^{2.0}$ is a live-cell single-molecule imaging assay to investigate

gene expression dynamics of -ssRNA virus RSV by visualizing translation of viral mRNAs. Like VIRIM, VIRIM[2.0] is based on SunTag translation imaging and some of the potential improvements explained in section 'MoonTag & MashTag - limitations and proposed improvements' are relevant for VIRIM[2.0] as well.

Translation-independent detection of viral transcripts could be an improvement of VIRIM[2.0]. Quantification of viral transcription with VIRIM[2.0] requires translation of the viral reporter mRNAs. However, a small fraction of viral reporter mRNAs is untranslating and therefore not included in the quantification. Visualization of RNAs using aptamers or PP7 stemloops would enable inclusion of such untranslating viral mRNAs in the quantification of the number of viral transcripts. Translation-independent visualization of RNAs is used to study cellular transcription (Janicki et al., 2004; Larson et al., 2011; Rodriguez and Larson, 2020), suggesting that translation-independent visualization of mRNAs can be used to study viral transcription as well.

Using VIRIM[2.0], we studied temporal dynamics of RSV transcription in single cells. However, VIRIM[2.0] only reports on transcription of a single viral gene. Based on smFISH analyses of RSV gene expression, there is coordination of expression between viral genes. To examine how RSV gene expression is coordinated over time, a multi-color live-cell single-molecule viral transcription assay is required. For instance, a new RSV strain could be generated that expresses multiple viral gene expression reporters to simultaneously measure viral gene expression of multiple viral genes. Alternatively, VIRIM[2.0] could be combined with smFISH-based analysis of viral mRNA. In particular, multiplexed or sequential smFISH may provide RSV transcriptome-wide viral gene expression insight (Chen et al., 2015; Eng et al., 2019). By combining VIRIM[2.0] with such transcriptome-wide measurements, both temporal dynamics of viral gene expression and expression of multiple viral genes could be analyzed.

In contrast to VIRIM, VIRIM[2.0] cannot be used to analyze viral replication, as VIRIM[2.0] does not enable visualization of replication products. It is currently unclear whether and how viral replication dynamics contribute to the dynamics of viral gene expression. For example, we uncovered that viral transcription rates vary between single cells. Differences in the number of viral genomes per cell likely contribute to this variation; if the viral replication occurs rapidly in a cell, many genomes may be formed that can all serve as a template for viral transcription. Additionally, it is currently unclear whether the concentration and expression of the RSV polymerase affects viral gene expression. The polymerase is required for both transcription and replication and differences in polymerase levels between cells could therefore contribute to heterogeneity in viral transcription and replication. To study the interplay between viral transcription, replication, and concentration of the polymerase, new tools would be required. For example, detection of viral genomes by labeling the genome or single-molecule imaging of the polymerase in combination with VIRIM[2.0] might provide a more insights into the early RSV infection dynamics.

BIOLOGICAL IMPLICATIONS AND FUTURE DIRECTIONS

HETEROGENEITY OF MRNA TRANSLATION

Using the Sun-, Moon-, and MashTag reporters, we uncovered that a substantial fraction of ribosomes translates mRNAs non-canonically: they select non-canonical translation initiation sites, translate multiple ORFs consecutively on the same mRNAs, or undergo stop codon readthrough (Chapter 2). Consequently, the proteome is more diverse than would be predicted based on annotations of coding sequences and new questions arise related to the causes and consequences of non-canonical translation.

Potential causes

As with anything in life, mistakes can occur. Variations in mRNA translation may be the result of random errors. However, non-canonical translation does not occur randomly. Instead, a subset of mRNAs is more susceptible to non-canonical translation compared to others. For instance, alternative start site selection varies a lot between single mRNAs originating from the same gene; on some mRNAs there is only canonical start site selection, whereas non-canonical translation initiation is the dominant form of translation on some other mRNAs (Chapter 2). Similarly, frameshifting on a viral frameshifting sequence is variable between mRNAs; a small fraction of mRNAs are responsible for the synthesis of the majority of proteins from the frameshifted coding sequence (Lyon et al., 2019). Moreover, the frequency of these non-canonical manners of translation can even change over time on single mRNAs. For example, simultaneous analysis of canonical and non-canonical start site selection on single mRNAs revealed that the rate of canonical and non-canonical translation initiation is constant on some mRNAs, whereas canonical and non-canonical translation initiation rates anticorrelate on some other mRNAs (Chapter 2). These examples illustrate that non-canonical translation cannot just be attributed to random variations during translation. Instead, unknown mechanisms may drive variation in translation between single mRNAs and on single mRNA over time.

The mRNA-to-mRNA heterogeneity may result from variations during mRNA synthesis, like selection of an alternative transcription start site, incorporation of an alternative nucleotide during transcription, or alternative splicing. As a result, mRNAs with different sequences can be formed from the same gene. If such an alternative sequence involves the Kozak sequence of an initiation site, it can affects the likelihood of translation initiation (Noderer et al., 2014). Notably, such variations to the sequence of mRNAs are static and would not change after synthesis of the mRNAs. Variations in the mRNA sequence can therefore not directly account for temporal changes of translation on single mRNAs.

The stability of a secondary structure near a translation initiation site affects the translation efficiency (Gu et al., 2010; Robbins-Pianka et al., 2010), suggesting that secondary RNA structures with variable stability may contribute to the temporal heterogeneity in mRNA translation. For instance, if a structure in a 5'UTR is formed, it may reduce the translation initiation efficiency of the canonical initiation site, which could result in increased initiation from another start site. Upon dissolving

the structure, however, the canonical initiation site can become dominant again. RNA structural dynamics may be controlled by regulators, such as helicases or translating ribosomes (Guenther et al., 2018; Ruijtenberg et al., 2020). Therefore, the formation of a structure may differ between mRNAs with the same sequence. A flexible secondary structure may thus explain, why we observe changes in the main translation initiation site over time on some mRNAs.

Additionally, heterogeneity in the recruitment of ribosomes to mRNAs may contribute to variations in mRNA translation over time. Recruitment of a ribosome to an mRNA depends on cap binding protein eIF4E and its interacting partner eIF4G. A single cap-eIF4E interaction can facilitate the recruitment of multiple consecutive ribosomes, resulting in a translation 'on' state of the mRNA. Upon dissociation of eIF4E, rebinding of an eIF4E is required to maintain the 'on' state. Instead, the mRNA goes into a translation 'off' state, if no new ribosomes can be recruited, for example if eIF4E availability is low (Chapter 3). Consequently, mRNAs from the same gene can be in a translation 'on' or 'off' state simultaneously, depending on whether each mRNA is bound by an eIF4E and therefore recruiting ribosomes. Binding and unbinding of eIF4E may therefore contribute to variations in translation over time on an mRNA. Such temporal fluctuations due to dynamics of ribosome recruitment are predicted to affect both canonical and non-canonical translation, as both types of translation require ribosomes. Indeed, canonical and non-canonical translation correlate over time on 56% of the mRNAs (Chapter 2), suggesting that dynamics in ribosome recruitment explains temporal variations in translation on the majority of mRNAs.

Moreover, ribosome composition is heterogenous and translation efficiency is influenced by ribosome composition. Different ribosome compositions can exist simultaneously in the same cell, causing gene-specific effects on translation efficiency (Shi and Barna, 2015). For example, specialized ribosomes with large 60S subunit RPL10A/uL1 preferentially translate mRNAs with a corresponding IRES-like element in the 5'UTR (Shi et al., 2017). Variations in the sequence of an mRNA and flexible secondary structures may thus influence the type of ribosomes that translate an mRNA.

Collectively, several different factors may account for the heterogeneity in mRNA translation between single mRNAs and over time on single mRNAs. Future experiments are required to understand the relative contribution of each factor to the dynamics and efficiency of mRNA translation.

Potential consequences

Extensive non-canonical translation has implications for the proteome and thereby cellular function. Non-canonical translation directly causes diversification of the proteome, resulting in a wider repertoire of polypeptides in a cell than previously appreciated. Non-canonical translation products include instable or dysfunctional polypeptides that can cause stress to a cell. For instance, N-terminally extended or truncated protein isoforms that may be dysfunctional can be formed during non-canonical translation (Hann et al., 1994).

There is an increasing understanding that many non-canonical translation events result in functional polypeptides, such as micropeptides involved in all kinds of

processes, like cytoskeleton dynamics or mitochondrial function (Chen et al., 2020; Couso and Patraquim, 2017). In some cases, a non-canonical translation product can even have a direct effect on the canonical translation product. For example, one of the uORFs of the mRNA encoding protein kinase C (PKC) encodes a small peptide that binds and inhibits PKC (Jayaram et al., 2021). The level of PKC activity in a cell therefore depends on the balance between translation of the inhibitor-encoding uORF and the PKC-encoding mainORF, suggesting that canonical and non-canonical translation must be well balanced.

Non-canonical translation also affects the proteome indirectly by reducing the production of canonical proteins; if many ribosomes utilize non-canonical start sites, less ribosomes use the canonical start site. The presence of a uORF is generally associated with a reduction in production of the proteins encoded by the main ORFs (Chew et al., 2016), exemplifying that non-canonical translation can affect the production rate of canonical proteins.

Moreover, diversification of a cell's proteome has implications for the peptides presented in major histocompatibility complexes (MHCs) that can be recognized by T cells, leading to an immune response. During T cell development, T cell receptors (TCR) that interact with 'self' peptides are excluded to prevent the formation of T cells that may induce an immune response upon binding to an MHC loaded with a self-peptide (Pontarotti et al., 2020). Widespread non-canonical translation results in a broad range of peptides that deviate from canonical peptides and non-canonical translation products are particularly efficient at generating peptides that are loaded in MHCs (Ruiz Cuevas et al., 2021). To avoid activation of a T cell response due to non-canonical translation products, TCRs binding to non-canonical peptides must be excluded during T cell development. It is therefore likely that there is extensive non-canonical translation during T cell development.

Additionally, frequent presentation of non-canonical peptides in TCRs is of relevance for the development of tumor-immunotherapies based on tumor-specific antigens (i.e. neoantigens). Due to the high indel frequency among tumors, tumors contain many frameshifting mutations resulting in neoantigens. Such neoantigens are promising candidates for the development of cancer vaccines (Roudko et al., 2020). However, our analyses with the MashTag reporters revealed that non-canonical translation occurs frequently. There may be overlap between neoantigens and non-canonical translation products. Selection of neoantigens for the development of cancer vaccines should therefore include analysis of non-canonical translation to exclude neoantigens that could also be the product of non-canonical translation.

Widespread non-canonical translation initiation also has implications for experimental designs, specifically engineering of knockouts using CRISPR/Cas9 by introducing a frameshifting mutation downstream of the canonical start site. We observed that leaky start site selection occurs frequently on various mRNAs and that a non-canonical start site downstream of the canonical start site can be used for translation initiation (Chapter 2). Substantial expression of an N-terminally truncated protein isoform is therefore expected from a gene that was knocked-out using the strategy of a frameshifting mutation. Indeed, frequent expression of knockout genes can be observed (Tuladhar et al., 2019), illustrating that it is essential to consider non-canonical translation while designing knockout strategies.

In sum, extensive heterogeneity in mRNA translation can influence the composition

I

II

III

IV

V

VI

&

of a cell's proteome and regulation of canonical and non-canonical translation impacts the diversity of the proteome.

DYNAMICS AND HETEROGENEITY DURING VIRUS INFECTIONS

Using VIRIM and VIRIM$^{2.0}$, we shed light on early virus infection dynamics and the virus-host competition.

Early virus infection dynamics

For all the viruses that were analyzed in this thesis, we found widespread variation in the dynamics of early infection between single cells (Chapter 4 and 5). For instance, 1) the duration of phase1 of picornavirus infections (i.e. translation of the first vRNA) varies from 15 minutes to several hours, 2) picornavirus infection leads to successful replication in approximately 80% of cells, but fails in the remaining cells, 3) upon activating the antiviral response pathways, the infection is successfully inhibited in the majority of cells, but unaffected in approximately 35% of the cells, 4) the transcription rate of RSV genes differs from a couple of mRNA per hour to hundreds of mRNAs per hour, and 5) upon RSV infection, there is potent activation of an antiviral response in only a fraction cells. The cause of each of these single cell heterogeneities is currently unknown and requires more investigations. Follow-up studies could focus on whether the variations between infections in single cells are stochastic or caused by differences between the cells or the presence of different virus variants. Even in a homogenous population of cells, only a small fraction of cells can launch an antiviral response, suggesting that cell-to-cell variations can contribute to single-cell heterogeneity in the virus infection dynamics (Doğanay et al., 2017; Patil et al., 2015). The presence of virus variants that lack fragments of the viral genome can also account for heterogeneity in the timing of viral replication and the virus's ability to suppress activation of an antiviral response (Sun et al., 2015; Vignuzzi and López, 2019). Detailed understanding of cellular and viral variations is required to understand the origins of heterogeneity in virus infections in single cells.

The genome of +ssRNA viruses is the template for both viral protein synthesis and for viral replication and both processes are essential for a successful infection and must therefore be coordinated. Using VIRIM, we revealed that the translation and replication of the incoming RNA molecule during a picornavirus infection are particularly heterogeneous in timing and success-rate (Chapter 3). We hypothesize that an unknown mechanism controls the translation-to-replication switch and thereby regulates the timing and success-rate of viral translation and replication. The so-called timer is insensitive to the concentration of viral proteins, as reducing the viral protein synthesis rate did not influence the translation-to-replication switch. Moreover, there is no correlation between the timing of the translation-to-replication switch and the success-rate of an infection, suggesting that the timer acts independently of the viral polymerase. More experiments are required to identify the mechanism that controls the translation-to-replication switch.

For instance, the role of host cell proteins like RNA-binding protein poly(rC) binding protein 2 (PCBP2) could be examined. Virus-induced cleavage of PCBP2 has been implicated in the translation-to-replication switch (Chase et al., 2014; Dave et

al., 2019). However, the role of PCBP2 in the translation-to-replication switch is inconclusive. Firstly, cleaved PCPB2 is only detected from 4 hours after infection onwards, whereas the translation-to-replication switch occurs earlier in infection in most cells. Secondly, PCBP2 cleavage is irreversible, whereas translation and replication can take place simultaneously and consecutively in the same cells. If cleaved PCBP2 would inhibit viral translation and stimulate viral replication, viral translation would be continuously suppressed after the translation-to-replication switch. Therefore, the translation-to-replication timer cannot depend on the cleavage of all PCBP2 in a cell. However, we found that virus-induced cleavage of host proteins, like eIF4G and nuclear porins, is initiated rapidly after the start of an infection. Low amounts of PCBP2 bellow Western blot detection thresholds may therefore be present early in infection and may aid in the first translation-to-replication switch. High-resolution analysis of PCBP2 binding and unbinding to viral RNA is required to address whether and how PCBP2 controls the translation-to-replication switch.

In contrast to +ssRNA viruses, the genome of the -ssRNA virus RSV cannot be translated directly into viral proteins but requires the transcription of viral genes into viral mRNAs. To ensure proper balance between viral proteins, the expression of viral genes needs to be coordinated. Based on smFISH analysis, the expression of RSV genes takes place in a gradient along the viral genome: upstream genes are expressed at a high rate compared to downstream genes (Chapter 5). Moreover, we identify that the viral polymerase can scan the genome bidirectionally to identify the gene starts for transcription initiation. Using VIRIM[2.0], we observed that viral expression fluctuates over time and seems to occur in a burst-like pattern. Many factors can control cellular transcriptional bursting, including cis-acting elements and the composition of the transcription start site (Larsson et al., 2019). However, there are no known cis-acting elements controlling viral gene expression and the gene starts of all viral genes are similar. It is therefore unknown which mechanism underlies the viral transcriptional bursts. Furthermore, it is unclear how the bursting of viral transcription relates to the coordination of expression between viral genes.

Virus-host competition

The outcome of a virus infection depends on a complex interplay between the virus and the host; host cells contain sensors to detect a virus and rapidly activate an antiviral response, whereas viruses have various mechanisms to evade detection by the host cell or to block launching of an antiviral response (Chapter 1). Using VIRIM, we uncovered that especially the replication of the first vRNA during a picornavirus infection is a bottleneck for successful infection and a potent target for multiple parallel antiviral mechanisms (Chapter 4). During replication of the first vRNA, there is only a single viral dsRNA molecule to be recognized by the host cell and only a brief window of approximately 30 minutes. In this short time, the host cell needs to sense this single molecule and ensure recruitment of the antiviral proteins to the dsRNA molecule to inhibit replication or degrade the vRNA. To understand the dynamics of host cell's detection of the infection and rapid suppression of the virus, more experiments are required. To this end, live-cell sensors that report on virus sensing or activation of antiviral response could be employed, such as a 2-5A sensor to monitor detection of dsRNA and activation of RNAse L (Chitrakar et al., 2019).

Particularly the replication of the first vRNA is a potent target of the antiviral response, suggesting that the outcome of an infection is largely determined during the early moments of a picornavirus infection. Similarly, there is an opposing trend between early RSV expression and activation of an antiviral response (Chapter 5). Although it is currently unclear how early and late RSV expression dynamics relate to each other and to activation of an antiviral response, this opposing trend suggests that successful launching of an antiviral response also depends on rapid sensing of an infection early during an infection. Moreover, these results suggest that if a virus manages to evade early detection, the virus is likely to win the virus-host competition. To address whether the virus-host competition is generally settled during early infection, expansion of the current analyses into more viruses is required.

Clinical applications

Using the single-molecule imaging approaches described in this thesis, we can identify the bottleneck for a successful infection and study how and when an innate antiviral response can inhibit an infection. We identified replication of the first vRNA as the bottleneck for a successful picornavirus infection and the prime target of the antiviral response (Chatper 4). This knowledge might be utilized to design antiviral therapies that target the bottleneck or support the antiviral response efficiently. Expanding the analyses into other viruses may reveal potential therapeutic targets for other viruses as well.

Analysis of coordination in expression between RSV genes suggests that the viral polymerase scans the genome bidirectionally to identify transcription start sites. Many antiviral drugs target viral polymerases (Fearns and Deval, 2016) and insight into the mechanism of scanning may help to design drugs to block scanning or transcription initiation and thereby inhibit RSV . Moreover, the polymerase is highly conserved between different viruses in the *Mononegavirales* family and insight into the RSV polymerase might also provide insight the polymerases of other viruses, which can be used to design therapies against other *Mononegavirales* (Cox and Plemper, 2015).

Insight into vulnerabilities in the life cycle of a virus might also be relevant for the design of therapies that take advantage of viruses, like oncolytic viral therapies. Such therapies use viruses that preferably infect tumor cells and have limited effect on non-cancerous cells(Sivanandam et al., 2019). Consequently, tumor cells are killed, and activation of an antiviral state in response to the infection may aid in overcoming immunosuppression by the tumor. There are currently several promising clinical trials with picornaviruses to treat tumors, ranging from bladder cancer to recurrent glioblastomas (Annels et al., 2019; Desjardins et al., 2018). Understanding the bottleneck in the life cycle of the therapeutic virus might facilitate the design of a virus that can overcome the bottleneck. Moreover, insight into the competition between the therapeutic virus and the tumor immune response might lead to improved viral tumor therapies.

CONCLUDING REMARKS

In this thesis, I have described the development of live-cell single-molecule imaging assays to study mRNA translation dynamics and to examine the dynamics of early RNA virus infections and the virus-host interplay. Based on these assays, mRNA translation initiation is heterogenous between single mRNAs from the same gene and over time on single mRNAs. Moreover, there is a lot of cell-to-cell variation in viral gene expression, replication, and the antiviral response during early virus infections. Collectively, these findings illustrate that live-cell single-molecule imaging is a powerful approach to study dynamic processes. The new imaging approaches will be valuable tools to study the causes and consequences of these single-molecule and cell-to-cell differences.
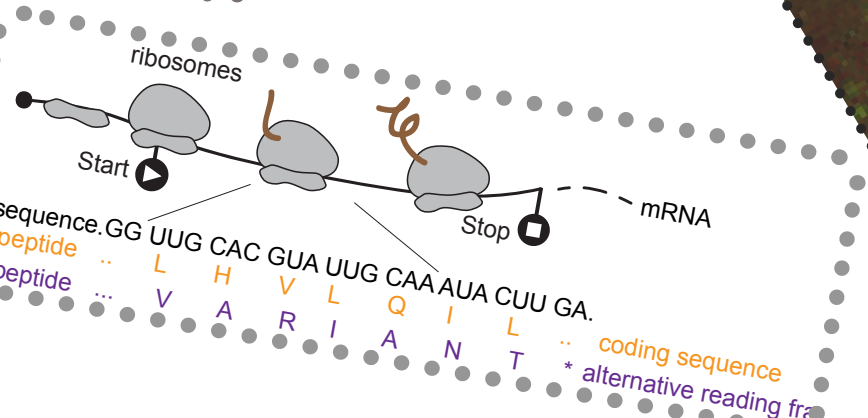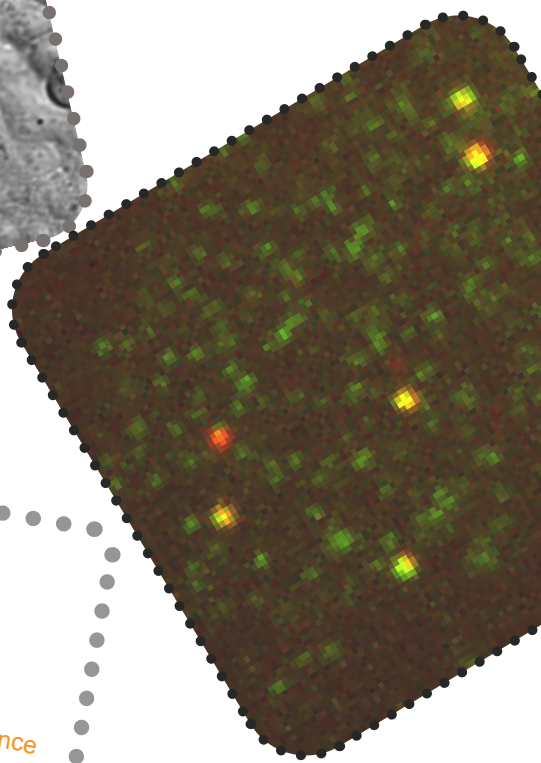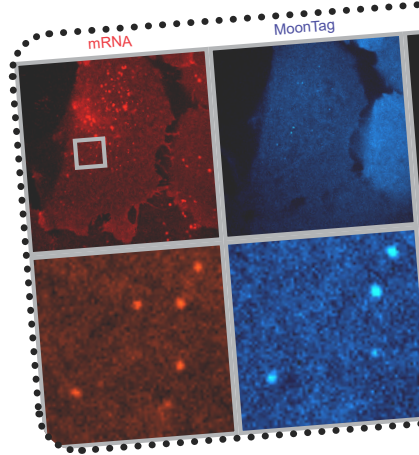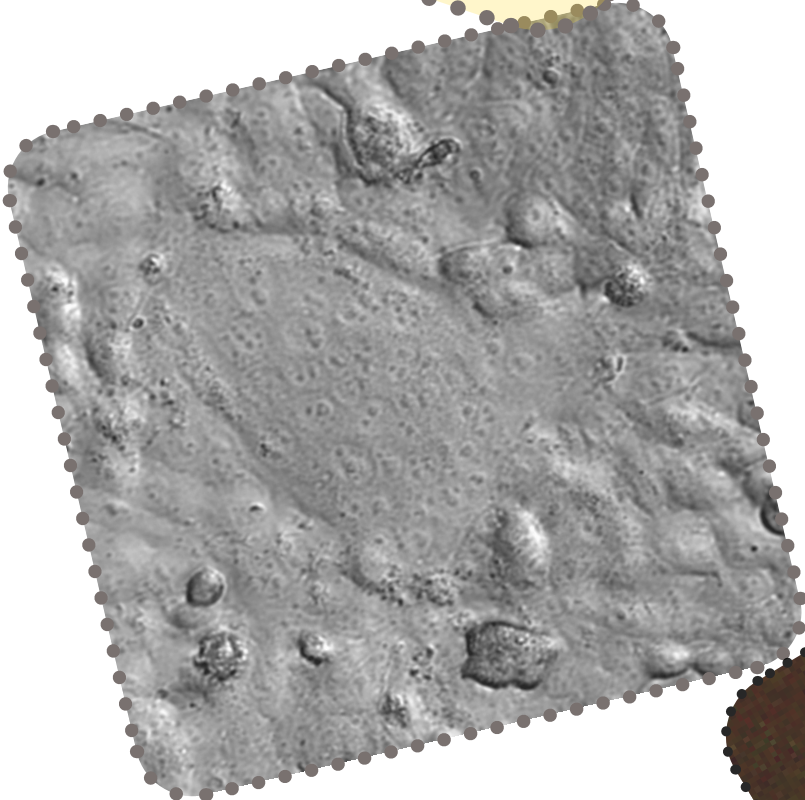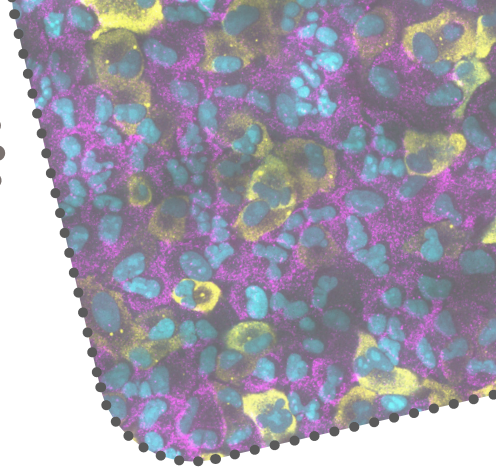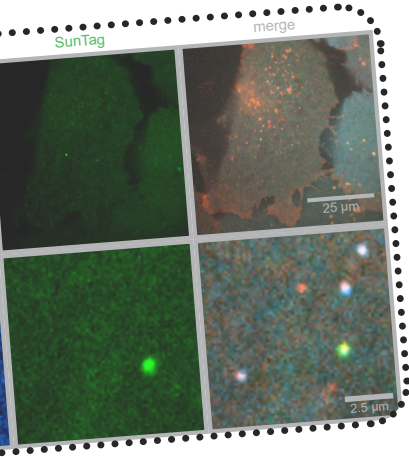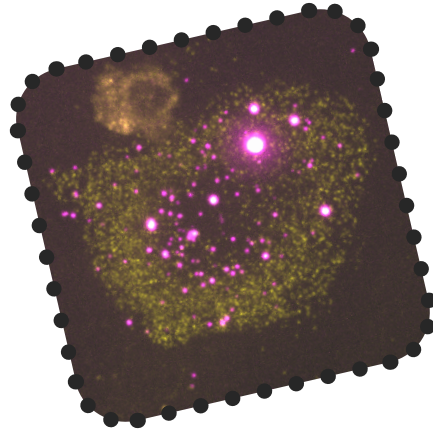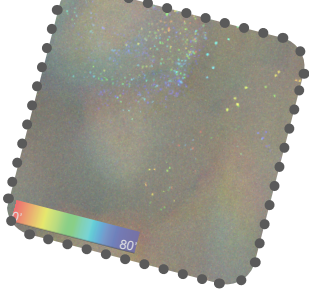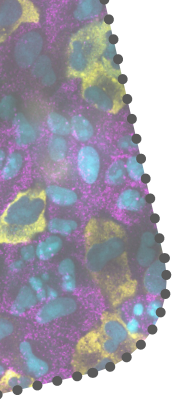
ribosomes

Start

Stop

mRNA

sequence.GG UUG CAC GUA UUG CAA AUA CUU GA.

peptide   ..   L   H   V   L   ..

peptide   ... V   A   R   I   A   N   T

* coding sequence

* alternative reading fra

Addendum
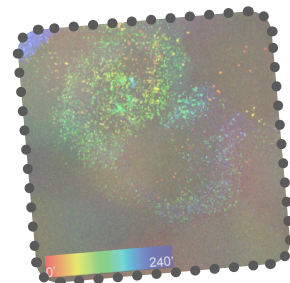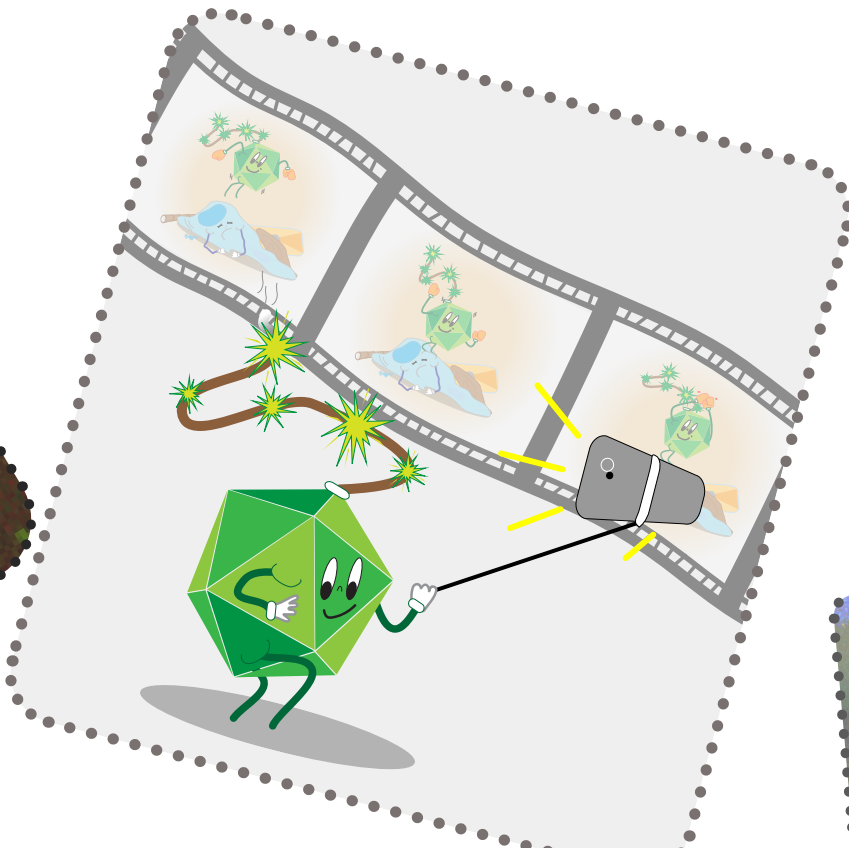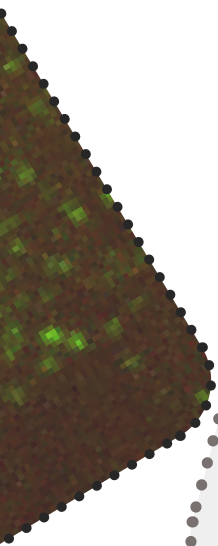
REFERENCES

Acker, M.G., Shin, B.S., Dever, T.E., and Lorsch, J.R. (2006). Interaction between eukaryotic initiation factors 1A and 5B is required for efficient ribosomal subunit joining. J. Biol. Chem. 281, 8469–8475.

Ahlquist, P., Noueiry, A.O., Lee, W.-M., Kushner, D.B., and Dye, B.T. (2003). Host Factors in Positive-Strand RNA Virus Genome Replication. J. Virol. 77, 8181–8186.

Ahmadian, G., Randhawa, J.S., and Easton, A.J. (2000). Expression of the ORF-2 protein of the human respiratory syncytial virus M2 gene is initiated by a ribosomal termination-dependent reinitiation mechanism. EMBO J. 19, 2681.

Aitken, C.E., and Lorsch, J.R. (2012). A mechanistic overview of translation initiation in eukaryotes. Nat. Struct. Mol. Biol. 19, 568–576.

Alain, T., Morita, M., Fonseca, B.D., Yanagiya, A., Siddiqui, N., Bhat, M., Zammit, D., Marcus, V., Metrakos, P., Voyer, L.-A., et al. (2012). eIF4E/4E-BP Ratio Predicts the Efficacy of mTOR Targeted Therapies. Cancer Res. 72, 6468–6476.

Alekhina, O.M., and Vassilenko, K.S. (2012). Translation initiation in eukaryotes: versatility of the scanning model. Biochemistry.Biokhimiia 77, 1465–1477.

Alessandra-Vigano, M., Ell, C.M., Kustermann, M.M.M., Aguilar, G., Matsuda, S., Zhao, N., Stasevich, T.J., Affolter, M., and Pyrowolakis, G. (2021). Protein manipulation using single copies of short peptide tags in cultured cells and in Drosophila melanogaster. Development 148.

Alexopoulou, L., Holt, A.C., Medzhitov, R., and Flavell, R.A. (2001). Recognition of double-stranded RNA and activation of NF-κB by Toll-like receptor 3. Nat. 2001 4136857 413, 732–738.

Aljabr, W., Touzelet, O., Pollakis, G., Wu, W., Munday, D.C., Hughes, M., Hertz-Fowler, C., Kenny, J., Fearns, R., Barr, J.N., et al. (2016). Investigating the Influence of Ribavirin on Human Respiratory Syncytial Virus RNA Synthesis by Using a High-Resolution Transcriptome Sequencing Approach. J. Virol. 90, 4876–4888.

Anderson, L.J., Jadhao, S.J., Paden, C.R., and Tong, S. (2021). Functional Features of the Respiratory Syncytial Virus G Protein. Viruses 13.

Andino, R., Silvera, D., Suggett, S.D., Achacoso, P.L., Miller, C.J., Baltimore, D., and Feinberg, M.B. (1994). Engineering Poliovirus as a Vaccine Vector for the Expression of Diverse Antigens. Science 265, 1448–1451.

Annels, N.E., Mansfield, D., Arif, M., Ballesteros-Merino, C., Simpson, G.R., Denyer, M., Sandhu, S.S., Melcher, A.A., Harrington, K.J., Davies, B., et al. (2019). Phase I Trial of an ICAM-1-Targeted Immunotherapeutic-Coxsackievirus A21 (CVA21) as an Oncolytic Agent Against Non Muscle-Invasive Bladder Cancer. Clin. Cancer Res. 25, 5818–5831.

Archer, S.K., Shirokikh, N.E., Beilharz, T.H., and Preiss, T. (2016). Dynamics of ribosome scanning and recycling revealed by translation complex profiling. Nature 535, 570–574.

Arnold, J.J., and Cameron, C.E. (2000). Poliovirus RNA-dependent RNA polymerase (3D(pol)). Assembly of stable, elongation-competent complexes by using a symmetrical primer-template substrate. J. Biol. Chem. 275, 5329–5336.

Atkins, J.F., Loughran, G., Bhatt, P.R., Firth, A.E., and Baranov, P. V (2016). Ribosomal frameshifting and transcriptional slippage: From genetic steganography and cryptography to adventitious use. Nucleic Acids Res. 44.

Aylett, C.H.S., Boehringer, D., Erzberger, J.P., Schaefer, T., and Ban, N. (2015). Structure of a yeast 40S-eIF1-eIF1A-eIF3-eIF3j initiation complex. Nat. Struct. Mol. Biol. 22.

Azimifar, S.B., Nagaraj, N., Cox, J., and Mann, M. (2014). Cell-Type-Resolved Quantitative Proteomics of Murine Liver. Cell Metab. 20, 1076–1087.

Baggen, J., Thibaut, H.J., Strating, J.R.P.M.P.M., and van Kuppeveld, F.J.M.M. (2018). The life cycle of non-polio enteroviruses and how to target it. Nat. Rev. Microbiol. 16, 368–381.

Balleza, E., Kim, J.M., and Cluzel, P. (2018). Systematic characterization of maturation time of fluorescent proteins in living cells. Nat. Methods 15, 47–51.

Barbosa, C., Peixeiro, I., Romão, L., Hentze, M., Zacchetti, D., and Romao, L. (2013). Gene expression regulation by upstream open reading frames and human disease. PLoS Genet. 9, e1003529.

Barik, S. (1992). Transcription of human respiratory syncytial virus genome RNA in vitro: requirement of cellular factor(s). J. Virol. 66, 6813.

Barr, J.N., Tang, X., Hinzman, E., Shen, R., and Wertz, G.W. (2008). The VSV polymerase can initiate at mRNA start sites located either up or downstream of a transcription termination signal but size of the intervening intergenic region affects efficiency of initiation. Virology 374, 361–370.

Barrows, N.J., Campos, R.K., Liao, K.C., Prasanth, K.R., Soto-Acosta, R., Yeh, S.C., Schott-Lerner, G., Pompon, J., Sessions, O.M., Bradrick, S.S., et al. (2018). Biochemistry and Molecular Biology of Flaviviruses. Chem. Rev. 118, 4448–4482.

Barton, D.J., and Flanegan, J.B. (1997). Synchronous replication of poliovirus RNA: initiation of negative-strand RNA synthesis requires the guanidine-inhibited activity of protein 2C. J. Virol. 71, 8482–8489.

Barton, D.J., Morasco, B.J., and Flanegan, J.B. (1999). Translating ribosomes inhibit poliovirus negative-strand RNA synthesis. J. Virol. 73, 10104–10112.

Belkowski, L.S., and Sen, G.C. (1987). Inhibition of vesicular stomatitis viral mRNA synthesis by interferons. J. Virol. 61, 653–660.

Belov, G.A., Evstafieva, A.G., Rubtsov, Y.P., Mikitas, O. V., Vartapetian, A.B., and Agol, V.I. (2000). Early alteration of nucleocytoplasmic traffic induced by some RNA viruses. Virology 275, 244–248.

Ben M'hadheb-Gharbi, M., Gharbi, J., Paulous, S., Brocard, M., Komaromva, A., Aouni, M., and Kean, K.M. (2006). Effects of the Sabin-like mutations in domain V of the internal ribosome entry segment on translational effciency of the Coxsackievirus B3. Mol Gen Genomics 276, 402–412.

Bermingham, A., and Collins, P.L. (1999). The M2–2 protein of human respiratory syncytial virus is a regulatory factor involved in the balance between RNA replication and transcription. Proc. Natl. Acad. Sci. U. S. A. 96, 11259.

Blondot, M.-L., Dubosclard, V., Fix, J., Lassoued, S., Aumont-Nicaise, M., Bontems, F., Eléouët, J.-F., and Sizun, C. (2012). Structure and Functional Analysis of the RNA- and Viral Phosphoprotein-Binding Domain of Respiratory Syncytial Virus M2-1 Protein. PLOS Pathog. 8, e1002734.

Boersma, S., Khuperkar, D., Verhagen, B.M.P., Sonneveld, S., Grimm, J.B., Lavis, L.D., and Tanenbaum, M.E. (2019). Multi-Color or Single-Molecule Imaging Uncovers Extensive Heterogeneity in mRNA Decoding. Cell 178, 458-472.e19.

Boersma, S., Rabouw, H.H., Bruurs, L.J.M., Pavlovič, T., van Vliet, A.L.W., Beumer, J., Clevers, H., van Kuppeveld, F.J.M., and Tanenbaum, M.E. (2020). Translation and Replication Dynamics of Single RNA Viruses. Cell 183, 1930-1945.e23.

Bohlen, J., Fenzl, K., Kramer, G., Bukau, B., and Teleman, A.A. (2020). Selective 40S Footprinting Reveals Cap-Tethered Ribosome Scanning in Human Cells. Mol. Cell 79, 561-574.e5.

Botos, I., Liu, L., Wang, Y., Segal, D.M., and Davies, D.R. (2009). The Toll-like Receptor 3:dsRNA signaling complex. Biochim. Biophys. Acta 1789, 667.

I

II

III

IV

V

VI

&

Bouillier, C., Rincheval, V., Sitterlin, D., Blouquit-Laye, S., Desquesnes, A., Eléouët, J.F., Gault, E., and Rameix-Welti, M.A. (2019). Generation, amplification, and titration of recombinant respiratory syncytial viruses. J. Vis. Exp. 2019, e59218.

Bowers, J.R., Readler, J.M., Sharma, P., and Excoffon, K.J.D.A. (2017). Poliovirus Receptor: More than a simple viral receptor. Virus Res. 242, 1.

Brandenburg, B., Lee, L.Y., Lakadamyali, M., Rust, M.J., Zhuang, X., and Hogle, J.M. (2007). Imaging Poliovirus Entry in Live Cells. PLoS Biol. 5, e183.

Braselmann, E., Rathbun, C., Richards, E.M., and Palmer, A.E. (2020). Illuminating RNA Biology: Tools for Imaging RNA in Live Mammalian Cells. Cell Chem. Biol. 27, 891–903.

Brauburger, K., Boehmann, Y., Krähling, V., and Mühlberger, E. (2016). Transcriptional Regulation in Ebola Virus: Effects of Gene Border Structure and Regulatory Elements on Gene Expression and Polymerase Scanning Behavior. J. Virol. 90, 1898–1909.

Brown, M.C., and Gromeier, M. (2015). Oncolytic Immunotherapy Through Tumor-specific Translation and Cytotoxicity of Poliovirus. Discov. Med. 19, 359.

Buchholz, U.J., Finke, S., and Conzelmann, K.-K. (1999). Generation of Bovine Respiratory Syncytial Virus (BRSV) from cDNA: BRSV NS2 Is Not Essential for Virus Replication in Tissue Culture, and the Human RSV Leader Region Acts as a Functional BRSV Genome Promoter. J. Virol. 73, 251–259.

Burke, J.M., Moon, S.L., Matheny, T., and Parker, R. (2019). RNase L Reprograms Translation by Widespread mRNA Turnover Escaped by Antiviral mRNAs. Mol. Cell.

Cai, A., Jankowska-Anyszka, M., Centers, A., Chlebicka, L., Stepinski, J., Stolarski, R., Darzynkiewicz, E., and Rhoads, R.E. (1999). Quantitative Assessment of mRNA Cap Analogues as Inhibitors of in Vitro Translation†. Biochemistry 38, 8538–8547.

Calvo, S.E., Pagliarini, D.J., and Mootha, V.K. (2009). Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proc. Natl. Acad. Sci. U. S. A. 106, 7507–7512.

Cao, D., Gao, Y., and Liang, B. (2021). Structural Insights into the Respiratory Syncytial Virus RNA Synthesis Complexes. Viruses 13.

Castagné, N., Barbier, A., Bernard, J., Rezaei, H., Huet, J.C., Henry, C., Da Costa, B., and Eléouët, J.F. (2004). Biochemical characterization of the respiratory syncytial virus P-P and P-N protein complexes and localization of the P protein oligomerization domain. J. Gen. Virol. 85, 1643–1653.

Chakrabarti, A., Jha, B.K., and Silverman, R.H. (2011). New insights into the role of RNase L in innate immunity. J. Interf. Cytokine Res. 31, 49–57.

Chao, J.A., Patskovsky, Y., Almo, S.C., and Singer, R.H. (2008). Structural basis for the coevolution of a viral RNA-protein complex. Nat. Struct. Mol. Biol. 15, 103–105.

Chappell, S.A., Edelman, G.M., and Mauro, V.P. (2006). Ribosomal tethering and clustering as mechanisms for translation initiation. Proc. Natl. Acad. Sci. 103, 18077–18082.

Chase, A.J., and Semler, B.L. (2012). Viral subversion of host functions for picornavirus translation and RNA replication. Future Virol. 7, 179–191.

Chase, A.J., Daijogo, S., and Semler, B.L. (2014). Inhibition of Poliovirus-Induced Cleavage of Cellular Protein PCBP2 Reduces the Levels of Viral RNA Replication. J. Virol. 88, 3192–3201.

Chen, J., Brunner, A.-D., Cogan, J.Z., Nuñez, J.K., Fields, A.P., Adamson, B., Itzhak, D.N., Li, J.Y., Mann, M., Leonetti, M.D., et al. (2020). Pervasive functional translation of noncanonical human open reading frames. Science (80-.). 367, 1140–1146.

Chen, K.H., Boettiger, A.N., Moffitt, J.R., Wang, S., and Zhuang, X. (2015). Spatially resolved, highly multiplexed RNA profiling in single cells. Science 348, aaa6090.

Chen, S., Short, J.A.L., Young, D.F., Killip, M.J., Schneider, M., Goodbourn, S., and Randall, R.E. (2010). Heterocellular induction of interferon by negative-sense RNA viruses. Virology 407, 247.

Chen, X., Kopecky, D., Mihalic, J., Jeffries, S., Min, X., Heath, J., Deignan, J., Lai, S., Fu, Z., Guimaraes, C., et al. (2012). Structure-guided design, synthesis, and evaluation of guanine-derived inhibitors of the eIF4E mRNA-cap interaction. J. Med. Chem. 55, 3837–3851.

Chew, G.L., Pauli, A., and Schier, A.F. (2016). Conservation of uORF repressiveness and sequence features in mouse, human and zebrafish. Nat. Commun. 7, 11663.

Chitrakar, A., Rath, S., Donovan, J., Demarest, K., Li, Y., Sridhar, R.R., Weiss, S.R., Kotenko, S. V, Wingreen, N.S., and Korennykh, A. (2019). Real-time 2-5A kinetics suggest that interferons β and λ evade global arrest of translation by RNase L. Proc. Natl. Acad. Sci. U. S. A. 116, 2103–2111.

Chouaib, R., Safieddine, A., Pichon, X., Imbert, A., Kwon, O.S., Samacoits, A., Traboulsi, A.M., Robert, M.C., Tsanov, N., Coleno, E., et al. (2020). A Dual Protein-mRNA Localization Screen Reveals Compartmentalized Translation and Widespread Co-translational RNA Targeting. Dev. Cell 54, 773-791.e5.

Cohen, C.J., Shieh, J.T.C., Pickles, R.J., Okegawa, T., Hsieh, J.-T., and Bergelson, J.M. (2001). The coxsackievirus and adenovirus receptor is a transmembrane component of the tight junction. Proc. Natl. Acad. Sci. 98, 15191–15196.

Collins, P.L., Dickens, L.E., Buckler-White, A., Olmsted, R.A., Spriggs, M.K., Camargo, E., and Coelingh, K. V (1986). Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order. Proc. Natl. Acad. Sci. U. S. A. 83, 4594.

Coots, R.A., Liu, X.M., Mao, Y., Dong, L., Zhou, J., Wan, J., Zhang, X., and Qian, S.B. (2017). m6A Facilitates eIF4F-Independent mRNA Translation. Mol. Cell 68.

Couso, J.P., and Patraquim, P. (2017). Classification and function of small open reading frames. Nat. Rev. Mol. Cell Biol. 18, 575–589.

Cowton, V.M., McGivern, D.R., and Fearns, R. (2006). Unravelling the complexities of respiratory syncytial virus RNA synthesis. J. Gen. Virol. 87, 1805–1821.

Cox, R., and Plemper, R.K. (2015). The paramyxovirus polymerase complex as a target for next-generation anti-paramyxovirus therapeutics. Front. Microbiol. 6.

Crick, F. (1970). Central dogma of molecular biology. Nature 227, 561–563.

Daffis, S., Szretter, K.J., Schriewer, J., Li, J., Youn, S., Errett, J., Lin, T.Y., Schneller, S., Zust, R., Dong, H., et al. (2010). 2′-O methylation of the viral mRNA cap evades host restriction by IFIT family members. Nature 468, 452–456.

Damgaard, C.K., and Lykke-Andersen, J. (2011). Translational coregulation of 5′TOP mRNAs by TIA-1 and TIAR. Genes Dev. 25, 2057–2068.

Das, S., Maiti, T., Das, K., and Maitra, U. (1997). Specific interaction of eukaryotic translation initiation factor 5 (eIF5) with the beta-subunit of eIF2. J. Biol. Chem. 272, 31712–31718.

Dave, P., George, B., Balakrishnan, S., Sharma, D.K., Raheja, H., Dixit, N.M., and Das, S. (2019). Strand-specific affinity of host factor hnRNP C1/C2 guides positive to negative-strand ratio in Coxsackievirus B3 infection. RNA Biol. 16, 1286–1299.

de Wit, E., Van Doremalen, N., Falzarano, D., and Munster, V.J. (2016). SARS and MERS: recent insights into emerging coronaviruses. Nat. Rev. Microbiol. 14, 523–534.

DeFilippis, V.R., Robinson, B., Keck, T.M., Hansen, S.G., Nelson, J.A., and Früh, K.J. (2006). Interferon Regulatory Factor 3 Is Necessary for Induction of Antiviral Genes during Human Cytomegalovirus Infection. J. Virol. 80, 1032.

Der, S.D., Zhou, A., Williams, B.R.G., and Silverman, R.H. (1998). Identification of genes differentially regulated by interferon α, β, or γ using oligonucleotide arrays. Proc. Natl. Acad. Sci. U. S. A. 95, 15623.

Desjardins, A., Gromeier, M., Herndon, J.E., Beaubier, N., Bolognesi, D.P., Friedman, A.H., Friedman, H.S., McSherry, F., Muscat, A.M., Nair, S., et al. (2018). Recurrent Glioblastoma Treated with Recombinant Poliovirus. N. Engl. J. Med. 379, 150–161.

Dey, M., Cao, C., Dar, A.C., Tamura, T., Ozato, K., Sicheri, F., and Dever, T.E. (2005). Mechanistic Link between PKR Dimerization, Autophosphorylation, and eIF2α Substrate Recognition. Cell 122, 901–913.

Dinman, J.D. (2012). Mechanisms and implications of programmed translational frameshifting. Wiley Interdiscip. Rev. RNA 3, 661–673.

Doğanay, S., Lee, M.Y., Baum, A., Peh, J., Hwang, S.Y., Yoo, J.Y., Hergenrother, P.J., García-Sastre, A., Myong, S., and Ha, T. (2017). Single-cell analysis of early antiviral gene expression reveals a determinant of stochastic: IFNB1 expression. Integr. Biol. 9, 857–867.

Dohm, J.C., Lottaz, C., Borodina, T., and Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. Nucleic Acids Res. 36, e105.

Dolnik, O., Gerresheim, G.K., and Biedenkopf, N. (2021). New Perspectives on the Biogenesis of Viral Inclusion Bodies in Negative-Sense RNA Virus Infections. Cells 2021, Vol. 10, Page 1460 10, 1460.

Donovan, J., Rath, S., Kolet-Mandrikov, D., and Korennykh, A. (2017). Rapid RNase L–driven arrest of protein synthesis in the dsRNA response without degradation of translation machinery. RNA 23, 1660–1671.

Du Toit, A. (2020). Coronavirus replication factories. Nat. Rev. Microbiol. 2020 188 18, 411–411.

Duke, G.M., and Palmenberg, A.C. (1989). Cloning and Synthesis of Infectious Cardiovirus RNAs Containing Short, Discrete Poly(C) Tracts. J. Virol. 63, 1822–1826.

Dunkle, J.A., and Dunham, C.M. (2015). Mechanisms of mRNA frame maintenance and its subversion during translation of the genetic code. Biochimie 114, 90–96.

Dunn, J.G., Foo, C.K., Belletier, N.G., Gavis, E.R., and Weissman, J.S. (2013). Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster. Elife 2013.

Durfee, L.A., Lyon, N., Seo, K., and Huibregtse, J.M. (2010). The ISG15 Conjugation System Broadly Targets Newly Synthesized Proteins: Implications for the Antiviral Function of ISG15. Mol. Cell 38, 722.

Edelstein, A., Amodaj, N., Hoover, K., Vale, R., and Stuurman, N. (2010). Computer control of microscopes using µManager. Curr. Protoc. Mol. Biol. Chapter 14, Unit14.20.

Eliseev, B., Yeramala, L., Leitner, A., Karuppasamy, M., Raimondeau, E., Huard, K., Alkalaeva, E., Aebersold, R., and Schaffitzel, C. (2018). Structure of a human cap-dependent 48S translation pre-initiation complex. Nucleic Acids Res. 46.

Ellis, E.L., and Delbrück, M. (1939). The growth of bacteriophage. J. Gen. Physiol. 22, 365–384.

Eng, C.H.L., Lawson, M., Zhu, Q., Dries, R., Koulena, N., Takei, Y., Yun, J., Cronin, C., Karp, C., Yuan, G.C., et al. (2019). Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. Nature.

Erzberger, J.P., Stengel, F., Pellarin, R., Zhang, S., Schaefer, T., Aylett, C.H., Cimermancic, P., Boehringer, D., Sali, A., Aebersold, R., et al. (2014). Molecular architecture of the 40SeIF1eIF3 translation initiation complex. Cell 158, 1123–1135.

Estell, C., Stamatidou, E., El-Messeiry, S., and Hamilton, A. (2017). In situ imaging of mitochondrial translation shows weak correlation with nucleoid DNA intensity and no suppression during mitosis. J. Cell Sci. 130, 4193–4199.

Fearns, R., and Collins, P.L. (1999). Role of the M2-1 Transcription Antitermination Protein of Respiratory Syncytial Virus in Sequential Transcription. J. Virol. 73, 5852.

Fearns, R., and Deval, J. (2016). New antiviral approaches for respiratory syncytial virus and other mononegaviruses: Inhibiting the RNA polymerase. Antiviral Res. 134, 63–76.

Fearns, R., and Plemper, R.K. (2017). Polymerases of paramyxoviruses and pneumoviruses. Virus Res. 234, 87–102.

Felt, S.A., Sun, Y., Jozwik, A., Paras, A., Habibi, M.S., Nickle, D., Anderson, L., Achouri, E., Feemster, K.A., Cárdenas, A.M., et al. (2021). Detection of respiratory syncytial virus defective genomes in nasal secretions is associated with distinct clinical outcomes. Nat. Microbiol. 1–10.

Femino, A.M., Fay, F.S., Fogarty, K., and Singer, R.H. (1998). Visualization of single RNA transcripts in situ. Science 280, 585–590.

Feng, Q., Hato, S.V. V, Langereis, M.A.A., Zoll, J., Virgen-Slane, R., Peisley, A., Hur, S., Semler, B.L.L., van Rij, R.P., van Kuppeveld, F.J.M., et al. (2012). MDA5 detects the double-stranded RNA replicative form in picornavirus-infected cells. Cell Rep. 2, 1187–1196.

Fensterl, V., and Sen, G.C. (2011). The ISG56/IFIT1 Gene Family. J. Interf. Cytokine Res. 31, 71.

Ferreira, J.P., Noderer, W.L., Diaz de Arce, A.J., and Wang, C.L. (2014). Engineering ribosomal leaky scanning and upstream open reading frames for precise control of protein translation. Bioengineered 5, 186–192.

Feuer, R., Mena, I., Pagarigan, R., Slifka, M.K., and Whitton, J.L. (2002). Cell Cycle Status Affects Coxsackievirus Replication, Persistence, and Reactivation In Vitro. J. Virol. 76, 4430–4440.

Fitzgerald, K.A., McWhirter, S.M., Faia, K.L., Rowe, D.C., Latz, E., Golenbock, D.T., Coyle, A.J., Liao, S.-M., and Maniatis, T. (2003). IKKε and TBK1 are essential components of the IRF3 signaling pathway. Nat. Immunol. 2003 45 4, 491–496.

Flather, D., and Semler, B.L. (2015). Picornaviruses and nuclear functions: targeting a cellular compartment distinct from the replication site of a positive-strand RNA virus. Front. Microbiol. 6, 594.

Forrest, A.R.R., Kawaji, H., Rehli, M., Baillie, J.K., de Hoon, M.J.L., Haberle, V., Lassmann, T., Kulakovskiy, I. V, Lizio, M., Itoh, M., et al. (2014). A promoter-level mammalian expression atlas. Nature 507, 462–470.

Fraser, C.S., Lee, J.Y., Mayeur, G.L., Bushell, M., Doudna, J.A., and Hershey, J.W. (2004). The j-subunit of human translation initiation factor eIF3 is required for the stable binding of eIF3 and its subcomplexes to 40 S ribosomal subunits in vitro. J. Biol. Chem. 279, 8946–8956.

Frei, A.P., Bava, F.A., Zunder, E.R., Hsieh, E.W.Y., Chen, S.Y., Nolan, G.P., and Gherardini, P.F. (2016). Highly multiplexed simultaneous detection of RNAs and proteins in single cells. Nat. Methods 13, 269–275.

Fuentes, S., Tran, K.C., Luthra, P., Teng, M.N., and He, B. (2007). Function of the Respiratory Syncytial Virus Small Hydrophobic Protein. J. Virol. 81, 8361–8366.

Fusco, D.N., Pratt, H., Kandilas, S., Cheon, S.S.Y., Lin, W., Cronkite, D.A., Basavappa, M., Jeffrey, K.L., Anselmo, A., Sadreyev, R., et al. (2017). HELZ2 is an IFN effector mediating suppression of dengue virus. Front. Microbiol. 8.

Galloux, M., Risso-Ballester, J., Richard, C.-A., Fix, J., Rameix-Welti, M.-A., and Eléouët, J.-F. (2020). Minimal Elements Required for the Formation of Respiratory Syncytial Virus Cytoplasmic Inclusion Bodies In Vivo and In Vitro. MBio 11, 1–16.

Gamarnik, A. V, and Andino, R. (1998). Switch from translation to RNA replication in a positive-stranded RNA virus. Genes Dev. 12, 2293–2304.

Gandin, V., English, B.P., Freeman, M., Leroux, L.-P., Preibisch, S., Walpita, D., Jaramillo, M., and Singer, R.H. (2021). Cap-dependent translation initiation monitored in living cells. BioRxiv 2021.05.21.445166.

Gao, F.-B., Richter, J.D., and Cleveland, D.W. (2017). Rethinking Unconventional Translation in Neurodegeneration. Cell 171, 994–1000.

García, M.A., Gil, J., Ventoso, I., Guerra, S., Domingo, E., Rivas, C., and Esteban, M. (2006). Impact of Protein Kinase PKR in Cell Biology: from Antiviral to Antiproliferative Action. Microbiol. Mol. Biol. Rev. 70, 1032.

Garcia-Moreno, M., Noerenberg, M., Ni, S., Järvelin, A.I., González-Almela, E., Lenz, C.E., Bach-Pages, M., Cox, V., Avolio, R., Davis, T., et al. (2019). System-wide Profiling of RNA-Binding Proteins Uncovers Key Regulators of Virus Infection. Mol. Cell 196-211.e11.

Garmaroudi, F.S., Marchant, D., Hendry, R., Luo, H., Yang, D., Ye, X., Shi, J., and McManus, B.M. (2015). Coxsackievirus B3 replication and pathogenesis. Future Microbiol. 10, 629–653.

Gaspar, I., Wippich, F., Ephrussi, A., Gáspár, I., Wippich, F., and Ephrussi, A. (2018). Terminal Deoxynucleotidyl Transferase Mediated Production of Labeled Probes for Single-molecule FISH or RNA Capture. BIO-PROTOCOL 8.

Genuth, N.R., and Barna, M. (2018). The Discovery of Ribosome Heterogeneity and Its Implications for Gene Regulation and Organismal Life (Cell Press).

Ghildyal, R., Baulch-Brown, C., Mills, J., and Meanger, J. (2003). The matrix protein of Human respiratory syncytial virus localises to the nucleus of infected cells and inhibits transcription. Arch. Virol. 2003 1487 148, 1419–1429.

Ghildyal, R., Ho, A., and Jans, D.A. (2006). Central role of the respiratory syncytial virus matrix protein in infection. FEMS Microbiol. Rev. 30, 692–705.

Giess, A., Torres Cleuren, Y.N., Tjeldnes, H., Krause, M., Bizuayehu, T.T., Hiensch, S., Okon, A., Wagner, C.R., and Valen, E. (2020). Profiling of Small Ribosomal Subunits Reveals Modes and Regulation of Translation Initiation. Cell Rep. 31.

Gingras, A., Gygi, P., Raught, B., Polakiewicz, R., Abraham, R., Hoekstra, M., Aebersold, R., and Sonenberg, N. (1999). Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. Genes Dev. 13, 1422.

González-Sanz, R., Mata, M., Bermejo-Martín, J., Álvarez, A., Cortijo, J., Melero, J.A., and Martínez, I. (2016). ISG15 Is Upregulated in Respiratory Syncytial Virus Infection and Reduces Virus Growth through Protein ISGylation. J. Virol. 90, 3428.

Götzke, H., Kilisch, M., Martínez-Carranza, M., Sograte-Idrissi, S., Rajavel, A., Schlichthaerle, T., Engels, N., Jungmann, R., Stenmark, P., Opazo, F., et al. (2019). A rationally designed and highly versatile epitope tag for nanobody-based purification, detection and manipulation of proteins Affiliation. BioRxiv 640771.

Gould, P.S., and Easton, A.J. (2005). Coupled Translation of the Respiratory Syncytial Virus M2 Open Reading Frames Requires Upstream Sequences*.

Gould, P.S., Dyer, N.P., Croft, W., Ott, S., and Easton, A.J. (2014). Cellular mRNAs access second ORFs using a novel amino acid sequence-dependent coupled translation termination–reinitiation mechanism. RNA 20, 373.

Grandvaux, N., Servant, M.J., tenOever, B., Sen, G.C., Balachandran, S., Barber, G.N., Lin, R., and Hiscott, J. (2002). Transcriptional Profiling of Interferon Regulatory Factor 3 Target Genes: Direct Involvement in the Regulation of Interferon-Stimulated Genes. J. Virol. 76, 5532.

Griffiths, C.D., Bilawchuk, L.M., McDonough, J.E., Jamieson, K.C., Elawar, F., Cen, Y., Duan, W., Lin, C., Song, H., Casanova, J.-L., et al. (2020). IGF1R is an entry receptor for respiratory syncytial virus. Nature 1–5.

Grimm, J.B., English, B.P., Chen, J., Slaughter, J.P., Zhang, Z., Revyakin, A., Patel, R., Macklin, J.J., Normanno, D., Singer, R.H., et al. (2015). A general method to improve fluorophores for live-cell and single-molecule microscopy. Nat. Methods 12, 244–250.

Gromeier, M., and Nair, S.K. (2018). Recombinant Poliovirus for Cancer Immunotherapy. Annu. Rev. Med. 69, 289–299.

Groskreutz, D., Monick, M., Powers, L., Yarovinsky, T., Look, D., and Hunninghake, G. (2006). Respiratory syncytial virus induces TLR3 protein and protein kinase R, leading to increased double-stranded RNA responsiveness in airway epithelial cells. J. Immunol. 176, 1733–1740.

Gu, W., Zhou, T., and Wilke, C.O. (2010). A Universal Trend of Reduced mRNA Stability near the Translation-Initiation Site in Prokaryotes and Eukaryotes. PLOS Comput. Biol. 6, e1000664.

Guenther, U.-P., Weinberg, D.E., Zubradt, M.M., Tedeschi, F.A., Stawicki, B.N., Zagore, L.L., Brar, G.A., Licatalosi, D.D., Bartel, D.P., Weissman, J.S., et al. (2018). The helicase Ded1p controls use of near-cognate translation initiation codons in 5′ UTRs. Nature 559, 130–134.

Guimaraes, J.C., and Zavolan, M. (2016). Patterns of ribosomal protein expression specify normal and malignant human cells. Genome Biol. 2016 171 17, 1–13.

Gunišová, S., Hronová, V., Mohammad, M.P., Hinnebusch, A.G., and Valášek, L.S. (2018). Please do not recycle! Translation reinitiation in microbes and higher eukaryotes. FEMS Microbiol. Rev. 42, 165.

Guo, F., Li, S., Caglar, M.U., Mao, Z., Liu, W., Woodman, A., Arnold, J.J., Wilke, C.O., Huang, T.J., and Cameron, C.E. (2017). Single-Cell Virology: On-Chip Investigation of Viral Infection Dynamics. Cell Rep. 21, 1692–1704.

Gustin, K.E., and Sarnow, P. (2001). Effects of poliovirus infection on nucleo-cytoplasmic trafficking and nuclear pore complex composition. EMBO J. 20, 240–249.

Hambidge, S.J., and Sarnow, P. (1992). Translational enhancement of the poliovirus 5′ noncoding region mediated by virus-encoded polypeptide 2A. Proc. Natl. Acad. Sci. U. S. A. 89, 10272–10276.

Han, Y., Donovan, J., Rath, S., Whitney, G., Chitrakar, A., and Korennykh, A. (2014). Structure of Human RNase L Reveals the Basis for Regulated RNA Decay in the IFN Response. Science (80-.). 343, 1244–1248.

Hann, S., King, M., Bentley, D., Anderson, C., and Eisenman, R. (1988). A non-AUG translational initiation in c-myc exon 1 generates an N-terminally distinct protein whose synthesis is disrupted in Burkitt's lymphomas. Cell 52, 185–195.

Hann, S.R., Dixit, M., Sears, R.C., and Sealy, L. (1994). The alternatively initiated c-Myc proteins differentially regulate transcription through a noncanonical DNA-binding site. Genes Dev. 8, 2441–2452.

Hardy, R.W., and Wertz, G.W. (1998). The Product of the Respiratory Syncytial Virus M2 Gene ORF1 Enhances Readthrough of Intergenic Junctions during Viral Transcription. J. Virol. 72, 520.

Hashimshony, T., Senderovich, N., Avital, G., Klochendler, A., de Leeuw, Y., Anavy, L., Gennert, D., Li, S., Livak, K.J., Rozenblatt-Rosen, O., et al. (2016). CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. Genome Biol. 17, 77.

Hijano, D.R., Vu, L.D., Kauvar, L.M., Tripp, R.A., Polack, F.P., and Cormier, S.A. (2019). Role of type I interferon (IFN) in the respiratory syncytial virus (RSV) immune response and disease severity. Front. Immunol. 10, 566.

Hinnebusch, A.G., Ivanov, I.P., and Sonenberg, N. (2016). Translational control by 5′-untranslated regions of eukaryotic mRNAs. Science 352, 1413–1416.

Hoek, T.A., Khuperkar, D., Lindeboom, R.G.H., Sonneveld, S., Verhagen, B.M.P., Boersma, S., Vermeulen, M., and Tanenbaum, M.E. (2019). Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay. Mol. Cell 75, 324-339.e11.

Hoffmann, H.-H., Schneider, W.M., and Rice, C.M. (2015). Interferons and viruses: an evolutionary arms race of molecular interactions. Trends Immunol. 36, 124–138.

Horikami, S.M., Curran, J., Kolakofsky, D., and Moyer, S.A. (1992). Complexes of Sendai virus NP-P and P-L proteins are required for defective interfering particle genome replication in vitro. J. Virol. 66, 4901.

Htar, M.T.T., Yerramalla, M.S., Moïsi, J.C., and Swerdlow, D.L. (2020). The burden of respiratory syncytial virus in adults: a systematic review and meta-analysis. Epidemiol. Infect. 148.

Hu, M.J., Bogoyevitch, M.A., and Jans, D.A. (2020). Impact of respiratory syncytial virus infection on host functions: Implications for antiviral strategies. Physiol. Rev. 100, 1527–1594.

Hur, S. (2019). Double-Stranded RNA Sensors and Modulators in Innate Immunity. Annu. Rev. Immunol. 37, 349.

Hussain, T., Llaser, J.L., Fernandez, I.S., Munoz, A., Martin-Marcos, P., Savva, C.G., Lorsch, J.R., Hinnebusch, A.G., Ramakrishnan, V., Ll??cer, J.L., et al. (2014). Structural Changes Enable Start Codon Recognition by the Eukaryotic Translation Initiation Complex. Cell 159, 597–607.

Iacono, M., Mignone, F., and Pesole, G. (2005). uAUG and uORFs in human and rodent 5'untranslated mRNAs. Gene 349, 97–105.

Iadevaia, V., Caldarola, S., Tino, E., Amaldi, F., and Loreni, F. (2008). All translation elongation factors and the e, f, and h subunits of translation initiation factor 3 are encoded by 5'-terminal oligopyrimidine (TOP) mRNAs. RNA 14.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science (80-.). 324, 218–223.

Ingolia, N.T.T., Lareau, L.F.F., and Weissman, J.S.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell 147, 789–802.

Iwaski, S., Floor, S.N., and Ingolia, N.T. (2016). Rocaglates convert DEAD-box protein eIF4A into a sequence-selective translational repressor. Nature 534.

Jackson, R.J., Hellen, C.U., and Pestova, T.V (2012). Termination and post-termination events in eukaryotic translation. Adv. Protein Chem. Struct. Biol. 86, 45–93.

Janicki, S.M., Tsukamoto, T., Salghetti, S.E., Tansey, W.P., Sachidanandam, R., Prasanth, K.V., Ried, T., Shav-Tal, Y., Bertrand, E., Singer, R.H., et al. (2004). From Silencing to Gene Expression: Real-Time Analysis in Single Cells. Cell 116, 683.

Jayaram, D.R., Frost, S., Argov, C., Liju, V.B., Anto, N.P., Muraleedharan, A., Ben-Ari, A., Sinay, R., Smoly, I., Novoplansky, O., et al. (2021). Unraveling the hidden role of a uORF-encoded peptide as a kinase inhibitor of PKCs. Proc. Natl. Acad. Sci. 118.

Jefferies, H.B., Reinhard, C., Kozma, S.C., and Thomas, G. (1994). Rapamycin selectively represses translation of the "polypyrimidine tract" mRNA family. Proc. Natl. Acad. Sci. U. S. A. 91, 4441.

Jemielity, J., Fowler, T., Zuberek, J., Stepinski, J., Lewdorowicz, M., Niedzwiecka, A., Stolarski, R., Darzynkiewicz, E., and Rhoads, R.E. (2003). Novel "anti-reverse" cap analogs with superior translational properties. RNA 9, 1108–1122.

Jiang, P., Liu, Y., Ma, H.-C., Paul, A.V., and Wimmer, E. (2014). Picornavirus Morphogenesis. Microbiol. Mol. Biol. Rev. 78, 418.

John, S.P., Sun, J., Carlson, R.J., Cao, B., Bradfield, C.J., Song, J., Smelkinson, M., and Fraser, I.D.C. (2018). IFIT1 Exerts Opposing Regulatory Effects on the Inflammatory and Interferon Gene Programs in LPS-Activated Human Macrophages. Cell Rep. 25, 95–106.e6.

Johnstone, T.G., Bazzini, A.A., and Giraldez, A.J. (2016). Upstream ORFs are prevalent translational repressors in vertebrates. EMBO J. 35, 706–723.

Kallewaard, N.L., Bowen, A.L., and Crowe, J.E. (2005). Cooperativity of actin and microtubule elements during replication of respiratory syncytial virus. Virology 331, 73–81.

Kapp, L.D., and Lorsch, J.R. (2004). GTP-dependent recognition of the methionine moiety on initiator tRNA by translation factor eIF2. J. Mol. Biol. 335, 923–936.

Kashiwagi, K., Takahashi, M., Nishimoto, M., Hiyama, T.B., Higo, T., Umehara, T., Sakamoto, K., Ito, T., and Yokoyama, S. (2016). Crystal structure of eukaryotic translation initiation factor 2B. Nature 531.

Kearse, M.G., and Wilusz, J.E. (2017). Non-AUG translation: A new start for protein synthesis in eukaryotes. Genes Dev. 31.

Khuperkar, D., Hoek, T.A., Sonneveld, S., Verhagen, B.M.P., Boersma, S., and Tanenbaum, M.E. (2020). Quantification of mRNA translation in live cells using single-molecule imaging. Nat. Protoc. 15, 1–28.

Kim, B., Arcos, S., Rothamel, K., Jian, J., Rose, K.L., McDonald, W.H., Bian, Y., Reasoner, S., Barrows, N.J., Bradrick, S., et al. (2020). Discovery of Widespread Host Protein Interactions with the Pre-replicated Genome of CHIKV Using VIR-CLASP. Mol. Cell 78, 624–640.e7.

Kim, J.H., Lee, S.R., Li, L.H., Park, H.J., Park, J.H., Lee, K.Y., Kim, M.K., Shin, B.A., and Choi, S.Y. (2011). High Cleavage Efficiency of a 2A Peptide Derived from Porcine Teschovirus-1 in Human Cell Lines, Zebrafish and Mice. PLoS One 6, e18556.

Kitagawa, Y., Zhou, M., Yamaguchi, M., Komatsu, T., Takeuchi, K., Itoh, M., and Gotoh, B. (2009). Human metapneumovirus M2-2 protein inhibits viral transcription and replication. Microbes Infect. 12, 135–145.

Klasse, P.J. (2015). Molecular determinants of the ratio of inert to infectious virus particles. Prog. Mol. Biol. Transl. Sci. 129, 285–326.

Koch, A., Aguilera, L., Morisaki, T., Munsky, B., and Stasevich, T.J. (2020). Quantifying the dynamics of IRES and cap translation with single-molecule resolution in live cells. Nat. Struct. Mol. Biol. 1–10.

Kocmik, I., Piecyk, K., Rudzinska, M., Niedzwiecka, A., Darzynkiewicz, E., Grzela, R., and Jankowska-Anyszka, M. (2018). Modified ARCA analogs providing enhanced translational properties of capped mRNAs. Cell Cycle 17, 1624–1636.

Kolakofsky, D., Le Mercier, P., Iseni, F., and Garcin, D. (2004). Viral RNA polymerase scanning and the gymnastics of Sendai virus RNA synthesis. Virology 318, 463–473.

Kolupaeva, V.G., Lomakin, I.B., Pestova, T.V, and Hellen, C.U. (2003). Eukaryotic initiation factors 4G and 4A mediate conformational changes downstream of the initiation codon of the encephalomyocarditis virus internal ribosomal entry site. Mol. Cell. Biol. 23, 687–698.

Koo, B.K., Stange, D.E., Sato, T., Karthaus, W., Farin, H.F., Huch, M., Van Es, J.H., and Clevers, H. (2012). Controlled gene expression in primary Lgr5 organoid cultures. Nat. Methods 9, 81–83.

Kozak, M. (1978). How do eucaryotic ribosomes select initiation regions in messenger RNA? Cell 15, 1109–1123.

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44, 283–292.

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. 15, 8125–8148.

Kozak, M. (1991). A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes. Gene Expr. 1, 111–115.

Kräusslich, H.G., Nicklin, M.J., Toyoda, H., Etchison, D., and Wimmer, E. (1987). Poliovirus proteinase 2A induces cleavage of eucaryotic initiation factor 4F polypeptide p220. J. Virol. 61, 2711–2718.

Kumar, P., Sweeney, T.R., Skabkin, M.A., Skabkina, O.V., Hellen,

C.U.T., and Pestova, T. V. (2014). Inhibition of translation by IFIT family members is determined by their ability to interact selectively with the 5'-terminal regions of cap0-, cap1- and 5'pppmRNAs. Nucleic Acids Res. 42, 3228–3245.

Kuo, L., Fearns, R., and Collins, P.L. (1996). The structurally diverse intergenic regions of respiratory syncytial virus do not modulate sequential transcription by a dicistronic minigenome. J. Virol. 70, 6143.

Lai, J.K.F., Sam, I.-C., and Chan, Y.F. (2016). The Autophagic Machinery in Enterovirus Infection. Viruses 8.

Lamphear, B.J., Kirchweger, R., Skern, T., and Rhoads, R.E. (1995). Mapping of functional domains in eukaryotic protein synthesis initiation factor 4G (eIF4G) with picornaviral proteases. Implications for Cap-dependent and Cap-independent translational initiation. J. Biol. Chem. 270, 21975–21983.

Lanke, K.H.W., van der Schaar, H.M., Belov, G.A., Feng, Q., Duijsings, D., Jackson, C.L., Ehrenfeld, E., and van Kuppeveld, F.J.M. (2009). GBF1, a Guanine Nucleotide Exchange Factor for Arf, Is Crucial for Coxsackievirus B3 RNA Replication. J. Virol. 83, 11940–11949.

Larson, D.R., Zenklusen, D., Wu, B., Chao, J.A., and Singer, R.H. (2011). Real-Time Observation of Transcription Initiation and Elongation on an Endogenous Yeast Gene. Science 332, 475.

Larsson, A.J.M., Johnsson, P., Hagemann-Jensen, M., Hartmanis, L., Faridani, O.R., Reinius, B., Segerstolpe, Å., Rivera, C.M., Ren, B., and Sandberg, R. (2019). Genomic encoding of transcriptional burst kinetics. Nature 565, 251.

Lee, A.S., Kranzusch, P.J., Doudna, J.A., and Cate, J.H. (2016). eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. Nature 536, 96–99.

Lee, J.Y., Wing, P.A., Gala, D.S., Noerenberg, M., Järvelin, A.I., Titlow, J., Zhuang, X., Palmalux, N., Iselin, L., Thompson, M.K., et al. (2021). Absolute quantitation of individual SARS-CoV-2 RNA molecules: a new paradigm for infection dynamics and variant differences. BioRxiv 2021.06.29.450133.

Lee, K.-M., Chen, C.-J., and Shih, S.-R. (2017). Regulation Mechanisms of Viral IRES-Driven Translation. Trends Microbiol. 25, 546–561.

Lee, S.S., Liu, B., Lee, S.S., Huang, S.-X.X., Shen, B., and Qian, S.-B.B. (2012). Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. Proc. Natl. Acad. Sci. U. S. A. 109, E2424-32.

Lemaire, P.A., Anderson, E., Lary, J., and Cole, J.L. (2008). Mechanism of PKR Activation by dsRNA. J. Mol. Biol. 381, 351–360.

Leyrat, C., Renner, M., Harlos, K., Huiskonen, J.T., and Grimes, J.M. (2014). Drastic changes in conformational dynamics of the antiterminator M2-1 regulate transcription efficiency in pneumovirinae. Elife 2014.

Li, Y., Banerjee, S., Wang, Y., Goldstein, S.A., Dong, B., Gaughan, C., Silverman, R.H., and Weiss, S.R. (2016). Activation of RNase L is dependent on OAS3 expression during infection with diverse human viruses. Proc. Natl. Acad. Sci. U. S. A. 113, 2241–2246.

Liljeroos, L., Krzyzaniak, M.A., Helenius, A., and Butcher, S.J. (2013). Architecture of respiratory syncytial virus revealed by electron cryotomography. Proc. Natl. Acad. Sci. 110, 11133–11138.

Lin, Y., Li, F., Huang, L., Polte, C., Duan, H., Fang, J., Sun, L., Xing, X., Tian, G., Cheng, Y., et al. (2020). eIF3 Associates with 80S Ribosomes to Promote Translation Elongation, Mitochondrial Homeostasis, and Muscle Health. Mol. Cell 79, 575-587.e7.

Lind, C., and Aqvist, J. (2016). Principles of start codon recognition in eukaryotic translation initiation. Nucleic Acids Res. 44, 8425–8432.

Lonberg-Holm, K., Crowell, R., and Philipson, L. (1976). Unrelated animal viruses share receptors. Nature 259, 679–681.

Lutje Hulsik, D., Liu, Y., Strokappe, N.M., Battella, S., El Khattabi, M., McCoy, L.E., Sabin, C., Hinz, A., Hock, M., Macheboeuf, P., et al. (2013). A gp41 MPER-specific llama VHH requires a hydrophobic CDR3 for neutralization but not for antigen recognition. PLoS Pathog. 9, e1003202.

Lykke-Andersen, S., and Jensen, T.H. (2015). Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nat. Rev. Cell Biol. 16, 665–677.

Lyon, K., Aguilera, L.U., Morisaki, T., Munsky, B., and Stasevich, T.J. (2019). Live-Cell Single RNA Imaging Reveals Bursts of Translational Frameshifting. Mol. Cell 75, 172-183.e9.

Lyon, K.R., Aguilera, L.U., Morisaki, T., Munsky, B., and Stasevich, T.J. (2018). Live-cell single RNA imaging reveals bursts of translational frameshifting. bioRxiv 478040.

Lyubimova, A., Itzkovitz, S., Junker, J.P., Fan, Z.P., Wu, X., and Van Oudenaarden, A. (2013). Single-molecule mRNA detection and counting in mammalian tissue. Nat. Protoc. 8, 1743–1758.

Maag, D., Fekete, C.A., Gryczynski, Z., and Lorsch, J.R. (2005). A conformational change in the eukaryotic translation preinitiation complex and release of eIF1 signal recognition of the start codon. Mol. Cell 17, 265–275.

Mair, F., Erickson, J.R., Voillet, V., Simoni, Y., Bi, T., Tyznik, A.J., Martin, J., Gottardo, R., Newell, E.W., and Prlic, M. (2020). A Targeted Multi-omic Analysis Approach Measures Protein Expression and Low-Abundance Transcripts on the Single-Cell Level. Cell Rep. 31, 107499.

Marintchev, A., Edmonds, K.A., Marintcheva, B., Hendrickson, E., Oberer, M., Suzuki, C., Herdy, B., Sonenberg, N., and Wagner, G. (2009). Topology and regulation of the human eIF4A/4G/4H helicase complex in translation initiation. Cell 136, 447–460.

Martin-Marcos, P., Cheung, Y.-N., and Hinnebusch, A.G. (2011). Functional Elements in Initiation Factors 1, 1A, and 2β Discriminate against Poor AUG Context and Non-AUG Start Codons. Mol. Cell. Biol. 31, 4814–4831.

Mastrangelo, P., and Hegele, R.G. (2013). Host and Viral Factors in Respiratory Syncytial Virus Infection. Curr. Pediatr. Rep. 1, 149–157.

Mateju, D., Eichenberger, B., Voigt, F., Eglinger, J., Roth, G., and Chao, J.A. (2020). Single-Molecule Imaging Reveals Translation of mRNAs Localized to Stress Granules. Cell 0.

Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A. V., Patil, D.P., Linder, B., Pickering, B.F., Vasseur, J.J., Chen, Q., et al. (2017). Reversible methylation of m6Amin the 5′ cap controls mRNA stability. Nature 541.

McGivern, D.R., Collins, P.L., and Fearns, R. (2005). Identification of Internal Sequences in the 3′ Leader Region of Human Respiratory Syncytial Virus That Enhance Transcription and Confer Replication Processivity. J. Virol. 79, 2449.

Mears, H. V., Emmott, E., Chaudhry, Y., Hosmillo, M., Goodfellow, I.G., and Sweeney, T.R. (2019). Ifit1 regulates norovirus infection and enhances the interferon response in murine macrophage-like cells. Wellcome Open Res. 4, 82.

Mendelsohn, C.L., Wimmer, E., and Racaniello, V.R. (1989). Cellular receptor for poliovirus: Molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily. Cell 56, 855–865.

Merrick, W.C. (2015). eIF4F: a retrospective. J. Biol. Chem. 290, 24091–24099.

Meyer, K.D., Patil, D.P., Zhou, J., Zinoviev, A., Skabkin, M.A., Elemento, O., Pestova, T. V., Qian, S.B., and Jaffrey, S.R. (2015). 5'UTR m6A Promotes Cap-Independent Translation. Cell 163.

Moon, S.L., Morisaki, T., Khong, A., Lyon, K., Parker, R., and Stasevich, T.J. (2019). Multicolour single-molecule tracking of mRNA interactions with RNP granules. Nat. Cell Biol. 21, 162–168.

Morisaki, T., Lyon, K., DeLuca, K.F., DeLuca, J.G., English, B.P., Zhang, Z., Lavis, L.D., Grimm, J.B., Viswanathan, S., Looger, L.L., et al. (2016). Real-time quantification of single RNA translation dynamics in living cells. Science (80-.). 352, 1425–1429.

Moudy, R.M., Harmon, S.B., Sullender, W.M., and Wertz, G.W. (2003). Variations in transcription termination signals of human respiratory syncytial virus clinical isolates affect gene expression. Virology 313, 250–260.

Mukherjee, A., Morosky, S.A., Delorme-Axford, E., Dybdahl-Sissoko, N., Oberste, M.S., Wang, T., and Coyne, C.B. (2011). The Coxsackievirus B 3Cpro Protease Cleaves MAVS and TRIF to Attenuate Host Type I Interferon and Apoptotic Signaling. 7, e1001311.

Muraro, M.J., Dharmadhikari, G., Grün, D., Groen, N., Dielen, T., Jansen, E., van Gurp, L., Engelse, M.A., Carlotti, F., de Koning, E.J.P., et al. (2016). A Single-Cell Transcriptome Atlas of the Human Pancreas. Cell Syst. 3, 385–394.e3.

Nag, N., Lin, K.Y., Edmonds, K.A., Yu, J., Nadkarni, D., Marintcheva, B., and Marintchev, A. (2016). eIF1A/eIF5B interaction network and its functions in translation initiation complex assembly and remodeling. Nucleic Acids Res. 44, 7441–7456.

Nam, H., Hwang, B.J., Choi, D.Y., Shin, S., and Choi, M. (2020). Tobacco etch virus (TEV) protease with multiple mutations to improve solubility and reduce self-cleavage exhibits enhanced enzymatic activity. FEBS Open Bio 10, 619.

Neupane, R., Pisareva, V.P., Rodriguez, C.F., Pisarev, A. V, and Fernández, I.S. (2020). A complex IRES at the 5'-UTR of a viral mRNA assembles a functional 48S complex via an uAUG intermediate. Elife 9.

Niessing, D., Blanke, S., and Jäckle, H. (2002). Bicoid associates with the 5'-cap-bound complex of caudal mRNA and represses translation. Genes Dev. 16, 2576–2582.

Noderer, W.L., Flockhart, R.J., Bhaduri, A., Diaz de Arce, A.J., Zhang, J., Khavari, P.A., and Wang, C.L. (2014). Quantitative analysis of mammalian translation initiation sites by FACS-seq. Mol. Syst. Biol. 10, 748–748.

Noton, S.L., and Fearns, R. (2015). Initiation and regulation of paramyxovirus transcription and replication. Virology 479–480, 545–554.

Noton, S.L., Tremaglio, C.Z., and Fearns, R. (2019). Killing two birds with one stone: How the respiratory syncytial virus polymerase initiates transcription and replication. PLOS Pathog. 15, e1007548.

Ogino, T., and Green, T.J. (2019). RNA Synthesis and Capping by Non-segmented Negative Strand RNA Viral Polymerases: Lessons From a Prototypic Virus. Front. Microbiol. 10.

Oh, H.S., Pathak, H.B., Goodfellow, I.G., Arnold, J.J., and Cameron, C.E. (2009). Insight into Poliovirus Genome Replication and Encapsidation Obtained from Studies of 3B-3C Cleavage Site Mutants. J. Virol. 83, 9370–9387.

Okumura, F., Okumura, A.J., Uematsu, K., Hatakeyama, S., Zhang, D.-E., and Kamura, T. (2013). Activation of Double-stranded RNA-activated Protein Kinase (PKR) by Interferon-stimulated Gene 15 (ISG15) Modification Down-regulates Protein Translation. J. Biol. Chem. 288, 2839.

Paek, K.Y., Hong, K.Y., Ryu, I., Park, S.M., Keum, S.J., Kwon, O.S., and Jang, S.K. (2015). Translation initiation mediated by RNA looping. Proc. Natl. Acad. Sci. U. S. A. 112, 1041–1046.

Palam, L.R., Baird, T.D., and Wek, R.C. (2011). Phosphorylation of eIF2 facilitates ribosomal bypass of an inhibitory upstream ORF to enhance CHOP translation. J. Biol. Chem. 286, 10939–10949.

Passmore, L.A., Schmeing, T.M., Maag, D., Applefield, D.J., Acker, M.G., Algire, M.A., Lorsch, J.R., and Ramakrishnan, V. (2007). The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome. Mol. Cell 26, 41–50.

Patil, S., Fribourg, M., Ge, Y., Batish, M., Tyagi, S., Hayot, F., and Sealfon, S.C. (2015). Single-cell analysis shows that paracrine signaling by first responder cells shapes the interferon-β response to viral infection. Sci. Signal. 8.

Patton, J.T., Davis, N.L., and Wertz, G.W. (1984). N protein alone

satisfies the requirement for protein synthesis during RNA replication of vesicular stomatitis virus. J. Virol. 49, 303.

Patursky-Polischuk, I., Stolovich-Rain, M., Hausner-Hanochi, M., Kasir, J., Cybulski, N., Avruch, J., Rüegg, M., Hall, M., and Meyuhas, O. (2009). The TSC-mTOR pathway mediates translational activation of TOP mRNAs by insulin largely in a raptor- or rictor-independent manner. Mol. Cell. Biol. 29, 640–649.

Paul, A. V, and Wimmer, E. (2015). Initiation of protein-primed picornavirus RNA synthesis. Virus Res. 206, 12–26.

Paul, D., and Bartenschlager, R. (2015). Flaviviridae Replication Organelles: Oh, What a Tangled Web We Weave. Https://Doi.Org/10.1146/Annurev-Virology-100114-055007 2, 289–310.

Peabody, D.S. (1989). Translation initiation at non-AUG triplets in mammalian cells. J. Biol. Chem. 264, 5031–5035.

Perng, Y.-C., and Lenschow, D.J. (2018). ISG15 in antiviral immunity and beyond. Nat. Rev. Microbiol. 2018 167 16, 423–439.

Pestova, T V, and Kolupaeva, V.G. (2002). The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. Genes Dev. 16, 2906–2922.

Peter, D., Igreja, C., Weber, R., Wohlbold, L., Weiler, C., Ebertsch, L., Weichenrieder, O., and Izaurralde, E. (2015). Molecular architecture of 4E-BP translational inhibitors bound to eIF4E. Mol. Cell 57, 1074–1087.

Pichlmair, A., Lassnig, C., Eberle, C.A., Górna, M.W., Baumann, C.L., Burkard, T.R., Búrckstúmmer, T., Stefanovic, A., Krieger, S., Bennett, K.L., et al. (2011). IFIT1 is an antiviral protein that recognizes 5'-triphosphate RNA. Nat. Immunol. 12, 624–630.

Pichon, X., Bastide, A., Safieddine, A., Chouaib, R., Samacoits, A., Basyuk, E., Peter, M., Mueller, F., and Bertrand, E. (2016). Visualization of single endogenous polysomes reveals the dynamics of translation in live human cells. J. Cell Biol. 214, 769–781.

Piedra, F.-A., Qiu, X., Teng, M.N., Avadhanula, V., Machado, A.A., Kim, D.-K., Hixson, J., Bahl, J., and Piedra, P.A. (2020). Non-gradient and genotype-dependent patterns of RSV gene expression. PLoS One 15, e0227558.

Pisarev, A. V, Kolupaeva, V.G., Pisareva, V.P., Merrick, W.C., Hellen, C.U., and Pestova, T. V (2006). Specific functional interactions of nucleotides at key -3 and +4 positions flanking the initiation codon with components of the mammalian 48S translation initiation complex. Genes Dev. 20, 624–636.

Pisareva, V.P., Pisarev, A. V Komar, A.A., Hellen, C.U., and Pestova, T. V (2008). Translation initiation on mammalian mRNAs with structured 5'UTRs requires DExH-box protein DHX29. Cell 135, 1237–1250.

Pontarotti, P., Abi-Rached, L., Yeh, J.H., and Paganini, J. (2020). Self-Peptidome Variation Shapes Individual Immune Responses. Trends Genet. 37, 414–420.

Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. Nat. Methods 5, 877–879.

Ramanan, V., Trehan, K., Ong, M.L., Luna, J.M., Hoffmann, H.H., Espiritu, C., Sheahan, T.P., Chandrasekar, H., Schwartz, R.E., Christine, K.S., et al. (2016). Viral genome imaging of hepatitis C virus to probe heterogeneous viral infection and responses to antiviral therapies. Virology 494, 236–247.

Rameix-Welti, M.A., Le Goffic, R., Hervé, P.L., Sourimant, J., Rémot, A., Riffault, S., Yu, Q., Galloux, M., Gault, E., and Eléouët, J.F. (2014). Visualizing the replication of respiratory syncytial virus in cells and in living mice. Nat. Commun. 5, 1–10.

Rath, S., Prangley, E., Donovan, J., Demarest, K., Wingreen, N.S., Meir, Y., and Korennykh, A. (2019). Concerted 2-5A-Mediated mRNA Decay and Transcription Reprogram Protein Synthesis in the dsRNA Response. Mol. Cell.

Reed, L.J., and Munech, H. (1938). A simple method of estimating fifty percent endpoints. Am. J. Epidemiol. 27, 493–497.

Rincheval, V., Lelek, M., Gault, E., Bouillier, C., Sitterlin, D., Blou-

quit-Laye, S., Galloux, M., Zimmer, C., Eleouet, J.F., and Rameix-Welti, M.A. (2017). Functional organization of cytoplasmic inclusion bodies in cells infected by respiratory syncytial virus. Nat. Commun. 8, 1–11.

Robbins-Pianka, A., Rice, M.D., and Weir, M.P. (2010). The mRNA landscape at yeast translation initiation sites. Bioinformatics 26, 2651–2655.

Rodriguez, J., and Larson, D.R. (2020). Transcription in Living Cells: Molecular Mechanisms of Bursting. Annu. Rev. Biochem. 89, annurev-biochem-011520-105250.

Roudko, V., Cimen Bozkus, C., Blank, S. V, Greenbaum, B., Correspondence, N.B., Bozkus, C.C., Orfanelli, T., Mcclain, C.B., Carr, C., O'donnell, T., et al. (2020). Shared Immunogenic Poly-Epitope Frameshift Mutations in Microsatellite Unstable Tumors. Cell 183, 1–16.

Roy, B., Vaughn, J.N., Kim, B.H., Zhou, F., Gilchrist, M.A., and Von Arnim, A.G. (2010). The h subunit of eIF3 promotes reinitiation competence during translation of mRNAs harboring upstream open reading frames. RNA 16, 748–761.

Ruijtenberg, S., Sonneveld, S., Cui, T.J., Logister, I., De Steenwinkel, D., Xiao, Y., Macrae, I.J., Joo, C., and Tanenbaum, M.E. (2020). mRNA structural dynamics shape Argonaute-target interactions. Nat. Struct. Mol. Biol. 1–12.

Ruiz Cuevas, M.V., Hardy, M.P., Hollý, J., Bonneil, É., Durette, C., Courcelles, M., Lanoix, J., Côté, C., Staudt, L.M., Lemieux, S., et al. (2021). Most non-canonical proteins uniquely populate the proteome or immunopeptidome. Cell Rep. 34, 108815.

Russell, A.B., Trapnell, C., and Bloom, J.D. (2018). Extreme heterogeneity of influenza virus infection in single cells. Elife 7.

Sabbah, A., Chang, T.H., Harnack, R., Frohlich, V., Tominaga, K., Dube, P.H., Xiang, Y., and Bose, S. (2009). Activation of innate immune antiviral response by NOD2. Nat. Immunol. 10, 1073.

Saeed, M., Kapell, S., Hertz, N.T., Wu, X., Bell, K., Ashbrook, A.W., Mark, M.T., Zebroski, H.A., Neal, M.L., Flodström-Tullberg, M., et al. (2020). Defining the proteolytic landscape during enterovirus infection. PLOS Pathog. 16, e1008927.

Sakaue-Sawano, A., Kurokawa, H., Morimura, T., Hanyu, A., Hama, H., Osawa, H., Kashiwagi, S., Fukami, K., Miyata, T., Miyoshi, H., et al. (2008). Visualizing Spatiotemporal Dynamics of Multicellular Cell-Cycle Progression. Cell 132, 487–498.

Samuel, C.E. (2001). Antiviral actions of interferons. Clin. Microbiol. Rev. 14, 778–809.

Sato, T., Stange, D.E., Ferrante, M., Vries, R.G.J., Van Es, J.H., Van Den Brink, S., Van Houdt, W.J., Pronk, A., Van Gorp, J., Siersema, P.D., et al. (2011). Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. Gastroenterology 141, 1762–1772.

Schlee, M., Roth, A., Hornung, V., Hagmann, C.A., Wimmenauer, V., Barchet, W., Coch, C., Janke, M., Mihailovic, A., Wardle, G., et al. (2009). Recognition of 5′-triphosphate by RIG-I helicase requires short blunt double-stranded RNA as contained in panhandle of negative strand virus. Immunity 31, 25.

Schneider, W.M., Chevillotte, M.D., and Rice, C.M. (2014). Interferon-Stimulated Genes: A Complex Web of Host Defenses. Annu. Rev. Immunol. 32, 513–545.

Schoggins, J.W. (2018). Recent advances in antiviral interferon-stimulated gene biology. F1000Research 7.

Schoggins, J.W., Wilson, S.J., Panis, M., Murphy, M.Y., Jones, C.T., Bieniasz, P., and Rice, C.M. (2011). A diverse range of gene products are effectors of the type I interferon antiviral response. Nature 472, 481–485.

Schueren, F., and Thoms, S. (2016). Functional Translational Readthrough: A Systems Biology Perspective. PLOS Genet. 12, e1006196.

Schulte, M.B., and Andino, R. (2014). Single-Cell Analysis Uncovers Extensive Biological Noise in Poliovirus Replication. J. Virol. 88, 6205–6212.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature 473, 337–342.

Sean, P., Nguyen, J.H.C., and Semler, B.L. (2009). Altered interactions between stem-loop IV within the 5′ noncoding region of coxsackievirus RNA and poly(rC) binding protein 2: Effects on IRES-mediated translation and viral infectivity. Virology 389, 45–58.

Sedeyn, K., Schepens, B., and Saelens, X. (2019). Respiratory syncytial virus nonstructural proteins 1 and 2: Exceptional disrupters of innate immune responses. PLoS Pathog. 15, e1007984.

Shaikh, F.Y., Crowe, J.E., and Jr (2013). Molecular mechanisms driving respiratory syncytial virus assembly (NIH Public Access).

Sharma, S., tenOever, B.R., Grandvaux, N., Zhou, G.-P., Lin, R., and Hiscott, J. (2003). Triggering the Interferon Antiviral Response Through an IKK-Related Pathway. Science (80-.). 300, 1148–1151.

Shi, T., McAllister, D.A., O'Brien, K.L., Simoes, E.A.F., Madhi, S.A., Gessner, B.D., Polack, F.P., Balsells, E., Acacio, S., Aguayo, C., et al. (2017a). Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. Lancet 390, 946–958.

Shi, Z., and Barna, M. (2015). Translating the Genome in Time and Space: Specialized Ribosomes, RNA Regulons, and RNA-Binding Proteins. Annu. Rev. Cell Dev. Biol. 31.

Shi, Z., Fujii, K., Kovary, K.M., Genuth, N.R., Röst, H.L., Teruel, M.N., and Barna, M. (2017b). Heterogeneous Ribosomes Preferentially Translate Distinct Subpools of mRNAs Genome-wide. Mol. Cell 67, 71-83.e7.

Shirokikh, N.E., and Preiss, T. (2018). Translation initiation by cap-dependent ribosome recruitment: Recent insights and open questions. Wiley Interdiscip. Rev. RNA 9.

Siddiqui, N., and Sonenberg, N. (2015). Signalling to eIF4E in cancer. Biochem. Soc. Trans. 43.

Silverman, R.H. (2007). Viral Encounters with 2′,5′-Oligoadenylate Synthetase and RNase L during the Interferon Antiviral Response. J. Virol. 81, 12720–12729.

Sin, J., Mangale, V., Thienphrapa, W., Gottlieb, R.A., and Feuer, R. (2015). Recent progress in understanding coxsackievirus replication, dissemination, and pathogenesis. Virology 484, 288–304.

Singer, Z.S., Ambrose, P.M., Danino, T., and Rice, C.M. (2021). Quantitative measurements of early alphaviral replication dynamics in single cells reveals the basis for superinfection exclusion. Cell Syst.

Singh, C.R., Yamamoto, Y., and Asano, K. (2004). Physical association of eukaryotic initiation factor (eIF) 5 carboxyl-terminal domain with the lysine-rich eIF2beta segment strongly enhances its binding to eIF3. J. Biol. Chem. 279, 49644–49655.

Sivanandam, V., LaRocca, C.J., Chen, N.G., Fong, Y., and Warner, S.G. (2019). Oncolytic Viruses and Immune Checkpoint Inhibition: The Best of Both Worlds. Mol. Ther. - Oncolytics 13, 93–106.

Smit, W.L., Spaan, C.N., Johannes de Boer, R., Ramesh, P., Martins Garcia, T., Meijer, B.J., Vermeulen, J.L.M., Lezzerini, M., MacInnes, A.W., Koster, J., et al. (2020). Driver mutations of the adenoma-carcinoma sequence govern the intestinal epithelial global translational capacity. Proc. Natl. Acad. Sci. 201912772.

Sokabe, M., and Fraser, C.S. (2014). Human eukaryotic initiation factor 2 (eIF2)-GTP-Met-tRNAi ternary complex and eIF3 stabilize the 43 S preinitiation complex. J. Biol. Chem. 289, 31827–31836.

Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell 136, 731–745.

Sonneveld, S., Verhagen, B.M.P., and Tanenbaum, M.E. (2020). Heterogeneity in mRNA Translation. Trends Cell Biol. 30, 606–618.

Sonneveld, S., Verhagen, B.M.P., and Tanenbaum, M.E. (2020). Heterogeneity in mRNA Translation. Trends Cell Biol. 30, 606–618.

Souii, A., Ben M'hadheb-Gharbi, M., and Gharbi, J. (2013). Role of RNA Structure Motifs in IRES-Dependent Translation Initiation of the Coxsackievirus B3: New Insights for Developing Live-Attenuated Strains for Vaccines and Gene Therapy. Mol. Biotechnol. 2013 552 55, 179–202.

Spann, K.M., Tran, K.-C., Chi, B., Rabin, R.L., and Collins, P.L. (2004). Suppression of the Induction of Alpha, Beta, and Gamma Interferons by the NS1 and NS2 Proteins of Human Respiratory Syncytial Virus in Human Epithelial Cells and Macrophages. J. Virol. 78, 4363.

Staals, R.H.J., Bronkhorst, A.W., Schilders, G., Slomovic, S., Schuster, G., Heck, A.J.R., Raijmakers, R., and Pruijn, G.J.M. (2010). Dis3-like 1: A novel exoribonuclease associated with the human exosome. EMBO J. 29, 2358–2367.

Stetson, D.B., and Medzhitov, R. (2006). Type I Interferons in Host Defense. Immunity 25, 373–381.

Sun, D., Chen, S., Cheng, A., and Wang, M. (2016). Roles of the Picornaviral 3C Proteinase in the Viral Life Cycle and Host Cells. Viruses 2016, Vol. 8, Page 82 8, 82.

Sun, L., Liu, S., and Chen, Z.J. (2010). SnapShot: Pathways of Antiviral Innate Immunity. 436 Cell 140.

Sun, Y., and López, C.B. (2016). Preparation of Respiratory Syncytial Virus with High or Low Content of Defective Viral Particles and Their Purification from Viral Stocks. BIO-PROTOCOL 6.

Sun, Y., and López, C.B. (2017). The innate immune response to RSV: Advances in our understanding of critical viral and host factors. Vaccine 35, 481–488.

Sun, Y., Jain, D., Koziol-White, C.J., Genoyer, E., Gilbert, M., Tapia, K., Jr., R.A.P., Hodinka, R.L., and López, C.B. (2015). Immunostimulatory Defective Viral Genomes from Respiratory Syncytial Virus Promote a Strong Innate Antiviral Response during Infection in Mice and Humans. PLOS Pathog. 11, e1005122.

Sunbul, M., Lackner, J., Martin, A., Englert, D., Hacene, B., Grün, F., Nienhaus, K., Nienhaus, G.U., and Jäschke, A. (2021). Super-resolution RNA imaging using a rhodamine-binding aptamer with fast exchange kinetics. Nat. Biotechnol. 1–5.

Suzuki, Y., Chin, W.-X., Han, Q., Ichiyama, K., Lee, C.H., Eyo, Z.W., Ebina, H., Takahashi, H., Takahashi, C., Tan, B.H., et al. (2016). Characterization of RyDEN (C19orf66) as an Interferon-Stimulated Cellular Inhibitor against Dengue Virus Replication. PLOS Pathog. 12, e1005357.

Svitkin, Y. V, Pause, A., Haghighat, A., Pyronnet, S., Witherell, G., Belsham, G.J., and Sonenberg, N. (2001). The requirement for eukaryotic initiation factor 4A (eIF4A) in translation is in direct proportion to the degree of mRNA 5′ secondary structure. RNA 7, 382–394.

Svitkin, Y. V., Cammack, N., Minor, P.D., and Almond, J.W. (1990). Translation deficiency of the sabin type 3 poliovirus genome: Association with an attenuating mutation C472 → U. Virology 175, 103–109.

Svitkin, Y. V., Maslova, S. V., and Agol, V.I. (1985). The Genomes of attenuated and virulent poliovirus strains differ in their in vitro translation efficiencies. Virology 147, 243–252.

Szamecz, B., Rutkai, E., Cuchalova, L., Munzarova, V., Herrmannova, A., Nielsen, K.H., Burela, L., Hinnebusch, A.G., and Valasek, L. (2008). eIF3a cooperates with sequences 5′ of uORF1 to promote resumption of scanning by post-termination ribosomes for reinitiation on GCN4 mRNA. Genes Dev. 22, 2414–2425.

Takai, Y., Miyoshi, J., Ikeda, W., and Ogita, H. (2008). Nectins and nectin-like molecules: roles in contact inhibition of cell movement and proliferation. Nat. Rev. Mol. Cell Biol. 2008 98 9, 603–615.

Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., and Vale, R.D. (2014). A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. Cell 159, 635–646.

Tanenbaum, M.E., Stern-Ginossar, N., Weissman, J.S., and Vale, R.D. (2015). Regulation of mRNA translation during mitosis. Elife 4.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. Nat. Methods 2009 65 6, 377–382.

Taniuchi, S., Miyake, M., Tsugawa, K., Oyadomari, M., and Oyadomari, S. (2016). Integrated stress response of vertebrates is regulated by four eIF2α kinases. Sci. Rep. 6, 32886.

Tawar, R.G., Duquerroy, S., Vonrhein, C., Varela, P.F., Damier-Piolle, L., Castagne, N., MacLellan, K., Bedouelle, H., Bricogne, G., Bhella, D., et al. (2009). Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. Science (80-.). 326, 1279–1283.

Tayyari, F., Marchant, D., Moraes, T.J., Duan, W., Mastrangelo, P., and Hegele, R.G. (2011). Identification of nucleolin as a cellular receptor for human respiratory syncytial virus. Nat. Med. 2011 179 17, 1132–1135.

Thoreen, C.C., Chantranupong, L., Keys, H.R., Wang, T., Gray, N.S., and Sabatini, D.M. (2012). A unifying model for mTORC1-mediated regulation of mRNA translation. Nature 486.

Thorne, L.G., and Goodfellow, I.G. (2014). Norovirus gene expression and replication. J. Gen. Virol. 95, 278–291.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. 32.

Tremaglio, C.Z., Noton, S.L., Deflubé, L.R., and Fearns, R. (2013). Respiratory Syncytial Virus Polymerase Can Initiate Transcription from Position 3 of the Leader Promoter. J. Virol. 87, 3196.

Truitt, M.L., Conn, C.S., Shi, Z., Pang, X., Tokuyasu, T., Coady, A.M., Seo, Y., Barna, M., and Ruggero, D. (2015). Differential Requirements for eIF4E Dose in Normal Development and Cancer. Cell 162.

Tuladhar, R., Yeu, Y., Tyler Piazza, J., Tan, Z., Rene Clemenceau, J., Wu, X., Barrett, Q., Herbert, J., Mathews, D.H., Kim, J., et al. (2019). CRISPR-Cas9-based mutagenesis frequently provokes on-target mRNA misregulation. Nat. Commun. 2019 101 10, 1–10.

Ullmer, W., and Semler, B.L. (2016). Diverse Strategies Used by Picornaviruses to Escape Host RNA Decay Pathways. Viruses 8.

Unbehaun, A., Borukhov, S.I., Hellen, C.U., and Pestova, T. V (2004). Release of initiation factors from 48S complexes during ribosomal subunit joining and the link between establishment of codon-anticodon base-pairing and hydrolysis of eIF2-bound GTP. Genes Dev. 18, 3078–3093.

Valarcher, J.-F., Furze, J., Wyld, S., Cook, R., Conzelmann, K.-K., and Taylor, G. (2003). Role of Alpha/Beta Interferons in the Attenuation and Immunogenicity of Recombinant Bovine Respiratory Syncytial Viruses Lacking NS Proteins. J. Virol. 77, 8426–8439.

van der Linden, L., Vives-Adrián, L., Selisko, B., Ferrer-Orta, C., Liu, X., Lanke, K., Ulferts, R., De Palma, A.M., Tanchis, F., Goris, N., et al. (2015). The RNA Template Channel of the RNA-Dependent RNA Polymerase as a Target for Development of Antiviral Therapy of Multiple Genera within a Virus Family. PLOS Pathog. 11, e1004733.

van der Schaar, H.M., Melia, C.E., van Bruggen, J.A.C., Strating, J.R.P.M., van Geenen, M.E.D., Koster, A.J., Bárcena, M., and van Kuppeveld, F.J.M. (2016). Illuminating the Sites of Enterovirus Replication in Living Cells by Using a Split-GFP-Tagged Viral Protein H. MSphere 1, e00104-16.

Vignuzzi, M., and López, C.B. (2019). Defective viral genomes are key drivers of the virus–host interaction. Nat. Microbiol. 4, 1075–1087.

Vissers, M., Remijn, T., Oosting, M., Jong, D.J. de, Diavatopoulos, D.A., Hermans, P.W.M., and Ferwerda, G. (2012). Respiratory syncytial virus infection augments NOD2 signaling in an IFN-β-dependent manner in human primary cells. Eur. J. Immunol. 42, 2727–2735.

Viswanathan, S., Williams, M.E., Bloss, E.B., Stasevich, T.J., Speer, C.M., Nern, A., Pfeiffer, B.D., Hooks, B.M., Li, W.-P., English, B.P., et al. (2015). High-performance probes for light and electron microscopy. Nat. Methods 12, 568–576.

Wagner, S., Herrmannová, A., Hronová, V., Gunišová, S., Sen, N.D., Hannan, R.D., Hinnebusch, A.G., Shirokikh, N.E., Preiss, T., and Valášek, L.S. (2020). Selective Translation Complex Profiling Reveals Staged Initiation and Co-translational Assembly of Initiation Factor Complexes. Mol. Cell 79, 546-560.e7.

Walsh, D., Mathews, M.B., and Mohr, I. (2013). Tinkering with translation: Protein synthesis in virus-infected cells. Cold Spring Harb. Perspect. Biol. 5, a012351.

Wang, C., Han, B., Zhou, R., and Zhuang, X. (2016). Real-Time Imaging of Translation on Single mRNA Transcripts in Live Cells. Cell 165, 990–1001.

Wang, X., Hou, J., Quedenau, C., and Chen, W. (2016b). Pervasive isoform-specific translational regulation via alternative transcription start sites in mammals. Mol. Syst. Biol. 12, 875.

Wang, Y., Ma, L., Stipkovits, L., Szathmary, S., Li, X., and Liu, Y. (2018). The strategy of picornavirus evading host antiviral responses: Nonstructural proteins suppress the production of IFNs. Front. Microbiol. 9.

Weber, F., Wagner, V., Rasmussen, S.B., Hartmann, R., and Paludan, S.R. (2006). Double-Stranded RNA Is Produced by Positive-Strand RNA Viruses and DNA Viruses but Not in Detectable Amounts by Negative-Strand RNA Viruses. J. Virol. 80, 5059.

Weingarten-Gabbay, S., Elias-Kirma, S., Nir, R., Gritsenko, A.A., Stern-Ginossar, N., Yakhini, Z., Weinberger, A., and Segal, E. (2016). Systematic discovery of cap-independent translation sequences in human and viral genomes. Science (80-.). 351.

Wessels, E., Duijsings, D., Niu, T.K., Neumann, S., Oorschot, V.M., de Lange, F., Lanke, K.H.W., Klumperman, J., Henke, A., Jackson, C.L., et al. (2006). A Viral Protein that Blocks Arf1-Mediated COP-I Assembly by Inhibiting the Guanine Nucleotide Exchange Factor GBF1. Dev. Cell 11, 191–201.

Wickenhagen, A., Sugrue, E., Lytras, S., Kuchi, S., Noerenberg, M., Turnbull, M.L., Loney, C., Herder, V., Allan, J., Jarmson, I., et al. (2021). A prenylated dsRNA sensor protects against severe COVID-19. Science (80-.).

Wu, B., Eliscovich, C., Yoon, Y.J., and Singer, R.H. (2016). Translation dynamics of single mRNAs in live cells and neurons. Science (80-.). 352.

Wu, J., and Chen, Z.J. (2014). Innate Immune Sensing and Signaling of Cytosolic Nucleic Acids.

Yan, X., Hoek, T.A., Vale, R.D., and Tanenbaum, M.E. (2016). Dynamics of Translation of Single mRNA Molecules in Vivo. Cell 165, 976–989.

Yanagiya, A., Suyama, E., Adachi, H., Svitkin, Y. V., Aza-Blanc, P., Imataka, H., Mikami, S., Martineau, Y., Ronai, Z.A., and Sonenberg, N. (2012). Translational Homeostasis via the mRNA Cap-Binding Protein, eIF4E. Mol. Cell 46, 847–858.

Zawatzky, R., De Maeyer, E., and De Maeyer-Guignard, J. (1985). Identification of individual interferon-producing cells by in situ hybridization. Proc. Natl. Acad. Sci. U. S. A. 82.

Zhang, L., Smit-McBride, Z., Pan, X., Rheinhardt, J., and Hershey, J.W. (2008). An oncogenic role for the phosphorylated h-subunit of human translation initiation factor eIF3. J. Biol. Chem. 283, 24047–24060.

Zhao, M., Zhang, J., Phatnani, H., Scheu, S., and Maniatis, T.

(2012). Stochastic Expression of the Interferon-β Gene. PLoS Biol. 10, 1001249.

Zhao, N., Kamijo, K., Fox, P., Oda, H., Morisaki, T., Sato, Y., Kimura, H., and Stasevich, T.J. (2018). A genetically encoded probe for imaging HA-tagged protein translation, localization, and dynamics in living cells and animals. bioRxiv 474668.

Zhou, J., Wan, J., Shu, X.E., Mao, Y., Liu, X.-M., Yuan, X., Zhang, X., Hess, M.E., Brüning, J.C., and Qian, S.-B. (2018). N 6 -Methyladenosine Guides mRNA Alternative Translation during Integrated Stress Response. Mol. Cell 69, 636–647.e7.

I

II

III

IV

V

VI

&

NEDERLANDSE SAMENVATTING

NIEUWE TECHNIEKEN OM MRNA TRANSLATIE EN RNA VIRUS INFECTIES IN BEELD TE BRENGEN

In onze cellen vinden continue allerlei processen plaats en deze processen zijn onderhevig aan veranderingen over tijd. De uitkomst van een proces is afhankelijk van veel factoren, zoals de exacte tijd en plaats van een proces, de aanwezigheid en de hoeveelheid regulatoren, of de kans dat er fouten worden gemaakt tijdens een proces. Van de meeste processen in onze cellen is grotendeels bekend welke aspecten bij kunnen dragen aan de uitkomst van een proces. Echter, de methoden die veelal toegepast worden om een proces te volgen, maken het niet mogelijk om de dynamiek en variaties van een proces in kaart te brengen. In deze scriptie worden nieuwe 'Track-en-Trace' methoden ontwikkeld en gebruikt om in te zoomen op individuele moleculen in levende cellen en de moleculen te volgen over tijd. Met behulp van de nieuwe methoden worden de processen mRNA translatie en RNA virus infecties in beeld gebracht.

MRNA TRANSLATIE

Hoewel al onze cellen dezelfde genetische informatie (DNA) bevatten, is de identiteit van cellen verschillend. De vorm, functie, activiteiten en daarmee identiteit van een cel worden bepaald door welke eiwitten een cel maakt. Eiwit productie vindt plaats via twee stappen: transcriptie en translatie. Tijdens transcriptie wordt een deel van het DNA omgezet in mRNA moleculen, die fungeren als boodschappers. Deze mRNA moleculen worden vervolgens getransleerd (vertaald) door ribosomen naar het eiwit dat volgens de boodschap geproduceerd dient te worden. mRNA translatie door ribosomen kan een grote impact hebben op de eiwitten in een cel: de hoeveelheid ribosomen beïnvloedt hoeveel eiwit gevormd wordt en de samenstelling van een eiwit is afhankelijk van de precieze manier waarop een ribosoom een mRNA transleert. De dynamiek van mRNA translatie en variaties tijdens mRNA translatie zijn dus van belang voor de identiteit van een cel.

Om translatie van individuele mRNAs te kunnen bestuderen zijn nieuwe microscopie methoden ontwikkeld en gebruikt: de SunTag en MoonTag technieken. Door middel van SunTag en MoonTag is het mogelijk om translerende mRNAs fluorescerend op te laten lichten. De helderheid van de fluorescentie per mRNA molecuul wordt bepaald door de hoeveelheid ribosomen die dat mRNA transleren en kan dus worden gebruikt om te bepalen hoeveel eiwitten er met een enkel mRNA gemaakt worden. Bovendien is het mogelijk om de kwaliteit van eiwit productie per individueel mRNA te bestuderen met de SunTag en MoonTag methoden. In de zogenaamde MashTag methode leidt de productie van een correct eiwit tot MoonTag fluorescentie, terwijl de productie van een ongebruikelijk eiwit SunTag fluorescentie genereert. De verhouding tussen SunTag en MoonTag fluorescentie per mRNA is vervolgens een indicatie voor de kwaliteit van eiwit productie. De nieuwe methoden kunnen dus

worden gebruikt om de kwantiteit en kwaliteit van mRNA translatie te onderzoeken.

De SunTag, MoonTag, en MashTag methoden zijn gebruikt om in detail de translatie van individuele mRNAs te volgen over tijd en hebben nieuwe inzichten over mRNA translatie opgeleverd. Allereerst blijkt dat translatie van een enkel mRNA molecuul erg kan fluctueren over tijd; in plaats van een continue snelheid van eiwit productie kan een mRNA molecuul switchen tussen een 'uit' of 'aan' status. Als een mRNA 'uit' staat, worden er geen ribosomen naar het mRNA gebracht om te transleren, terwijl er wel ribosomen worden geworven als een mRNA 'aan' staat. Daarnaast blijkt de kwaliteit van mRNA translatie veel lager te zijn dan eerder werd gedacht en erg te verschillen tussen mRNA moleculen die dezelfde boodschap hebben. Door sommige mRNAs worden hoofdzakelijk ongebruikelijk eiwitten geproduceerd, terwijl andere mRNAs vrijwel uitsluitend leiden tot de productie van correcte eiwitten. Tenslotte kan de kwaliteit van eiwit productie veranderen over tijd: een mRNA molecuul met hoge kwaliteit eiwit productie kan enkele minuten later switchen naar productie van vooral ongebruikelijke eiwitten. De productie van ongebruikelijke eiwitten in een cel kan het functioneren van een cel verstoren of kan mogelijk worden toegepast door een cel om de hoeveelheid correcte eiwitten te controleren. De nieuwe microscopie technieken kunnen worden toegepast om te onderzoeken wat de precieze functie van ongebruikelijke eiwit productie is en hoe een cel eiwit productie kan managen.

RNA VIRUS INFECTIES

Er zijn veel verschillende virussen en ze kunnen allerlei ziektes in mens en dier veroorzaken met grote gevolgen voor onze maatschappij. Een typische virus infectie begint met een enkel virus deeltje dat een cel binnendringt en probeert te vermenigvuldigen, zodat nieuwe virus deeltjes worden gevormd. De vorming van nieuwe virus deeltjes vergt de replicatie van het virale genoom en de productie van virale eiwitten. Bovendien moet een virus de antivirale verdedigingsmechanismen van de cel zien te ontwijken of onschadelijk zien te maken. Als de verdediging van de cel in staat is om een infectie te herkennen en neutraliseren, wordt verspreiding van een virus voorkomen. De uitkomst van een virus infectie is dus afhankelijk van een complexe wisselwerking tussen virale processen (virus replicatie, virale eiwit productie, ontwijking van afweermechanismen, etc.) en cellulaire processen (herkenning van infectie, snelle activatie van afweermechanismen, volledige blokkade van virus, etc.).

Positieve RNA (+RNA) en negatieve RNA (-RNA) virussen gebruiken RNA moleculen als hun genetische materiaal. +RNA virussen, zoals polio-, corona-, Zika-, en rhinovirussen, beginnen meteen na infectie met de productie van virale eiwitten. -RNA virussen, zoals influenza-, mazelen-, Ebola-, en RS (respiratoir syncytieel)-virussen, moeten eerst hun genoom omzetten in virale mRNAs voor de productie van virale eiwitten. Beide soorten virussen produceren virale polymerase eiwitten die het genoom van een virus repliceren en voor de replicatie moet tijdelijk de viral eiwit productie (van +RNA virussen) of de productie van virale mRNAs (van -RNA virussen) worden gepauzeerd. Aangezien veel huidige technieken niet gevoelig genoeg zijn om een infectie van het allereerste begin waar te nemen en te volgen, is het moeilijk om

te onderzoeken hoe een virus replicatie, virale eiwit productie, en/of virale mRNA productie controleert en hoe en wanneer de antivirale afweermechanismen in actie komen. Nieuwe microscopie technieken (genaamd VIRIM voor +RNA virussen en VIRIM$^{2.0}$ voor -RNA virussen) zijn ontwikkeld en toegepast om in levende cellen een infectie vanaf het allereerste begin te kunnen volgen en bestuderen.

Door gebruik te maken van VIRIM is ontdekt dat infecties met een picronavirus (een familie van +RNA virussen met o.a. rhinovirus en poliovirus) een aantal fases doorloopt. Tijdens de eerste fase wordt het virale genoom uit het infecterende virusdeeltje gebruikt om virale eiwitten te produceren. Vervolgens wordt het virale genoom in de tweede fase gebruikt voor de replicatie tot nieuwe genomen. De timing van de eerste fase en het slagingskans van de tweede fase verschillen enorm tussen individuele cellen. In sommige cellen, zijn de eerste twee fases afgerond binnen een uur na infectie, terwijl alleen al de eerste fase meer dan 8 uur kan duren in enkele andere cellen en de tweede fase nooit plaatsvindt in een fractie van cellen. De tweede fase (de genoom replicatie) is een belangrijk moment voor de antivirale afweer. Stimulatie van de antivirale afweer leidt tot een forse afname in de slagingskans van de tweede fase, maar heeft geen sterk effect op de eerste fase of op latere stadia van de infectie. De tweede fase, een vroeg moment in een infectie, is dus een belangrijk doel van de afweermechanismen en daarmee uitermate relevant voor uitkomst van een infectie. VIRIM kan worden gebruikt om ook voor andere +RNA virussen te ontdekken wat de virale zwaktes tijdens een infectie zijn, die kunnen worden gebruikt door afweermechanismen of kunnen dienen als een begin van de ontwikkeling van antivirale medicijnen.

Met behulp van VIRIM$^{2.0}$ is onderzocht hoe de virale mRNA productie van het -RNA virus RS-virus plaatsvindt en hoe de virale mRNA productie zich verhoudt tot de antivirale afweer. Om te kunnen repliceren heeft het RS-virus verschillende virale eiwitten nodig, die elk de productie van een mRNA vereisen. Er is strikte coördinatie in de productie van de verschillende virale mRNAs en de mate van productie hangt af van de locatie in het virale genoom waarvandaan een viraal mRNA moet worden geproduceerd. Bovendien is er grote variatie in de mate van virale mRNA productie tussen individuele cellen en over tijd. Gecombineerd onderzoek naar de virale mRNA productie en activatie van de antivirale afweermechanismen in dezelfde cellen heeft aangetoond dat cellen met een lage virale mRNA productie snelheid een hoge mate van antivirale afweer hebben en andersom. De mate en coördinatie van RS-virus mRNA productie speelt dus mogelijk een rol in de beslissing of mogelijkheid van een cel om de antivirale afweer op te starten. Met VIRIM$^{2.0}$ kan onderzocht worden hoe de activatie van een antivirale afweermechanismen zich verhoudt tot de productie van virale mRNAs van -RNA virussen, zoals het RS-virus.

In conclusie, de nieuwe methoden brengen dynamische processen in beeld, zoals mRNA translatie en RNA virus infectie. De methoden maken het mogelijk om over tijd in levende cellen te volgen hoe de processen plaatsvinden, hoe ze gecontroleerd worden in een cel, en hoe vaak er fouten optreden tijdens een proces. Bovendien laat het werk in deze scriptie zien hoe de ontwikkeling van nieuwe methoden kan leiden tot meer inzicht in mRNA translatie en RNA virus infecties.

I

II

III

IV

V

VI

&

LAY SUMMARY

NEW IMAGING TOOLS TO STUDY MRNA TRANSLATION AND RNA VIRUS INFECTIONS

Many processes take place all the time in our cells and these processes are subjected to changes over time. The outcome of a process depends on many aspects, such as the exact time and location of a process, the presence of regulators, or the chance that mistakes happen during the process. For most processes, there is extensive knowledge of the key players involved in the process. However, currently used methods to study a process often lack the ability to follow a process over time in live cells and thus study the dynamics of the process. In this thesis, various 'Track-and-Trace' methods are developed and used to zoom in on individual molecules in live cells. Using the new methods, the processes of mRNA translation and RNA virus infections are studied.

MRNA TRANSLATION

Even though all cells contain the same genetic code (DNA), cells in our body have widely different identities. The shape, activity, and function - and thus the identity of a cell - are determined by the proteins in a cell. Protein production takes place via two steps: transcription and translation. During transcription, a section of the DNA is converted into mRNA molecules that serve as messengers. These mRNA molecules are then translated into proteins by ribosomes, according to the message of the mRNAs. mRNA translation by ribosomes can greatly impact the proteins in a cell: the number of ribosomes per mRNA determines the amount of proteins synthesized per mRNA and the composition of a protein depends on the exact mode of translation. The dynamics of mRNA translation and the variations during mRNA translation thus impact the identity of a cell.

To study translation of individual mRNAs, new imaging methods are developed and applied: the SunTag and MoonTag assays. The SunTag and MoonTag assay enable the fluorescence labeling of an mRNA as it is translated by ribosomes. The brightness of the fluorescence corresponds to the number of ribosomes that translate an mRNA and can thus be used to calculate how much proteins are produced from a single mRNA. Moreover, combining SunTag and MoonTag into yet another assay (called MashTag) enables investigating the quality of protein production per mRNA. Using the MashTag assay, synthesis of a correct protein leads to MoonTag fluorescence, whereas production of an unusual protein gives rise to SunTag fluorescence. The ratio between MoonTag and SunTag fluorescence per mRNA is then a read-out for the quality of protein production. In sum, the new imaging methods can be used to analyze the quantity and quality of mRNA translation.

The SunTag, MoonTag, and MashTag methods have been applied to follow translation of individual mRNAs over time and have provided new insight into mRNA translation.

First, translation of single mRNAs is subjected to changes over time. Rather than steady-state protein production per mRNA, an mRNA molecule can switch between an 'on' and an 'off' state. In the 'off' state, no ribosomes are recruited to an mRNA and no proteins are synthesized from that mRNA, whereas new ribosomes can be recruited to the mRNA in its 'on' state. Second, the quality of mRNA translation is lower that previously expected and can vary widely between mRNA molecules with the same message. Some mRNAs are primarily used to produce unusual proteins, whereas other molecules almost exclusively lead to correct proteins. Finally, the quality of protein production per mRNA can change over time: an mRNA with high quality protein production can switch to mostly unusual protein production within minutes. Synthesis of unusual proteins can affect the function of a cell, but a cell may also employ unusual protein production to control the rate of correct protein production. The new imaging methods can be used to examine the regulation and consequences of unusual protein production.

RNA VIRUS INFECTION

There are many different viruses that can cause diseases in humans and animals, which can greatly impact our society. A virus infection typically starts with a single virus entering a single cell and the virus then tries to multiply to form extra viruses. Multiplication of a virus requires replication of the viral genome and synthesis of viral proteins. Moreover, a virus must evade the cell's antiviral defense mechanisms. If the cell's defense mechanisms manage to recognize and neutralize an infection, viral spread is prevented. The outcome of an infection therefore depends on a complex interplay between viral processes (including virus replication, viral protein production, evading the cell's antiviral defenses, etc.) and cellular processes (including recognition of an infection, quick activation of the antiviral defense mechanisms, and complete neutralization of the infection, etc.).

Positive RNA (+RNA) and negative RNA (-RNA) viruses use RNA molecules to store their genetic information. +RNA viruses, such as polio-, corona-, Zika-, and rhinoviruses, start producing viral proteins immediately after the start of an infection. In contrast, -RNA viruses, like influenza-, measles-, Ebola-, and RS (respiratory syncytial)- viruses have to convert their genome into viral mRNAs to enable viral protein production. Both types of viruses generate viral polymerase proteins to replicate their genomes. To ensure that the polymerase can replicate a genome, the viral protein production (of +RNA viruses) or synthesis of viral mRNAs (of -RNA viruses) needs to be temporarily paused. Currently available techniques generally lack sensitivity to detect an infection from the very first moment. It is therefore challenging to analyze viral replication, viral protein production, and/or viral mRNA synthesis and to examine how and when the antiviral defense mechanisms become active. New imaging approaches (named VIRIM and VIRIM[2.0] for +RNA or -RNA viruses respectively) are developed and applied to follow virus infection from the very start in live cells.

VIRIM is used to study picornaviruses, a family of +RNA viruses including rhinovirus and poliovirus, and revealed that early virus infection progresses via several infection

phases. During the first phase, the viral genome from the infecting particle is used to produce viral proteins. Next, during the second phase, the same genome molecule is used for replication. The timing and success-rate of these first two phases vary a lot between single cells. In some cells, both phases are completed within 1 hour after the start of the infection. In contrast, the first phase can last more than 8 hours in some other cells and phase 2 is not even completed in a fraction of cells. The second phase (replication of the genome) is an important moment for the antiviral response. Stimulation of the defense mechanisms results in a strong decrease in the success-rate of the second phase, however the first phase or later processes of the infection are unaffected. The earliest replication of the first genome is therefore a key target of the cell's antiviral defense mechanisms and the outcome of an infection is mainly determined very early in infection. VIRIM can be applied to other +RNA viruses to discover weaknesses of these viruses, that may represent important targets for the antiviral defenses or could help in developing new antiviral therapies.

VIRIM$^{2.0}$ is applied to investigate viral mRNA synthesis by the -RNA virus RSV and to study the relation between viral mRNA synthesis and the cell's antiviral defense mechanisms. RSV replication requires several viral proteins and the production of each of the viral proteins depends on the synthesis of viral mRNAs. There is strong coordination between the synthesis of the viral mRNAs and the the synthesis depends on the location along the viral genome wherefrom a viral mRNA is produced. Moreover, the rate of viral mRNA syntheses varies a lot between single cells and is subjected to fluctuations over time. Combined analysis of the viral mRNA synthesis rate and activation of the antiviral defenses in the same cells uncovered an anticorrelation in viral activity and the cell's ability or decision to launch an antiviral response. The rate and regulation of RSV mRNA synthesis may therefore influence the activation of antiviral defense mechanisms. VIRIM$^{2.0}$ will be a valuable tool to study how the activation of antiviral defenses relates to the synthesis of viral mRNAs by -RNA viruses, like RSV.

In conclusion, the new imaging assays can be used to study the dynamics and heterogeneity of mRNA translation and RNA virus infections. With the new assays, these processes can be followed over time in live cells and can be used to uncover how dynamic processes take place in cells, understand how these processes are regulated, and the assays exemplify how the development of new methods can lead to new biological insights.

ABOUT THE AUTHOR

Sanne Boersma was born on August 3rd 1993 in Amsterdam. She attended the Scholengemeenschap Cambium in Zaltbommel and completed here VWO degree in 2012. Sanne developed an interest in biology and decided to start the Bachelor's studies in Biomedical Sciences at Utrecht University in 2011. She completed the Bachelor *cum laude* in 2014 and continued her education at Utrecht University with the Master program Cancer Stem cells and Developmental biology. During the Master, Sanne did a research internship with the group of Prof. Sander van der Heuvel at the Developmental biology department of Utrecht University, focusing on new tools for precise genetic modifications in *C. elegans.* Sanne did a second research internship at the lab of Prof. Iain Cheeseman at the Whitehead Institute in Cambridge (USA), where she studied the mitotic assembly of protein complexes on centromeres. After obtaining her Master's degree *cum laude* in 2016, Sanne started her PhD studies in the lab of Dr. Marvin Tanenbaum at the Hubrecht Institute in Utrecht. During her PhD, she developed new tools to study mRNA translation and RNA virus infections, and the results are presented in this thesis. Sanne will move to Cambridge (USA) to join the research group of Prof. Xiaowei Zhuang at Harvard University as a post-doctoral fellow.

I

II

III

IV

V

VI

&

LIST OF PUBLICATIONS

- **<u>Boersma, S.</u>***, Khuperkar, D.*, Verhagen, B.M.P., Sonneveld, S., Grimm, J.B., Lavis, L.D., and Tanenbaum, M.E. (2019). Multi-Color Single-Molecule Imaging Uncovers Extensive Heterogeneity in mRNA Decoding. Cell *178*, 458-472.e19.

- Hoek, T.A.*, Khuperkar, D.*, Lindeboom, R.G.H., Sonneveld, S., Verhagen, B.M.P., **<u>Boersma, S.</u>**, Vermeulen, M., and Tanenbaum, M.E. (2019). Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay. Mol. Cell *75*, 324-339.e11.

- Khuperkar, D.*, Hoek, T.A.*, Sonneveld, S., Verhagen, B.M.P., **<u>Boersma, S.</u>**, and Tanenbaum, M.E. (2020). Quantification of mRNA translation in live cells using single-molecule imaging. Nat. Protoc. *15*, 1–28.

- **<u>Boersma, S.</u>**, Rabouw, H.H.[§], Bruurs, L.J.M.[§], Pavlovič, T., van Vliet, A.L.W., Beumer, J., Clevers, H., van Kuppeveld, F.J.M.[#], and Tanenbaum, M.E.[#] (2020). Translation and Replication Dynamics of Single RNA Viruses. Cell *183*, 1930-1945.e23.

- Krijger, P.H.L.*, Hoek, T.A.*, **<u>Boersma, S.</u>**, Donders, L.I.P.M., Broeders, M.M.C., Pieterse, M., Toonen, P.W., Logister, I., Verhagen, B.M.P., Verstegen, M.J.A.M., Verstegen, T.W., van Ravesteyn, R.J.T.M., Roymans, F., Mattiroli, J.,Vandesompele, M., Nijhuis, S., Meijer, A., van Weert, E.,Dekker, F.J., Dom, R., Ruijtenbeek, L.B.J., van der Velden[#], J.H.B., van de Bovenkamp, M., Bosch[#], W., de Laat[#], and M.E., Tanenbaum[#] (2021). A public–private partnership model for COVID-19 diagnostics. Nat. Biotechnol. *39*, 1182–1184.

Shared authorship: [*] Co-first authors; [§] Co-second authors; [#] Co-last authors

ACKNOWLEDGEMENT / DANKWOORD

Science is not done in solitude; many people have helped with experiments and have kept me sane. I am very proud of the final product of my PhD and I am incredible grateful for all the support that I received during my PhD time! Just a few people are highlighted here, but many more have contributed to the great journey of my PhD: thank you all!

Allereerst mijn begeleider, Marvin: dankjewel voor vijf geweldige jaren! Ik heb veel geleerd van onze wekelijkse/dagelijkse meetings, die soms konden uitlopen op lange discussies omdat we allebei helemaal opgingen in de details van mijn projecten. Toen ik bij de groep kwam, was je nog maar een jaar bezig als groepsleider. Inmiddels zijn het lab en de interesses van het lab gegroeid. Nog altijd ben je op de hoogte van alle details en heb je scherpe blik op alle projecten. Ik heb bewondering voor je passie voor wetenschap en hoe jij al vroeg in een project weet te beslissen welke richting weleens leuk kan uitpakken. Je hebt een geweldig leuke en diverse groep van mensen in je team weten te verzamelen en ik kijk uit naar de ontdekkingen die jullie in de toekomst zullen doen!

Wouter, dankjewel dat je mijn promoter wilde zijn. Hoewel we elkaar weinig hebben gesproken in de beginjaren van mijn PhD, hebben we dat ruimschoots goedgemaakt met de vele meetingen voor het Beast project om de test capaciteit te verbeteren tijdens de Covid-19 pandemie.

To the members of the reading and assessment committee: thanks for critical reading and approving the work presented in this thesis. I look forward to discussing my PhD work with you during the defense.

Mijn paranimfen, Lotte en Marleen: heel erg bedankt dat jullie me willen bijstaan op de dag van de verdediging! Lotte, ik ken niemand die zo trots is op haar nerdiness en dat ook nog eens zo creatief weet uit te drukken als jij. Onze ontelbare koffiepauzetjes, city trips, tuin centrum uitjes, klaagsessies, en kletspraatjes waren een hoogtepunt van mijn PhD. Marly, ik weet dat je absoluut niet begrijpt waarom iemand het leuk vindt om dagen door te brengen in een donkere microscopie kelder, maar toch was je enthousiast voor me als ik een doorbraak dacht te hebben. Door de Covid-19 crisis was je afgelopen jaren opeens vlakbij en konden we lekker vaak afspreken. Hopelijk kun jij ook snel weer reizen en duiken in een of andere tropische oceaan.

Current and former members of the Tanenbaum lab: thank you all for the great discussions, €2-challenges, borrels, and activities! I have been very lucky to have collaborated with and worked alongside such great colleagues.

Bram, we hebben eindeloos gekletst over de interpretatie van data, nieuwe hypotheses, politiek, en schaats,- en wielrenwedstrijden. Succes met het oplossen van het translation initiation mystery! Deepak, it is amazing that we managed to complete the MashTag project during our camping-in-the-lab months. Eating pizza

has not been the same, ever since you left the lab. Hopefully you are having fun in the UK and I hope that I inspired you to eat a salad every now and then! Iris, geweldig hoe jij je inzet voor het welzijn van PhD studenten. Veel succes met het afronden van je eigen PhD! Ive, zonder jou zou het lab binnen een week in elkaar storten. Jij zorgt dat alles op rolletjes verloopt en dat er geen rare dingen gebeuren. Daarnaast ben je ook een luisterend oor en heb ik regelmatig even stoom bij je mogen afblazen, waarvoor dank! Max, I am sorry for bothering you with all of my struggles with Harry and Meghan. Have fun with your PhD, but don't go home too late! Micha, thanks for your patience and eagerness to help and best of luck with your PhD! Lenno, hoewel onze projecten vrij ver uit elkaar lagen, heb ik veel gehad aan je scherpe vragen en lay-out suggesties. Sora, unfortunately the pandemic prevented us from interacting a lot, but I hope that your project works out well and perhaps you will even follow up on some of my earliest initiation factor observations! Stijn, ik kon altijd bij je aankloppen voor hulp, feedback, of artistieke interpretaties van data. Zelfs nadat je het lab al had verlaten, heb je me nog veel geholpen met de RSV modeling. Veel plezier en succes in de wereld van consultancy en data science! Suzan, het is al meer dan 8 (!) jaar geleden dat ik mijn eerste stage onder jouw begeleiding begon en heb geleerd hoe leuk wetenschap is. Ik weet zeker dat je een top lab aan het opzetten bent! Tim, je blijft natuurlijk altijd de senior PhD student! Jouw advies over mijn projecten en mijn aanpak heeft me vaak wakkergeschud. Dankjewel voor alle hulp en gezelligheid tijdens imaging sessions en feestjes!

It started as a fun side project, but it has evolved into the ever-growing Team Virus that (successfully?) tries to take over all offices of the Tanenbaum lab. To Team Virus: good luck with studying and naming all the new virus strains and thanks for the great collaborations! Huib and Lucas, even kletsen bij het koffiezetapparaat liep eigenlijk altijd uit op een discussie over hoe we nieuwe virussen konden ontwerpen en bestuderen. Dankjulliewel voor alle flauwe grappen! Dhanushika, it was a blast to work with you in the past two years. When I was about to leave the lab, I was struggling to let go of the RSV project. But I am very fortunate to leave it with you and Rupa and I am sure that you will manage to solve all - or at least many - RSV mysteries!

During my time in the Tanenbaum lab, many students have contributed to the projects and atmosphere in the lab. Particularly Tonja has been very important to me and my PhD thesis. Tonja, when I convinced you to work on the virus project, I had only done a single set of experiments. Somehow, you dared to join me in this exciting project, and I think it worked out very well. Good luck with your own PhD!

I also want to thank my collaborators outside the Tanenbaum lab. Frank en Arno: de samenwerking op het VIRIM project was zeer inspirerend en ik ben trots op het eindresultaat. Het was leuk om jullie input op m'n data te krijgen en na onze meetings had ik vaak bergen aan nieuwe experimenten om te doen. Marie-Anne, Jean-Francois, and Louis: thank you for introducing me to the intriguing world of RSV, for all your help with troubleshooting, and sharing lots of reagents. The entire Beast team: it has been an amazing experience to shift focus from basic science to diagnostics. Despite many struggles, we have managed to establish a great product that hopefully will help in combating the Covid-19 crisis.

The Hubrecht Institute is an amazing place to do science and make new friends. To the entire Hubrecht community: thank you for the great discussions and fun during the lunch seminars, BBQs, Masterclasses, borrels, gardening sessions, pie times, PhD retreats, holidays, dinners, and many more activities.

Ik ben ook erg dankbaar voor alle hulp vanuit de verschillende departmenten (FACS facility, technische service, IT helpdesk, mediakeuken, communicatie, etc.) en in het bijzonder voor alle hulp van de biologische veiligheidsfunctionarissen Marco en Sonja met het laten goedkeuren (en soms afkeuren) van m'n veiligheidsprotocollen door het RIVM.

Buiten het lab heb ik veel steun gehad van mijn vrienden en familie.

De garnalencocktail party (de naam uitleggen maakt het niet veel beter, dus ik laat het maar zo): het was heerlijk om met jullie te kunnen praten en klagen over het leven in de wetenschap en ik hoop dat we nog vele jaren contact houden.

Aan iedereen van de SMAV groep (wederom is het beter als ik de naam niet probeer uit te leggen): het is bijzonder hoe verschillend we door de jaren heen zijn geworden en hoe gezellig en als van ouds we het hebben als we samen eten, discussiëren, en elkaar belachelijk maken. Hoewel de meeste van jullie mijn interesses en keuzes niet begrijpen, kan ik wel altijd een luisterend oor vinden bij jullie.

En dan mijn familie: ik weet zeker dat jullie vol trots mijn verdediging zullen volgen. Hoewel het voor jullie soms lastig is dat ik naar het buitenland ga, kan ik altijd op jullie steun rekenen. Gelukkig weten jullie inmiddels hoe makkelijk het is om via videobellen contact te houden, zodat ik niets hoef te missen van alle chaos en gezelligheid aan de keukentafel.

Tenslotte bied ik aan m'n vrienden en familie mijn excuses aan voor alle afspraken waar ik een paar minuten of uren te laat kwam of eerder weg moest, omdat ik 'nog even iets op het lab wilde doen'. Ik heb enorm veel geluk dat jullie altijd voor me klaar staan. Dank voor al jullie geduld en support!

I

II

III

IV

V

VI

&