Reverse genetics and microRNAs in zebrafish

Erno Wienholds

Reverse genetics and microRNAs in zebrafish

Omgekeerde genetica en microRNA's in de zebravis

(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. W.H. Gispen, ingevolge het besluit van het College voor Promoties in het openbaar te verdedigen op dinsdag 8 november 2005 des ochtends te 10:30 uur

door

Erno Wienholds
geboren op 15 juni 1977, te Amsterdam

Promotor: Prof. Dr. R.H.A. Plasterk

Cover: miR-140 expression in cartilage of zebrafish embryo

ISBN: 90-393-4087-0

Printed by Wöhrmann Print Service

The work described in this thesis was performed at the Hubrecht Laboratory of the Royal Netherlands Academy of Arts and Sciences (KNAW), within the Graduate School for Developmental Biology, Utrecht, the Netherlands. This study was supported by a grant from the Netherlands Organization for Scientific Research (NWO).

Contents

Preface

The zebrafish (*Danio rerio*) has become an important genetic model organism in various research areas of modern biology. Originally, it was chosen as a vertebrate model system because its excellent properties that allow the study of embryonic development could be combined with powerful genetic analysis. Zebrafish eggs are fertilized externally and embryos subsequently develop synchronously outside of the body. Together with the relatively large egg size and optical transparency, these qualities make it possible to manipulate and study zebrafish embryonic development under a standard stereomicroscope. In addition, the embryonic development of zebrafish is extremely rapid compared to that of other vertebrate animals. During the first 24 hours of development embryos undergo early developmental processes, such as gastrulation and epiboly, and the primary organs become visible. After 48 hours of development a basic body plan is established and the major organs have been formed.

The other feature that makes zebrafish a good model organism is its power in genetic analysis. Although zebrafish have a relative long generation time (2-4 months), the ease of raising and keeping zebrafish at high densities, the large number of offspring from single crosses (typically more than 100 embryos each week), the long fertile lifespan (up to two years) and the ability to introduce mutations in the germ line at high frequencies by mutagenesis makes them ideally suited for forward mutagenesis screens. Such screens, for example, have led to the identification of thousands of mutants that are defective in developmental processes. More recent genetic screens have yielded many additional mutants in diverse biological processes, some of which are very reminiscent of human diseases. Great effort is undertaken to identify the genes responsible for these mutant phenotypes. However, the reverse genetic discovery of gene function, by inactivation of genes through knockout technologies, has not been possible in zebrafish until recently.

In this thesis, I describe the establishment of a knockout technology in zebrafish. Chapter 2 describes the creation of the first zebrafish knockout, the *rag1* gene, by target-selected mutagenesis. By resequencing a comprehensive library of mutagenized zebrafish, a null mutation in the *rag1* gene was identified. This mutation leads to a failure in rearrangement of the immunoglobulin locus and consequently results in immunodefiency. In Chapter 3, several refinements to this knockout technique are described. Instead of direct resequencing of mutagenized animals, the TILLING technique was used to pre-screen for point mutations in several genes. Using this method, thirteen potential zebrafish knockouts were identified. In Chapter 4, the current methodologies and potential future applications to do reverse genetic analysis in zebrafish are reviewed.

The main subject of this thesis is ‘microRNAs (miRNAs) in zebrafish development’. miRNAs are small non-coding RNA molecules that post-transcriptionally regulate gene expression. The base-pairing of miRNAs to target mRNAs results in translational inhibition or mRNA cleavage. Hundreds of miRNAs have been identified in various multicellular organisms and many miRNAs are evolutionarily conserved. Although the biological functions of most miRNAs are unknown, miRNAs are predicted to regulate up to 30% of the genes within the human genome. In the introduction (Chapter 1), the recent advances in miRNA biology are reviewed. Particularly, there is a focus on the roles of miRNAs in vertebrate development and disease. Chapter 5 describes the construction and analysis of the *dicer* knockout in zebrafish. Disruption of the miRNA-producing enzyme Dicer results in developmental arrest and failure to produce miRNAs, indicating that miRNAs are essential for vertebrate development. In Chapter 6, the miRNA expression patterns during zebrafish embryonic development are described. Most miRNA are expressed in a highly tissue-specific manner during segmentation and later stages, but not early in development, which

suggests that their role is not in tissue fate establishment but in differentiation or maintenance of tissue identity.

Chapter 1

Introduction:
MicroRNAs in animal development

Adapted from Wienholds and Plasterk (2005), *FEBS lett* 579: in press

MicroRNAs in animal development

1. The discovery of microRNAs

The founding microRNA, *lin-4*, was first thought to be an isolated case. The *lin-4* gene was identified by genetics in the nematode *Caenorhabditis elegans* (*C. elegans*) and is involved in developmental timing. Surprisingly, this gene does not encode a protein but codes for two small RNA molecules of ~60 and ~22 nucleotides in length [1]. The former is the precursor of the smaller and can form a secondary stem-loop structure that is phylogenetically conserved in nematodes. The smaller RNA molecule regulates the *lin-14* gene by base-paring to imperfect complementary sites in the 3´ untranslated region (UTR) of *lin-14* mRNA [2]. This form of post-transcriptional gene silencing by a small regulatory RNA was initially regarded as a rare phenomenon, specific for nematodes. However, the discovery of *let-7* changed this idea. *Let-7* is another small RNA that is also involved in developmental timing in *C. elegans* and similarly regulates the *lin-41* gene [3, 4]. Hence, *lin-4* and *let-7* were both called small temporal RNAs (stRNAs) for their role in developmental timing [5]. In contrast to the *lin-4/lin-14* pair, *let-7* and *lin-41* are evolutionarily conserved among a wide variety of multicellular organisms. Furthermore, *let-7* is also differentially expressed during development of various bilaterian animals [5].

This discovery implied that the post-transcriptional regulation of gene expression by these small RNAs was not restricted to nematodes and could be a general mechanism present in many multicellular organisms. Indeed, almost one hundred of such small regulatory RNAs were readily identified in *C. elegans*, *Drosophila* and human [6-8]. Since most of these newly identified small RNAs are not expressed in a temporal-specific manner, as *lin-4* or *let-7*, but in a spatial- or tissue-specific manner, the small regulatory RNAs are now called microRNAs (miRNAs). Currently, miRNAs have been identified in many other multicellular organisms, and it is estimated that vertebrate genomes may encode several hundred or perhaps up to a thousand different miRNAs [9, 10], which may regulate up to 20-30% of the genes [10, 11].

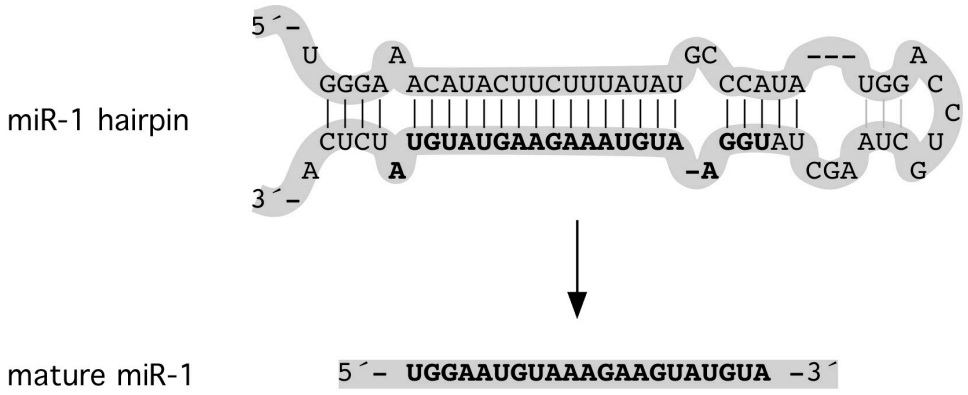


Fig. 1. Structures of the human precursor miR-1 hairpin, as encoded in the genome, and the mature miR-1. The mature miRNA is encoded in the 3´ arm of the precursor (indicated in bold).

2. Identification and expression of miRNAs

2.1. Identification of miRNAs

Currently, 2,909 miRNA genes have been annotated in miRBAse (release 7.0) [12]. To annotate a newly identified small RNA as miRNA it must fulfill several criteria. First, miRNAs are expressed as ~22-nucleotide RNA molecules. Second, miRNAs are encoded in the genome as phylogenetically conserved hairpin structures with a low free energy (see Fig. 1 for example). Third, miRNA expression levels are reduced in animals defective in miRNA biogenesis [13]. It is worth noting that not all miRNAs exactly fit these criteria. In particular, the requirements for annotation of closely related miRNAs in miRNA families or other species is less stringent [13]. Members of the miRNA gene class in animals are identified in several ways (Fig. 2).

The founding members, *lin-4* and *let-7*, were both identified in *C. elegans* based on their mutant phenotypes in forward genetic screens [1, 3]. Other forward genetic screens have only led to the identification of one additional miRNA in *C. elegans* (*lsy-6* [14]) and two additional miRNAs in *Drosophila* (*bantam* and *miR-14* [15, 16]). Although small in number, these five miRNAs are probably the best-studied miRNAs so far. The unbiased nature of forward genetics screens can discover miRNAs that cannot easily be detected by any other method. For example, the *lsy-6* miRNA could not be identified by either cDNA cloning nor computational predictions [14]. On the other hand, many miRNAs are likely to be missed in forward genetic screens because miRNA genes are relatively small genes (about 50 times smaller than the average protein coding gene), which are not often hit by the classical mutagens. In addition, many miRNAs might have redundant functions and therefore give only very subtle or no phenotypes when mutated [17].

To bypass the rare genetic discovery of miRNAs, several labs identified more than a hundred miRNAs by cDNA cloning and subsequent sequence analysis [6-8]. This approach has the advantage that it can be applied to any organism, even if little or no genomic information is available. In addition, miRNAs can be identified independent of their function, thus also allowing the identification of redundant miRNAs. By cDNA cloning, miRNAs have now been identified in diverse animals like nematodes [6, 8, 18], flies [7, 19], mammals [7, 20-26], frogs [27], fish [28, 29] and several mammalian viruses [30, 31].

One limitation of the cloning strategy is that there is a potential bias in cloning small RNA molecules. Highly expressed miRNAs or technically easy to clone miRNAs might overshadow the miRNAs that are only expressed at particular developmental stages, under specific circumstances, in specific cell types or which are otherwise difficult to clone. To overcome such a bias, several computational algorithms have been developed to predict miRNAs in nematode, fly and vertebrate genomes [18, 28, 32-35]. These approaches use genome comparisons to search for evolutionary conserved hairpin structures, which could encode precursor miRNAs (pre-miRNAs). Together, these algorithms predict that there are ~110 miRNAs in the *Drosophila* genome, ~120 to 300 miRNAs in the *C. elegans* genome, and ~255 miRNAs in vertebrate genomes such as human [18, 28, 32-35].

More recently, three independent computational approaches indicate that the actual number of miRNAs in vertebrates might be much higher [9, 10, 36]. We found, by phylogenetic shadowing of primate miRNAs, that there is a high degree of conservation in miRNA regions, but relatively low conservation in sequences immediately flanking these regions (see Fig. 3A for example). This feature was used to predict over 800 new miRNAs in mammalian genomes of which many are conserved in other vertebrate genomes. The expression of some (16) of these predicted miRNAs could be confirmed experimentally [9]. Xie and coworkers searched for conserved motifs in 3´ UTRs by systematic comparison of several mammalian genomes. More than half of the recovered motifs correspond with the seed sequences (nucleotide positions 2 to 8) of known miRNAs. The remaining motifs were used to predict 129

new miRNAs [10]. Finally, Bentwich and coworkers used an integrative approach, combining bioinformatic predictions with microarray analysis and sequence-directed cloning, to identify 89 new miRNAs and predict the total number of human miRNAs to be at least 800 [36]. Together, these studies predict that the total number of miRNAs in vertebrate genomes might be doubled and maybe approach as many as a thousand [9]. These candidate miRNAs require further experimental verification.

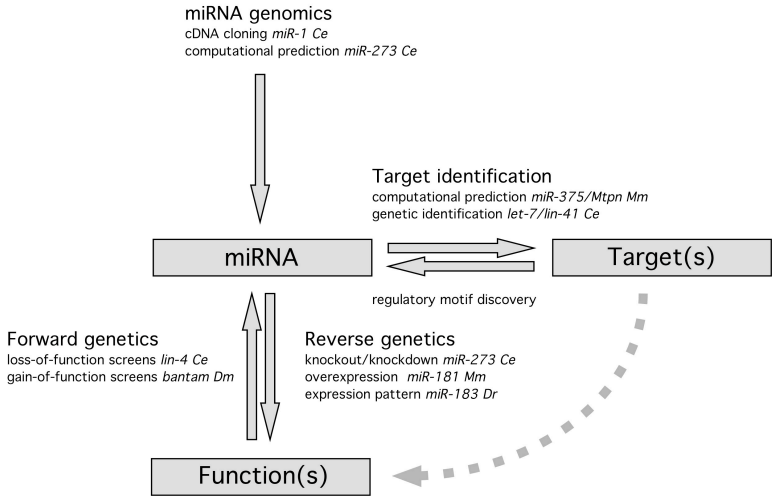


Fig. 2. Approaches for the identification of miRNA genes, target genes and functions of miRNAs in animals. For each approach an example miRNA is given that is mentioned in the text or literature. Abbreviations; Ce, *Caenorhabditis elegans*; Dm, *Drosophila melanogaster*; Dr, *Danio rerio* (zebrafish); Hs, *Homo Sapiens*; Mm, *Mus musculus*.

2.2. Expression of miRNAs

The distribution of miRNA genes within genomes is not random. More than half of known mammalian miRNA genes are within introns of host genes [37]. Microarray analysis indicates that many of these miRNAs are coexpressed with their host genes [38]. The location of some of these intronic miRNAs is evolutionarily conserved and they are similarly coexpressed with their host genes in different animals. For example, the *mir-126* gene is located in an intron of the *EGFL7* gene of mouse, human [37] and zebrafish. miR-126 and *EGFL7* are similarly expressed in endothelial cells of the heart and blood vessels in both humans and zebrafish [38, 39]. Such homologies suggest important and evolutionarily conserved roles for these miRNAs. Approximately 40% of human miRNA genes are in genomic clusters. Many of these miRNAs are also coexpressed [38], which could have biological advantages. Finally, miRNA genes are frequently located at fragile sites and genomic regions involved in cancers [40], suggesting that these miRNAs are involved in cellular processes impaired in cancers, such as cell growth, cell division and proliferation.

miRNAs have a wide variety of expression patterns. In *C. elegans* and *Drosophila* some miRNAs are differentially expressed in time during development, whereas others seem to be more ubiquitously expressed [5-8, 18, 19, 32, 35, 41]. For most of these miRNAs only the temporal expression patterns have been determined. However, reporter constructs show that some miRNAs are only expressed in specific cell types [14, 42].

The tissue-specific miRNA expression has been best studied in vertebrates. First, the cloning frequency from particular tissues or cells is different for many miRNAs. For example, several miRNAs are predominantly cloned from mouse heart, liver or brain tissues [21], embryonic stem cells [25, 26] or pancreatic islet cells [43]. Further

analysis by northern blots or microarrays reveals that many other vertebrate miRNAs are also tissue-specifically expressed [38, 39, 44-48]. Our recent data from *in situ* hybridizations in zebrafish embryos indicate that ~80% of the conserved vertebrate miRNAs that are expressed during embryonic development are tissue-specific (e.g. see Fig. 3B). This tissue-specificity is not restricted to only a few major organs, but virtually all zebrafish tissues and even individual cell types within tissues have specific expression of one or a few miRNAs. For example, miR-183 is specifically expressed in the hair cells of sensory epithelia [39].

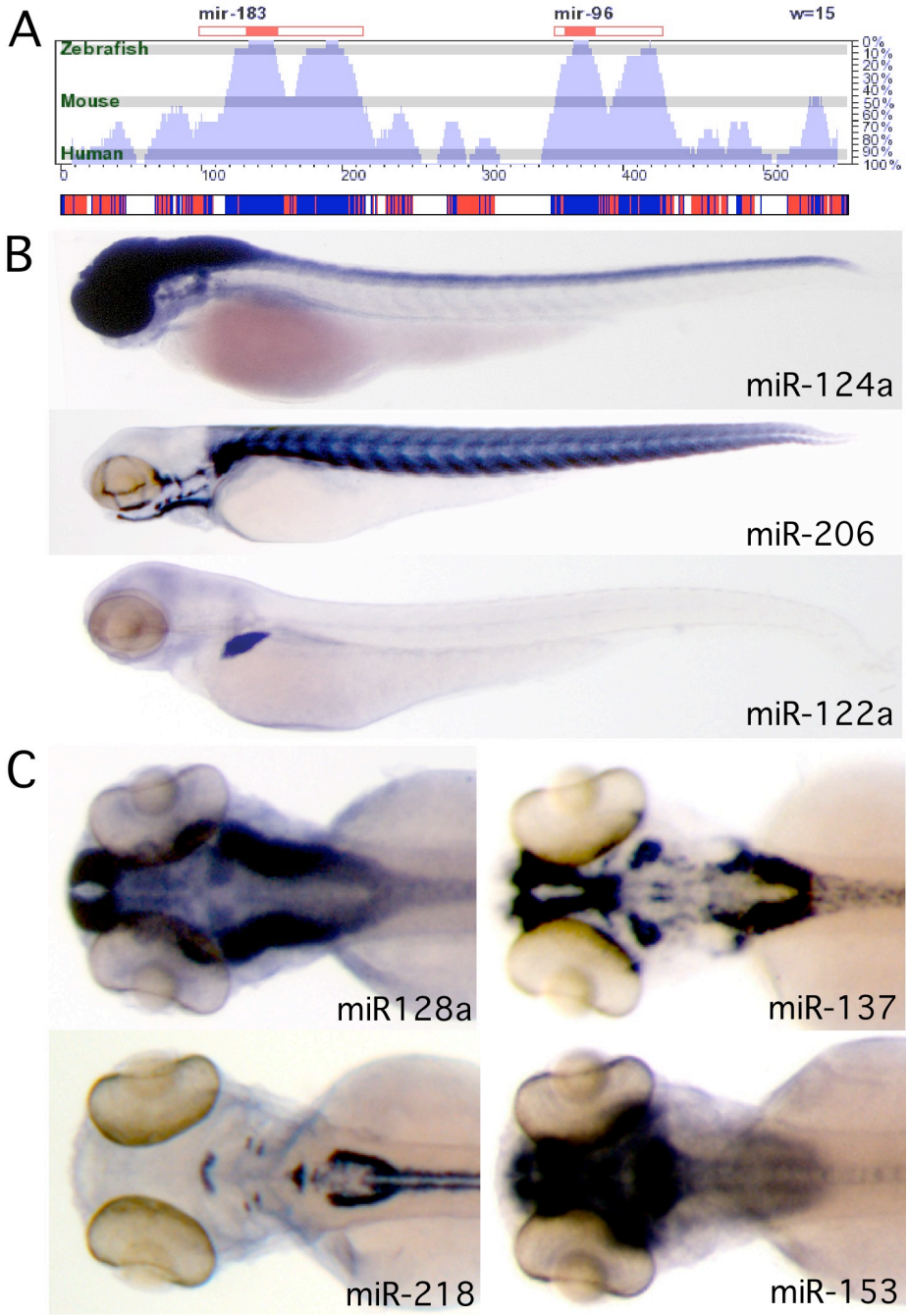


Fig. 3. Vertebrate miRNAs in zebrafish. (A) Phylogentic conservation of two clustered vertebrate miRNA genes. miRNA regions (indicated by the red boxes) are more conserved than flanking sequences. See [9] for details. (B) Tissue-specific miRNA expression in brain, muscles and liver in zebrafish embryos. (C) Example expression patterns of conserved brain-specific miRNAs in zebrafish embryos. Adapted with permission from [39].

Many miRNAs are similarly expressed in mammals and zebrafish, indicating an evolutionarily conserved function. For example, many miRNAs that are brain-specific in mammals have very distinct *in situ* expression patterns in the brain of zebrafish embryos (e.g. see Fig. 3C [39]). In addition, these miRNAs are also differentially expressed during mammalian brain development [24, 49], suggesting a conserved role in vertebrate brain function.

3. miRNA biogenesis

The mammalian miRNA biogenesis pathway (and that of other animals) can be divided into multiple steps (Fig. 4). During these steps, long primary miRNA transcripts are processed into short functional mature miRNAs, the effector molecules of miRNA-induced gene silencing.

3.1. Transcription

Most, if not all, miRNA genes are transcribed by RNA polymerase II into long primary miRNAs (pri-miRNAs) that contain 5´ caps and poly(A) tails [50, 51]. This is in contrast with transcription of most other small RNA molecules, such as tRNAs and U6 snRNAs, which are transcribed as short RNAs by RNA polymerase III. The pri-miRNA transcripts may encode one or several clustered precursor miRNAs (pre-miRNAs) that form imperfect secondary stem-loop structures (pri-miRNA hairpins). The future mature miRNA resides in one of the arms of the stem sequences (Fig. 1). The pri-miRNA hairpins can be located in both exonic and intronic sequences of the transcript. The pri-miRNA transcripts can also simultaneously code for proteins. For example, miRNA sequences have been found adjacent to mRNA sequences in expressed sequence tags (ESTs) [7, 52] and chimeric transcripts can produce both miRNAs and proteins [51]. After transcription, the maturation of pri-miRNAs into mature miRNA is a stepwise and compartmentalized process [53].

3.2. Nuclear pri-miRNA processing

pri-miRNAs are processed in the nucleus into ~70-80 nucleotide long pre-miRNAs. The nuclear processing of pri-miRNAs into pre-miRNAs is performed by the microprocessor complex, a multiprotein complex that contains the enzymes Drosha [54] and DGCR8/Pasha [55-58]. Drosha is a member of the RNase III family proteins in eukaryotes that is involved in preribosomal RNA processing [59, 60] and is the core nuclease that mediates the initial cleavage of the pri-miRNA [54]. Drosha contains a double-stranded RNA (dsRNA)-binding domain and two RNase III domains. The latter interact with each other to make up one processing center that generates pre-miRNAs with two-nucleotide 3´ overhangs [54, 57, 61]. In the microprocessor complex, Drosha interacts with the dsRNA-binding protein DGCR8 in mammals or Pasha in *Drosophila*. DGCR8 is a protein that is deleted in DiGeorge syndrome. Both Drosha and DGCR8/Pasha are necessary and sufficient for processing of pri-miRNAs [55-58]. The cleavage sites in the pri-miRNA hairpin are located approximately two helix turns from the junction between the distal loop and the stem of the hairpin [62]. Efficient processing further requires strong base-pairing at the cleavage sites [54], a stem extension of approximately one helix turn [62], single-stranded RNA extensions outside the pri-miRNA hairpin [63] and a large (≥10 nt) terminal loop sequence [62]. The Drosha-cleaved pre-miRNA hairpins are exported from the nucleus by Exportin-5 in the presence of the Ran guanosine triphosphate as cofactor. Probably, Exportin-5 exports only bonafide pre-miRNAs by recognizing the two-nucleotide 3´ overhangs [64-66].

3.3. Cytoplasmic pre-miRNA processing

In the cytoplasm, the pre-miRNAs are processed into ~22-nucleotide duplex miRNAs by another RNase III enzyme, Dicer [53, 67-69]. Dicer was originally discovered by its role in RNA interference (RNAi), where it processes long double-stranded RNA into small interfering RNAs (siRNAs) that mediate RNAi [70-72]. Dicer is also required for development of *C. elegans*, *Drosophila* and vertebrates [68, 69, 73-76], and for processing of pre-miRNAs during development. Dicer cleaves pre-miRNAs approximately two helix turns away from the pre-miRNA ends, before the loop sequence. These cuts release the ~22-nucleotide duplex miRNAs, which contain two-nucleotide 3´ overhangs on both ends, similar as siRNAs.

Like Drosha, Dicer contains two RNase III domains, which can form one processing center through intramolecular dimerization that generates two-nucleotide 3´ overhangs [77]. In addition, most Dicer proteins contain a DEAD-box RNA helicase domain, a domain with unknown function, a dsRNA-binding domain and a Piwi/Argonaute/Zwille (PAZ) domain. Critical residues of the PAZ domain are conserved between Dicer and Argonaute proteins. The Dicer PAZ domain probably recognizes the two-nucleotide 3´ overhangs of the pre-miRNA hairpins and thereby orientates the processing center correctly.

Dicer also interacts with dsRNA-binding domain proteins. In mammals, Dicer forms a complex with the dsRNA-binding protein TRBP (human immunodeficiency virus transactivating response RNA-binding protein). TRBP is essential for pre-miRNA processing by Dicer and RISC-mediated gene silencing. Furthermore, the Dicer-TRBP complex associates with the RISC component Argonaute 2, thereby recruiting Argonaute 2 to siRNAs bound by Dicer [78]. Mice with TRBP deletions usually die at the time of weaning, but occasional males survive. However, these males are sterile and have defects in spermatogenesis [79]. In *C. elegans*, Dicer associates with the dsRNA-binding protein RDE-4, which is required for RNAi. RDE-4 also interacts with the *C. elegans* Argonaute protein RDE-1 and a conserved DExH-box RNA helicase [80].

Drosophila has two Dicers, Dicer-1 and Dicer-2, which both lack different domains and have different functions. Dicer-1 lacks a functional helicase domain. However, Dicer-1 is essential for processing of pre-miRNAs and miRNA-induced translational repression [76]. Mutations in Dicer-1 give severe developmental defects [76] and abnormal stem cell division in the germline [81]. Dicer-2 lacks the PAZ domain. Dicer-2 is not required for miRNA biogenesis but is needed for the processing of siRNA precursors. Furthermore, it is dispensable for normal development [76]. Together, this suggests that the processing of pre-miRNAs by Dicer-1 is essential for *Drosophila* development. Nevertheless, both Dicer-1 and Dicer-2 are required for siRNA-directed mRNA cleavage [76].

Like mammalian Dicer, both *Drosophila* Dicer-1 and Dicer-2 are also in complexes with dsRNA-binding proteins, which are homologous to the mammalian TRBP and *C. elegans* RDE-4. However, each Dicer forms a complex with different dsRNA-binding proteins. Dicer-1 interacts with Loquacious (Loqs). Loqs is required for pre-miRNA processing by Dicer-1 and miRNA-directed silencing of a reporter gene [82-84]. Both Dicer-1 and Loqs are present in an Argonaute1-containing complex that has pre-miRNA-processing activity [83, 84], but not in an Argonaute2 complex that does not have this activity [83]. However, like Dicer-1, Loqs is also required for siRNA-induced silencing processes *in vivo* [82]. Removal of Loqs results in severe sterility in both males and females, presumably through a failure to maintain germline stem cells [82, 84]. Dicer-2 forms a complex with the dsRNA-binding domain R2D2. Unlike the other dsRNA-binding proteins, RdD2 is not essential for Dicer-2 function, but enhances siRNA-induced mRNA cleavage by the RISC complex [85].

3.4. miRNA and siRNA asymmetry

Little is known about the mechanism that unwinds miRNA and siRNA duplexes and the proteins performing this process. However, these duplexes are usually asymmetric in terms of internal thermodynamic stability. Unwinding of the duplexes starts at the ends with the lowest free energies. The miRNA strand that has its 5´ terminus at this end is the future mature miRNA (also called guide RNA) and is incorporated into RISC [86, 87]. The other strand, the miRNA* (star), is usually rapidly degraded. Occasionally, for more symmetric miRNA duplexes, it also remains stable. In *Drosophila*, strand selection of siRNA duplexes is mediated by the Dicer-2/R2D2 complex, in which R2D2 senses the most stable end of the siRNA duplex, thereby orienting the heterodimer and initiating the siRNA unwinding [85, 88]. Unwinding of siRNA duplexes is ATP-dependent [89]. However, complete unwinding of a siRNA duplex is not necessary for RNAi activity *in vivo* [90]. Recently, a reversible inhibitor of siRNA duplex unwinding has been found. This inhibitor could give more insight on the mechanism of unwinding [91]. The mechanism for unwinding of miRNA duplexes is not known either, but likely involves other dsRNA-binding partners of Dicer.

3.5. The RNA-induced silencing complex (RISC)

During or after unwinding, the mature miRNAs are loaded into a ribonucleoprotein complex, miRNP [23, 92], which is similar, but not necessarily identical, to RISC. RISC was originally identified as the siRNA-containing effector of RNAi that mediates degradation of the targeted mRNAs [93]. RISC may be present as several different complexes [94], which contain different proteins. The protein-components of miRNA-induced RISC (miRISC) and siRNA-induced RISC (siRISC) are similar. Furthermore, miRNAs can enter the siRISC and act as siRNAs [92] and siRNAs can act also as miRNAs [95, 96]. The only determinant is the sequence of the small RNA it contains in combination with an Argonaute family protein.

Argonaute proteins are prominently present in RISC [97]. Argonaute proteins are members of the PPD family proteins (also called the Argonaute family), which contain PAZ and PIWI domains. PPD family proteins are involved in several RNAi-related mechanisms in different organisms [68, 98-103]. Based on their sequence, the PPD family proteins can be classified into two subfamilies: the PIWI subfamily and the Argonaute (eIF2C) subfamily [104, 105]. Mammals have several PPD family protein members, belonging to either the PIWI subfamily or the Argonaute subfamily. Whereas mammalian PIWI proteins are specifically expressed in the testis, mammalian Argonautes are expressed in a wide variety of tissues [104, 105]. Members of both subfamilies are associated with Dicer [97, 105]. All mammalian Argonaute proteins bind mature miRNAs, but have different functions [23, 106-108]. In both mammals and *Drosophila*, Argonaute2 (eIF2C2) is the sole catalytic engine of RNAi that mediates the targeted RNA cleavage [97, 107-111]. The role for the other mammalian Argonautes is not yet known, but they could, for example, be involved in chromatin modification and heterochromatin silencing similar as in plants and *Drosophila* [112, 113]. In *Drosophila*, Argonaute1 is required for efficient RNAi [114], miRNA production or stability and miRNA-induced RNA cleavage [115]. Furthermore, *Drosophila* embryos that are defective in Argonaute1 show developmental defects and malformations of the nervous system [116].

The PAZ domain is conserved in Piwi, Argonaute and Zwille proteins. Structural studies indicate that the PAZ domain recognizes the two-nucleotide 3´ overhangs of duplex miRNAs and siRNAs [117-121]. The PAZ domain could thereby serve as an anchoring site for the 3´ end of the guide miRNA or siRNA and subsequently orientate the guide RNA in RISC. The PIWI domain of Argonaute2 is the catalytic domain that is involved in RNA cleavage. It is similar to RNase H and Endonuclease V [108-110, 122]. The 5´ nucleotide of the guide RNA in RISC is unpaired from the target RNA. The PIWI domain recognizes this nucleotide, which then serves as an anchor-

ing site that determines the cleavage site in the target RNA [109, 122-124]. A subregion of the PIWI domain, the PIWI-box, binds to Dicer [125, 126].

Other components that are associated with RISC activity in *Drosophila*, *C. elegans* and mammalian cells are fragile X mental retardation protein (FMRP), Tudor-SN (tudor staphylococcal nuclease) and VIG (Vasa intronic gene) [23, 127-130]. FMRP physically interacts with Dicer, Argonaute1/eIF2C1 and miRNAs [127, 129, 130]. Tudor-SN binds to hyper-edited dsRNA and promotes its cleavage [131]. The role for VIG is unknown, but it might be involved in miRNA function [128]. In mammals, the miRNA-programmed RISC further contains Gemin3 (a DEAD-box RNA helicase) and Gemin4 (a protein with unknown function) [23]. In *Drosohila*, RISC also contains the p68 RNA helicase (Dmp68) [129] and Armitage, which probably has a function during the incorporation of single-stranded siRNAs into RISC [132].

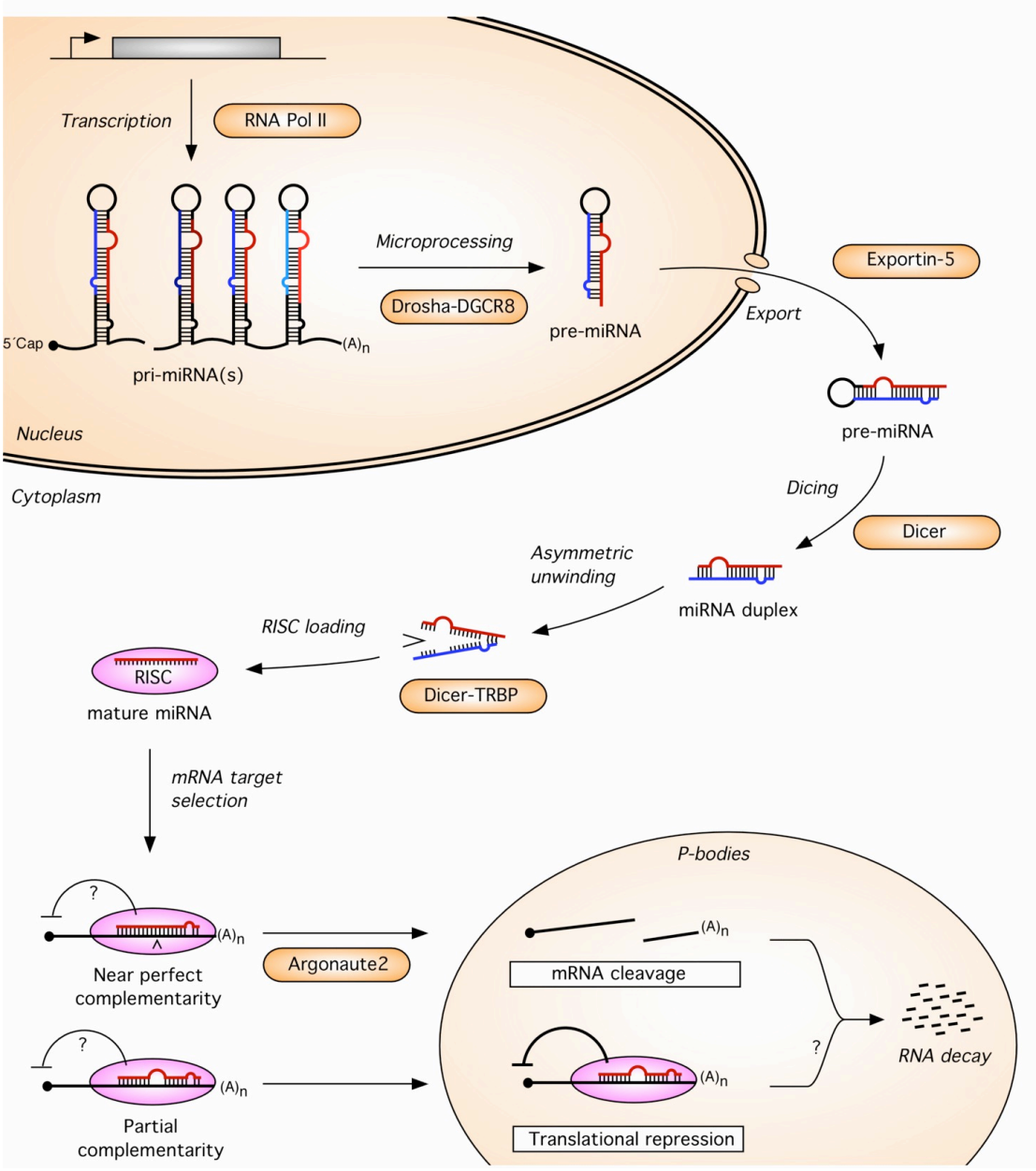


Fig. 4. A model for the miRNA biogenesis pathway and the miRNA silencing mechanisms in mammals. See text for details.

3.6. miRNA target selection

The most important determinant for target selection in animals (and plants) is the sequence that spans nucleotides two to eight in the 5´ parts of miRNAs, the seed sequences. Introduction of mismatches that disrupt pairing between the seed sequence and complementary target sites results in reduced miRNA function [133-137]. Furthermore, bases of the seed sequence contribute disproportionately to the target RNA-binding energy [138]. Compared to other parts of the miRNAs, the seed sequence is also phylogenetically the most conserved region in both miRNAs and complementary target sites [10, 11, 32, 136, 139-141].

The first nucleotide positions of miRNA target sites are enriched in adenosines [11], however in miRNAs themselves, this position is usually not conserved and it is also not involved in base-pairing [109, 122-124, 136, 138]. Base-pairs formed by the central and 3´ regions of siRNAs provide a helical geometry required for catalysis [138]. The 3´ ends of miRNAs are probably least required for miRNA function. They are evolutionary poorly conserved and mutations in this part of the miRNA or the complementary target sequences generally do not alter the ability of miRNAs to inhibit translation of reporter constructs [133-135, 137]. However, natural mismatches between the seed sequence and the complementary target site can be complemented by extensive base-pairing between the 3´ end of the miRNA and targets [136]. Furthermore, the 3´ ends of miRNAs are also suggested to be involved in discrimination between different members of a miRNA family, which could explain the extreme evolutionary conservation along the complete sequence for some miRNAs [136].

4. miRNA silencing mechanisms

In RISC, miRNAs can mediate downregulation of target gene activity by two modes: translational inhibition and target mRNA cleavage followed by mRNA decay (Fig. 4). The canonical model is that near-perfect complementarity results in cleavage of the target, whereas partial complementarity causes translational inhibition. The complementarity of the miRNA-target interaction is most likely the only factor that determines which mode occurs, since mismatched siRNAs do not cleave targets but inhibit translation and, conversely, miRNAs that normally only inhibit translation can also cleave perfect matched targets. Furthermore, the functionality of siRNAs and miRNAs is independent of the biogenesis pathway by which they are generated and incorporated into RISC [92, 95, 96, 142, 143]. Most natural miRNA-target interactions contain mismatches and bulges, which presumably prevent the targets from being cleaved and results in translational inhibition [86]. An alternative model to the strict choice between the two regulatory modes is that the modes work consecutively (Fig. 4). In this model, binding of the miRNAs with either partial or near perfect complementary target mRNAs cause translational inhibition of these target mRNAs. Next, the target mRNAs are cleaved by RISC and subsequently degraded if there is sufficient complementarity between the miRNA and the target. In contrast, partial complementarity between the miRNA and its target initially causes translational repression, but might eventually also lead to degradation.

4.1. Translational inhibition

The interaction of the *lin-4* RNA with the imperfect target sites of the *lin-14* mRNA in *C. elegans* leads to a reduction in lin-14 protein levels, but does not alter the mRNA levels. The simplest explanation for this observation is that the *lin-14* mRNA is translationally repressed or that nascent peptides are rapidly degraded [144, 145]. The polysome profiles of both the *lin-14* and *lin-28* target mRNAs are not altered upon the presence of *lin-4* RNA, which suggests that the inhibition is likely to occur after the initiation of translation [145, 146]. In contrast, recent experiments in human

cells show that endogenous *let-7* and the tethering of Argonaute proteins to reporter mRNAs inhibit translation at the initiation phase. Furthermore, cap-dependent translation is repressed by *let-7*, suggesting that the action miRNA-containing RISC involves recognition of the cap [147].

4.2. mRNA cleavage

The mechanism of miRNA-directed cleavage of perfect and near-perfect complementary target mRNAs is probably identical to that of siRNA-induced cleavage. Single-stranded siRNAs in the RISC complex guide the endonucleolytic cleavage of target mRNAs at a single site [106, 148]. A 5′ phosphate is required for single-stranded siRNAs to function [89, 148] and 5′ hydroxyl termini of siRNAs are phosphorylated *in vivo* [90]. The endonucleolytic cleavage reaction occurs via magnesium-dependent hydrolysis of a scissile phosphate that results in the release of a 5′ cleavage product, which contains a 3′-hydroxyl terminus, and a 3′ cleavage product, which contains a 5′-phosphate terminus [149, 150]. The sole catalytic engine of the RISC complex that mediates the cleavage reaction is the Argonaute 2 protein [107-110]. Each miRNA- or siRNA-containing RISC directs multiple rounds of mRNA cleavage [92]. Although the cleavage reaction itself does not require ATP [89, 92], these multiple turnovers are enhanced by ATP [92, 138], suggesting the involvement of an ATP-dependent helicase in release of the cleavage products.

The actual cleavage site in the target mRNA is located in the middle of the region spanned by the guiding siRNA, in between nucleotides 11 and 12. The scissile phosphate is determined by the 5′ end of the guiding siRNA [72, 149, 151]. Modifications of the 3′ ends of siRNAs are tolerated, whereas 5′ modifications are not [90, 106, 148, 151]. However, both the 5′ end and 3′ end of a siRNA are not necessarily required for cleavage reaction itself since perfect base-pairing of thirteen nucleotides, symmetrically surrounding the target scissile phosphate, is sufficient for efficient cleavage [150] and perfect base-paring of nucleotides 2 to 12 of a siRNA with a target still results low level cleavage [138]. Mismatches at the cleavage site can still induce cleavage, although at different scissile phosphate positions, indicating that the target scissile phosphate is not determined by the base-pairing of the siRNA with its target, but is rather intrinsic determinant of the siRNA [138, 150].

After endonucleolytic cleavage, an oligouridine track is added to the 3′ ends of the 5′ cleavage products, which presumably enhances mRNA decay by exonuclease action, starting at the 5′ end [152]. The 5′ cleavage products furthermore contain cap structures and the 3′ cleavage products end in a poly(A) tails, suggesting that decapping and deadenylation occurs after the initial cleavage [153]. In *Drosophila*, the decay of the 5′ cleavage products is mediated by the exosome, a complex of ten or more 3′-to-5′ exonucleases. This decay further requires homologs of the Ski complex proteins [153]. The 3′ cleavage products are degraded by the 5′-to-3′ exonuclease Xrn1 in *Drosophila* and by its homolog XRN4 in *Arabidopsis* [153, 154]. Xrn1p is also required for efficient RNAi in *C. elegans* [155], and XRN4 is linked to RdRp-dependent transgene silencing in *Arabidopsis* through the 3′ cleavage product [156].

Thus far, miRNA-induced mRNA cleavage has only been directly shown in animals for miR-196, which is able to cleave *HOXB8* mRNA in mouse embryos [157]. However, a recent study shows that several mRNAs with complementary target sites are downregulated by overexpression of cognate miRNAs in human cells [158]. Furthermore, in contrast with previous results, the regulation by *let-7* and *lin-4* in *C. elegans* also results in target mRNA degradation [159]. This could be the result of miRNA-induced cleavage, although degradation by general RNA decay is also likely.

4.3. miRNA regulation and P-bodies

Cytoplasmic processing bodies (P-bodies) are the sites where non-translated mRNAs are concentrated and depleted of the translation machinery. Inside P-bodies, mRNAs can be degraded by decapping and 5′ to 3′ decay [160, 161], but mRNAs

within P-bodies can also return to the translation machinery [162]. Efficient RNAi and the decay of mRNAs that are targeted for cleavage in plants and *C. elegans* depend on P-body components [153-155].

Mammalian Argonaute proteins are not randomly distributed in the cytoplasm, but concentrate in these P-bodies. These Argonaute proteins interact with P-body components, such as the decapping enzymes Dcp1a and Dcp1b. Colocalization of Argonaut proteins to P-bodies is dependent on the ability of the Argonaute proteins to bind miRNAs and the presence of RNA. Moreover, reporter mRNAs that are targeted for translational repression by endogenous or exogenous miRNAs become concentrated in P-bodies in a miRNA-dependant manner [147, 163, 164]. There is also a link between miRNAs and P-bodies in *C. elegans*. AIN-1 (a homolog of the candidate human neurological disease protein GW182) localizes to P-bodies and interacts with protein complexes containing and the Argonaute protein ALG-1, Dicer and miRNAs. Furthermore, overexpression of AIN-1 targets ALG-1 to the P-bodies [165].

Together, this indicates that P-bodies facilitate miRNA-induced translational repression and target miRNA degradation by cleavage or general decay. In this respect, it could explain the miRNA-dependent rapid decay of AU-rich elements-containing mRNAs [166]. However, the question remains whether the miRNA-dependent localization of target mRNAs to P-bodies is a cause or consequence of inhibiting protein synthesis [164].

5. Biological functions of miRNAs in animals

miRNAs have diverse biological functions (Table 1). Clues about the function of miRNAs in animals have been obtained by several approaches (Fig. 2). First, several miRNAs were identified by loss- and gain-of-function genetic screens in *C. elegans* and *Drosophila*. Since these miRNAs were found based on their biological functions, their described functions are likely the most accurate. Other genetically assigned miRNA functions come from reverse genetic approaches. These approaches include miRNA knockout or knockdown [167-170] and miRNA overexpression studies. Furthermore, the miRNA expression profiles, determined by, for example, microarray- and *in situ* analysis, have revealed specific miRNA expression patterns, which give hints to the functions of specific miRNAs.

miRNAs function through the regulation of target genes. Only a few (seven) of the miRNA target genes have been identified genetically by forward and reverse genetic screens. Most miRNA target genes come from computational predictions. However, only a handful of these predicted targets have been experimentally validated in *Drosophila* and human cells using reporter constructs and the majority remains to be verified [171, 172]. Recent miRNA target predictions indicate that for each *Drosophila* miRNA there are on average ~100 different target genes of which most are regulated through the seed site interactions [136]. miRNA target predictions in mammals indicate that ~10-30% of the genes might be under control of the currently known miRNAs [10, 11, 173] and that each mammalian miRNA regulates on average ~200 target genes through interaction between the seed sequence and the complementary target site [174]. Many genes have several target sites for either one miRNA or a few different miRNAs. In some cases, multiple target sites are essential for proper regulation *in vitro* or *in vivo* [133, 135, 136, 174, 175]. The combinatorial regulation by different miRNAs adds another layer of complexity that can result in more fine-tuned activity in diverse cell-types, similar to the combinatorial regulation by different transcription factors.

Table 1 Biological functions of miRNAs in animals and disease

miRNA	Target(s)	Function(s)	Refs
Ceanorhabditis elegans			
lin-4	*lin-14, lin-28*	Early developmental timing	[1, 144, 176]
let-7	*lin-41, hbl-1, daf-12, pha-4, ras*	Late developmental timing	[3, 4, 177-180]
lsy-6	*cog-1*	Left/right neuronal asymmetry	[14]
miR-273	*die-1*	Left/right neuronal asymmetry	[181]
Drosophila melanogaster			
bantam	*hid*	Programmed cell death	[15]
miR-14	*Drice?*	Programmed cell death and fat metabolism	[16]
miR-7	Notch targets?	Notch signaling	[141, 183]
Danio rerio			
miR-430	?	Brain morphogenesis	[185]
Mus musculus			
miR-196	*Hoxb8*	Developmental patterning	[157]
miR-181	?	Hematopoietic lineage differentiation	[44]
miR-1	*Hand2*	Cardiomyocyte differentiation and proliferation	[188]
miR-375	*Mtpn*	Insulin secretion	[43]
Human and other vertebrate cell lines			
miR-16	Several	AU-rich element-mediated mRNA instability	[166]
miR-32	Retrovirus *PFV-1*	Anitviral defense	[169]
miR-143	*Erk5?*	Adipocyte differentiation	[189]
SVmiRNAs	*SV40* viral mRNAs	Susceptibility to cytotoxic T cells	[207]
Cancer in humans			
miR-15-miR-16	?	Downregulated in B cell chronic lymphocyte leukemia	[199]
miR-143, miR-145	?	Downregulated in colonic adenocarcinoma	[200]
miR-155/BIC	?	Upregulated in diffuse large B cell lymphoma	[201, 202]
let-7	*Ras*?	Downregulated in lung cell carcinoma	[180, 203]
miR-17-92	?	Upregulated in B-cell lymphoma	[204]

5.1. Developmental timing in worms

The two best-studied miRNAs, *lin-4* and *let-7*, were found by forward genetics to act in the heterochronic pathway to regulate developmental timing in *C. elegans* [1, 3]. *Let-7* was identified as a suppressor of *lin-14* mutants [3], indicating that they act in the same pathway. Loss-of-function mutations in *lin-4* and *let-7* both cause retarded development but at different developmental stages. Whereas *lin-4* null mutants reiterate specific fates of the first larval stage at subsequent later stages, *let-7* null mutants reiterate larval cell fates at the adult stage. Conversely, overexpression of *let-7* gives precocious development, the opposite heterochronic phenotype [3]. The timing of the phenotypes corresponds with the onset of *lin-4* and *let-7* expression, which is early and late in development, respectively [1, 3]. Thus, *lin-4* is an early developmental timer and *let-7* a late developmental timer.

Two *lin-4* targets, *lin-14* and *lin-28*, were both identified as genes that act downstream of *lin-4* in the heterochronic pathway. Loss-of-function mutations in these genes cause phenotypes that are opposite to that of *lin-4* mutants: precocious cell fates of late developmental stages during the first and second larval stages [2, 176]. Furthermore, gain-of-function mutations in the *lin-14* 3′ UTR cause a phenotype similar to *lin-4* loss-of-function mutations [144]. *Lin-14* encodes a nuclear protein, which regulates the transition from the first to the second larval stage [2]. *Lin-28* encodes a cold-shock zinc finger protein that promotes the transition of the second to third larval stage [176]. Regulation by *lin-4* might not be restricted to the early stages since it is also required for downregulation of a *hbl-1::GFP* reporter gene, which has target sites for both *lin-4* and *let-7* in its 3′ UTR, in ventral nerve cord neurons of adults [177].

Multiple genes are regulated by *let-7* in tissue-specific manners. The *lin-41* target gene was identified in a genetic screen as strongest suppressor of the lethality of *let-7* mutants. It encodes a RBCC (ring finger, B box, coiled coil) protein. Null muta-

tions in *lin-41* cause the opposite phenotype to that of the *let-7* mutants, namely precocious expression of adult fates at larval stages [3, 4]. Overexpression of wild-type *lin-41* results in reiteration of larval fates, similar to *let-7* loss-of-function mutants [4]. *Lin-41* is partially redundant with the hunchback-like gene *hbl-1* in the hypodermal seam cells and negatively regulates the timing of the adult specification transcription factor LIN-29 [177, 178]. The defects in *let-7* mutant strains are also suppressed by loss of *hbl-1* function. *Hbl-1* is involved in timing events in neurons of the ventral nerve cord and vulval cells [177, 178]. Several other targets for *let-7* have been identified in an RNAi knockdown screen for genes that have *let-7* target sites and suppress *let-7* mutant lethality. This yielded the nuclear hormone receptor *daf-12*, which is another target in seam cells of the hypodermis and the forkhead transcription factor *pha-4*, which is an intestinal target. It also identified the zinc finger transcription factor *die-1*, the putative chromatin remodeling factor *lss-4* and the RAS ortholog *let-60* as probable targets [179]. In addition, recent experiments indicate that *let-60/ RAS* is regulated by *let-7* in the hypodermal cells and by the *let-7* family member miR-84 in vulval cells [180].

5.2. Neuronal left/right asymmetry in worms

A cascade of genes, involving two miRNAs, determines the left/right asymmetric expression of chemosensory receptor genes in the left (ASEL) and right (ASER) chemosensory neurons of *C. elegans* [14, 181]. The *lsy-6* miRNA is expressed in the ASEL neuron and inhibits the expression of its target, the Nkx-type homeobox gene *cog-1*. This ultimately leads to the expression of the GCY-7 chemosensory receptor in ASEL [14]. In the ASER neuron, miR-273 inhibits the translation of *die-1* mRNA. DIE-1 is a zinc-finger transcription factor needed for the transcription of *lsy-6*. Therefore, the expression of miR-273 leads to the down-regulation of *lsy-6* and subsequently to the expression of the GCY-5 chemosensory receptor in ASER [181]. Thus, inverse and sequential expression of two miRNAs leads to asymmetric expression of chemosensory receptors in neurons of *C. elegans*.

5.3. Programmed cell death in flies

In *Drosophila*, two miRNAs, *bantam* and miR-14, were genetically found to be involved in the regulation of programmed (apoptotic) cell death. The *bantam* miRNA was identified in a gain-of-function EP element insertion screen for genes that affect cell proliferation. Overexpression of *bantam* causes tissue overgrowth and inhibits proliferation-induced apoptosis [15]. In contrast, loss of bantam function is lethal. *Bantam* negatively regulates Hid protein expression *in vivo* and thereby blocks *hid*-induced apoptostis in the eye [15]. miR-14 was identified by a genetic screen for inhibitors of programmed cell death. Animals with miR-14 loss-of-function alleles have enhanced Reaper-induced cell death, whereas overexpression of miR-14 suppresses cell death. The absence of miR-14 is also characterized by semilethality, reduced lifespan and stress sensitivity. Furthermore, miR-14 is involved in the regulation of fat metabolism in a dose-dependent manner [16]. A potential target for miR-14 is the apoptotic effector caspase Drice. Drice is upregulated in the absence of miR-14, suggesting that the *drice* mRNA is regulated, either directly or indirectly, by miR-14 [16]. Other miRNAs that potentially act in the programmed cell death pathway are miR-2 and miR-13. They are predicted to regulate the proapoptotic genes *reaper*, *grim* and *sickle* [141]. miR-2 downregulates reporter constructs with the 3´ UTRs of these target genes *in vitro* and *in vivo* [141]. Furthermore, knockdown of miR-2 or miR-13 may give developmental defects [182].

5.4. Notch signaling in flies

Several miRNAs may be involved in regulating the Notch signaling pathway. Notch signaling is essential for proper patterning and development of all multicellular organisms. In *Drosophila*, the Notch target genes, encoding basic helix-loop-helix

(bHLH) repressors and Bearded family proteins, are post-transcriptionally regulated by different combinations of the K-box, GY-box and Brd-box motifs in their 3´ UTRs. These 6-7 nt motifs are evolutionary conserved [141, 183] and are complementary to the seed sequences of several miRNAs [140]. Reporter constructs with the 3´ UTRs of the bHLH and Bearded *Notch* target genes are downregulated by multiple members of miRNA families *in vivo* [141, 183]. The individual K-boxes, GY-boxes or Brd-boxes are necessary and sufficient for this miRNA-mediated regulation. The K-box is regulated by miR-2 and miR-11, the GY-box by miR-7, and the Brd-box by miR-4 and miR-79 [183]. Ectopic overexpression of miR-7 or a cluster of K-box-regulating miRNAs leads to reduced expression of downstream notch targets such as Cut. Furthermore, it causes notching of the wing margin, reduced wing vein spacing, thickening of the wing veins and ectopic microchaete bristles [141, 183]. These phenotypes are characteristic of reduced Notch signaling, supporting the idea that the Notch signaling pathway is regulated by miRNAs.

5.5. Early vertebrate development

Several observations show that miRNAs are essential for the normal development of mammals. First, mouse and human ES cells express a specific set of miRNAs that are downregulated upon differentiation into embryoid bodies [25, 26]. Second, ES cells that are deficient in *dicer* are viable, but do not form mature miRNAs and they fail to differentiate *in vitro* and *in vivo* [184]. Third, *dicer* mutant mouse embryos die before axis formation during gastrulation and have ES cell loss [73].

To study the global role of miRNAs in early embryonic development we knocked out the *dicer* gene in zebrafish [74]. Zebrafish embryos that are *dicer* mutant develop normally but arrest ~ 8 days after fertilization when they run out of maternal Dicer and are depleted of mature miRNAs. At this point all the major organs have been formed. This arrest suggests that miRNAs are essential for development and growth of tissues beyond this stage. Knockdown of maternal Dicer mRNA by morpholinos results in an earlier arrest and delay in miRNA production. However during the first 24 hrs these embryos still develop quite normally, which indicates that miRNAs are not essential for earliest developmental processes [74]. To also exclude a role for maternal Dicer protein Giraldez and coworkers generated maternal-zygotic mutants from the *dicer* knockout zebrafish [185]. As expected, these maternal-zygotic *dicer* mutants do not process pre-miRNAs into mature miRNAs. Surprisingly, they have only some mild defects during early development. They have intact axis formation and cell regionalization and differentiate into multiple cell types and tissues, indicating correct patterning. However, they show morphogenesis defects during gastrulation, brain formation, neural differentiation, somitogenesis and heart development. Together, this implies that miRNAs are not essential for cell fate determination and early patterning, but are essential for subsequent later steps in embryonic development [185]. This notion is further supported by the temporal and spatial expression patterns of conserved vertebrate miRNAs in zebrafish embryos [29, 39]. The majority of these miRNAs are not expressed early, but show highly tissue-specific expression during the later stages [29, 39], suggesting that their role is not in tissue fate establishment but in differentiation or maintenance of tissue identity [29, 39]. Interestingly, *dicer* mutant primordial germ cells give rise to viable germ lines in both male and female zebrafish, which can form embryos. This indicates that mature miRNAs are also not cell-autonomously required in gametes during germ-line development in zebrafish [185].

Why do *dicer* mutant mouse embryos without mature miRNAs die before axis formation while zebrafish *dicer* mutant embryos without mature miRNAs survive that stage? This may be explained by the differences in dynamics during early embryonic development. Compared to mouse development zebrafish early development is extremely rapid. One day after fertilization mouse embryos have undergone cell divi-

sion once. During the same period zebrafish embryos have established a basic body plan and have formed all major organs. Maybe most miRNAs are not produced fast enough, miRNA levels are to low or miRNA action is to slow to function in the rapid dividing cells of early zebrafish embryos. Alternatively, other RNAi-related mechanisms that are controlled by Dicer, such as the formation of heterochromatin structures and centromeric silencing [184, 186], may have a function in early development of mice, but not in zebrafish.

The only miRNAs that are highly expressed during early zebrafish development are members of the miR-430 miRNA family. The miR-430 miRNA family is expressed from a genomic cluster of more than 90 copies within 120 kb, is conserved in other fish species and related miRNAs are present in mammals and frogs [27, 29, 185]. In frogs, these related miRNAs are also expressed early in development [27]. Injection of miR-430 into maternal-zygotic *dicer* mutant zebrafish embryos rescues the brain morphogenesis defects and to some extend the other neuronal defects, indicating that the miR-430 miRNA family regulates neurogenesis in zebrafish [185].

5.6 Late vertebrate development

Some other miRNAs have a more specialized function at later stages of vertebrate development (Table 1). First, miR-196 is involved in HOX gene regulation. The *mir-10* and *mir-196* genes of various vertebrates reside in the homeobox (HOX) clusters. Like the HOX genes, miR-10 and miR-196 are colinearly expressed in time and space along the anterioposterior (head-to-tail) body axis [39, 187]. In 15-day old mouse embryos, miR-196 directs cleavage of *HOXB8* mRNA and probably also inhibits the *HOXC8*, *HOXD8* and *HOXA7* genes [157]. The expression of miR-196 is somewhat more posterior than the target HOX genes [187], which probably helps to define the posterior expression-boundary of the target HOX genes [157]. miR-10 may have a similar function in the HOX cluster since target sites for mir-10 have also been predicted in HOX genes, for example in *HOXA3* [11]. Second, the muscle-specific miR-1 regulates the balance between differentiation and proliferation of cardiomyocytes during heart development in mice [188]. Overexpression of miR-1 in the heart of transgenic mice results in a proliferation defect and failure of ventricular cardiomyocyte expansion, indicative of premature differentiation of cardiomyocytes. Hand2, a transcription factor that promotes ventricular cardiomyocyte expansion is a target of miR-1. Hand2 is downregulated by overexpression of miR-1 *in vivo*, suggesting that it is a true target during mouse heart development [188]. Third, the miR-181 miRNA modulates hematopoietic lineage differentiation in mice. miR-181 is preferentially expressed in B-lymphocytes of mouse bone marrow and the thymus, the primary lymphoid organ [39, 44]. Ectopic overexpression of miR-181 increases the fraction of B-lymphocytes and decreases the fraction of T-lymphocytes *in vitro* and *in vivo* in mice. Although the target for miR-181 is not known, this indicates that miR-181 regulates mouse hematopoietic lineage differentiation [44]. Fourth, miR-143 regulates human adipocyte differentiation. miR-143 is strongly expressed in adipose (fat) tissue and is upregulated during the differentiation of human pre-adipocytes into adipocytes. Knockdown of miR-143 prevents adipocyte-specific gene expression and the accumulation of triglycerides but increases ERK5 protein levels [189]. ERK5 is a predicted target gene of miR-143 [139]. These data suggest that miR-143 is normally involved in promoting adipocyte differentiation or functioning, possibly through the regulation of ERK5 protein levels [189]. Finally, some miRNAs may have a role in imprinting during development. The *mir-127* and *mir-136* genes reside in the human imprinted 14q32 domain and are expressed from the maternally inherited chromosome, in the antisense orientation to a retrotransposon-like gene (*Rtl1*), which is exclusively expressed from the paternal chromosome. Lack of maternal gene expression in this region has been associated with abnormal development in mouse and humans [190, 191].

5.7. Physiological functions of vertebrate miRNAs

Several miRNAs do not have an obvious role in vertebrate development, but rather act in diverse physiological and cellular processes (Table 1). The miR-375 miRNA is specifically expressed in murine pancreatic islets cells, where it regulates the *Myotrophin* (*Mtpn*) gene and thereby glucose-stimulated insulin exocytosis [43]. Recently, it has been shown that miR-375 may act synergistically with miR-124a and let-7b [174]. Besides expression in the pancreatic islets, miR-375 is also highly expressed in the pituitary gland of zebrafish embryos [39]. This may indicate a role for miR-375 in the secretion of other hormones or neuroendocrine products, perhaps by regulating *Mtpn* in the pituitary gland. A role for miR-122a is the exclusion of cationic amino acid transporter (CAT-1) protein from the liver through translation inhibition and mRNA cleavage [192]. miR-122a is highly expressed in adult livers and its expression is upregulated during mammalian liver development. The expression pattern of CAT-1, a target of miR-122a, is inversely correlated with miR-122a expression and thus absent from the liver [21, 192]. The functional consequence of this absence is however unknown. miR-122a may also have a role in mouse spermatogenesis. miR-122a is differently expressed in mouse testis during spermatogenesis and here probably regulates the target gene Transition protein 2 (*Tnp2)*. *Tnp2* is a post-transcriptionally regulated testis-specific gene that is involved in chromatin remodeling during mouse spermatogenesis [193]. miRNA knockdown experiments indicate that several other miRNAs have roles in cell proliferation, cell growth and cell death in cell cultures [170, 194]. For example, inhibition of miR-125b results in decreased proliferation of differentiated cancer cells [170]. Recently, it has been shown that miR-16 acts together with AU-rich element (ARE)-binding proteins in the degradation of ARE-containing mRNAs [166]. AREs are located in the 3´ UTR of various short-lived mRNAs such as those encoding cytokines and proto-oncogenes. Rapid decay of reporter constructs containing these AREs requires proteins of the miRNA machinery. The human miR-16 has only limited complementarity to AREs. However, in the presence of the ARE binding protein, tristetraprolin (TTP), this limited complementarity is enough to destabilize ARE containing mRNAs. Knockdown and overexpression of miR-16 inhibits and promotes degradation of ARE containing mRNAs, respectively [166]. The miR-16-dependent regulation of ARE mRNAs by ARE binding proteins suggests that we might expect several miRNAs to have a variety of potential activities as specificity factors for the interaction of mRNA regulatory proteins.

5.8. General function of miRNAs

Thus far miRNAs have only been identified in multicellular organisms and are notably absent from unicellular organisms. This absence indicates that miRNAs, in general, may be essential for organisms to differentiate into multiple cell- and tissue types and/or to keep cells in a particular differentiation state. Several recent observations suggest that miRNAs are indeed such regulators of differentiation. First, undifferentiated or poorly differentiated cells do not require miRNAs to survive. This is apparent from the fact that mouse ES cells that do not form miRNAs are viable but fail to differentiate [184] and zebrafish germ cells do not need miRNAs for their survival and contribution to the germ line [185]. Second, most miRNAs are not expressed in early zebrafish development when cells are undifferentiated [29, 39], but have highly tissue-specific expression at later stages when the most of cell types have been formed [39]. Third, in agreement with this absence, miRNAs are also not essential for tissue fate establishment during early zebrafish development, but are essential for later developmental steps [185] and tissue growth and/or functioning [74]. Fourth, many types of human cancer cells have reduced miRNA expression compared to their fully differentiated tissue of origin [195]. Fifth, vertebrate animals encode hundreds different miRNAs that are expected to regulate up to 30% of the genes [10, 11]. Many of these miRNAs are widely conserved. Sixth, miRNAs have a high molecular abundance per cell. In *C. elegans* some miRNAs have been estimated to be

present in as many as 50,000 copies per cell, which is approximately 500-fold higher than the level of a typical worm mRNA [32]. Finally, transfection of the muscle-specific miR-1 and the brain-specific miRNA miR-124 into human HeLa cells shift the mRNA expression profile to that of muscle and brain cells, respectively [158].

From an evolutionary point of view, the regulation of gene expression by miRNAs may be an easy and flexible innovation, nevertheless crucial for cellular differentiation. The small size and simplicity of miRNA genes suggests that they can arise relatively easily and frequently *de novo* in animal and plants genomes. The independent creation of miRNAs is reflected by the high diversity between miRNA families and the lack of overlap between plant and animal miRNAs [196, 197]. Furthermore, a few DNA base-pair changes corresponding to the seed sequences of miRNAs are likely to considerably alter the repertoire of target genes. In parallel, target genes can easily be subsumed under the control of miRNAs through changes of only a few base-pairs in the DNA corresponding to their 3´ UTRs, similar to the acquisition of *cis*-regulatory sites in the promoters of genes. Single miRNAs are able to regulate about hundred target genes [158] and the combinatorial action of miRNAs is expected to regulate the expression of thousands of mRNAs [196]. The dampening of such a number of genes might be essential for the initiation or maintenance of tissue differentiation [39, 196]. In addition, the post-transcriptional regulation of gene expression might also dampen fluctuations in gene expression. This dampening results in stable protein levels, which might also be required to keep cells differentiated.

6. miRNAs and disease

The high number of miRNA genes, the diverse expression patterns and the abundance of potential miRNA targets suggest that miRNAs are likely to be involved in a broad spectrum of human diseases. In addition, components required for miRNA processing and/or function have also been implicated in fragile X mental retardation [127], DiGeorge syndrome [56-58] and cancer [198], pointing at a general involvement of miRNAs in disease.

6.1. miRNAs and cancer

More than half of the human miRNA genes are located at sites known to be involved in cancers, such as fragile sites, minimal regions of loss of heterozygosity, minimal regions of amplification, or common breakpoint regions. Such locations suggest that some miRNAs are involved in tumorigenesis [40]. miRNA expression profiles of a large number of human tumor samples show that miRNAs are sometimes upregulated [40], but generally downregulated in tumors [195]. Furthermore, nearly all miRNAs are differentially expressed in different cancers. Since the miRNA profiles reflect the developmental lineage and differentiation state of tumors, these profiles can be used to classify poorly differentiated tumors [195]. The roles for some miRNAs have been investigated in more detail (Table 1).

The cluster of the *mir-15* and *mir-16* genes lies in a region that is deleted in more than half of B cell chronic lymphocyte leukemias and expression of miR-15 and miR-16 is downregulated in the majority of these leukemias [199]. The *mir-143* and *mir-145* genes also reside in a genomic cluster. Expression of both miRNAs is downregulated in colon cancer tissue as well as in several cell lines derived from other types of cancers [200]. The *mir-155* gene lies in the non-coding *BIC* RNA transcript. Both the expression of *BIC* RNA and miR-155 are upregulated in several types of lymphomas, especially in diffuse large B cell lymphomas [201, 202]. Patients with activated B cell lymphomas have worse prognosis than patients with germinal center lymphomas, which is reflected by the absolute levels of miR-155 in these cancers [202]. The human *let-7* miRNA is downregulated in several lung cancers and is able to inhibit growth of lung cancer cells *in vitro* [180, 203]. In addition, *let-7* levels are

lowest in patients with the lowest postoperative survival; thus, *let-7* is also a good indicator for the prognosis of these types of cancers. The *mir-17-92* cluster is located in a region that is often amplified in human B-cell lymphomas and miRNA levels from this cluster are increased in B-cell lymphomas. Overexpression of the *mir-17-92* cluster prevents apoptosis and accelerates the formation of c-Myc-induced B-cell lymphomas in mouse models [204]. Interestingly, c-Myc itself also upregulates the expression of miRNAs from the *mir-17-92* cluster and the cell cycle promoter E2F1, a target for some of these miRNAs, resulting in a tightly controlled proliferative signal [205].

6.2. miRNAs and viruses

miRNAs have been identified in several mammalian viruses, including many members of the herpesvirus family [30], such as Epstein-Barr virus [31] and Kaposi sarcoma-associated virus [206], simian virus 40 [207] and possibly the human immunodeficiency virus [208]. For most of these miRNAs the functions are unknown, but they are expected to regulate the expression of viral and host genes for their survival and propagation in infected cells. For example, two miRNAs encoded by simian virus 40 target early viral mRNAs for cleavage. This reduces the expression of the viral T antigens, which makes the infected cells less susceptible to cytotoxic T cells and thus enhances the probability of successful infection [207]. Recently it has also been shown that endogenous miRNAs can mediate antiviral defense. The endogenous miR-32 prevents the accumulation of the retrovirus primate foamy virus type 1 (PFV-1) in human cells. However, a PFV-1-encoded protein, Tas, is able to partially suppress the miRNA-induced inhibition of PFV-1 [169]. Together, this indicates that the miRNA machinery is involved various aspects of human viral infections.

7. Conclusions

Since the discovery of *lin-4* and *let-7* it has become apparent that miRNAs form an important and abundant class of post-transcriptional gene regulators, which are widely present in multicellular organisms, ranging from plants to humans. Animals encode hundreds of miRNAs, of which the vast majority have unknown functions. Nevertheless, the limited set of characterized miRNAs indicates that miRNAs can act in diverse biological processes. In addition, mice and zebrafish that are defective in miRNA production or function show that miRNAs are essential for vertebrate development, and are likely to be involved in differentiation and/or maintenance of tissue and cell identity. The current set of miRNAs is predicted to regulate several thousands of target mRNAs, which may go up to 30% of all protein-coding genes [10, 11, 173]. This number might still increase because many additional miRNAs are predicted, even up to a thousand in vertebrate genomes [9]. The verification of the existence of these miRNAs and their interaction with target genes will be the key to find the function of all individual miRNAs during development, disease and other cellular processes.

References

1. Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.
2. Wightman, B., Burglin, T.R., Gatto, J., Arasu, P., and Ruvkun, G. (1991). Negative regulatory sequences in the lin-14 3'-untranslated region are necessary to generate a temporal switch during Caenorhabditis elegans development. Genes Dev *5*, 1813-1824.
3. Reinhart, B.J., et al. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature *403*, 901-906.
4. Slack, F.J., et al. (2000). The lin-41 RBCC gene acts in the C. elegans heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. Mol Cell *5*, 659-669.
5. Pasquinelli, A.E., et al. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature *408*, 86-89.
6. Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science *294*, 858-862.
7. Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. Science *294*, 853-858.
8. Lee, R.C., and Ambros, V. (2001). An extensive class of small RNAs in Caenorhabditis elegans. Science *294*, 862-864.
9. Berezikov, E., et al. (2005). Phylogenetic shadowing and computational identification of human microRNA genes. Cell *120*, 21-24.
10. Xie, X., et al. (2005). Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. Nature *434*, 338-345.
11. Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell *120*, 15-20.
12. Griffiths-Jones, S. (2004). The microRNA Registry. Nucleic Acids Res *32 Database issue*, D109-111.
13. Ambros, V., et al. (2003). A uniform system for microRNA annotation. Rna *9*, 277-279.
14. Johnston, R.J., and Hobert, O. (2003). A microRNA controlling left/right neuronal asymmetry in Caenorhabditis elegans. Nature *426*, 845-849.
15. Brennecke, J., Hipfner, D.R., Stark, A., Russell, R.B., and Cohen, S.M. (2003). bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila. Cell *113*, 25-36.
16. Xu, P., Vernooy, S.Y., Guo, M., and Hay, B.A. (2003). The Drosophila MicroRNA Mir-14 Suppresses Cell Death and Is Required for Normal Fat Metabolism. Curr Biol *13*, 790-795.
17. Lai, E.C. (2003). microRNAs: runts of the genome assert themselves. Curr Biol *13*, R925-936.
18. Ambros, V., Lee, R.C., Lavanway, A., Williams, P.T., and Jewell, D. (2003). MicroRNAs and Other Tiny Endogenous RNAs in C. elegans. Curr Biol *13*, 807-818.
19. Aravin, A.A., et al. (2003). The small RNA profile during Drosophila melanogaster development. Dev Cell *5*, 337-350.
20. Lagos-Quintana, M., Rauhut, R., Meyer, J., Borkhardt, A., and Tuschl, T. (2003). New microRNAs from mouse and human. Rna *9*, 175-179.
21. Lagos-Quintana, M., et al. (2002). Identification of tissue-specific microRNAs from mouse. Curr Biol *12*, 735-739.
22. Kim, J., et al. (2004). Identification of many microRNAs that copurify with polyribosomes in mammalian neurons. Proc Natl Acad Sci U S A *101*, 360-365.
23. Mourelatos, Z., et al. (2002). miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. Genes Dev *16*, 720-728.
24. Miska, E.A., et al. (2004). Microarray analysis of microRNA expression in the developing mammalian brain. Genome Biol *5*, R68.
25. Houbaviy, H.B., Murray, M.F., and Sharp, P.A. (2003). Embryonic stem cell-specific MicroRNAs. Dev Cell *5*, 351-358.
26. Suh, M.R., et al. (2004). Human embryonic stem cells express a unique set of microRNAs. Dev Biol *270*, 488-498.
27. Watanabe, T., et al. (2005). Stage-specific expression of microRNAs during Xenopus development. FEBS Lett *579*, 318-324.
28. Lim, L.P., Glasner, M.E., Yekta, S., Burge, C.B., and Bartel, D.P. (2003). Vertebrate microRNA genes. Science *299*, 1540.
29. Chen, P.Y., et al. (2005). The developmental miRNA profiles of zebrafish as determined by small RNA cloning. Genes Dev *19*, 1288-1293.
30. Pfeffer, S., et al. (2005). Identification of microRNAs of the herpesvirus family. Nat Methods *2*, 269-276.
31. Pfeffer, S., et al. (2004). Identification of virus-encoded microRNAs. Science *304*, 734-736.
32. Lim, L.P., et al. (2003). The microRNAs of Caenorhabditis elegans. Genes Dev *17*, 991-1008.
33. Lai, E.C., Tomancak, P., Williams, R.W., and Rubin, G.M. (2003). Computational identification of Drosophila microRNA genes. Genome Biol *4*, R42.
34. Ohler, U., Yekta, S., Lim, L.P., Bartel, D.P., and Burge, C.B. (2004). Patterns of flanking sequence conservation and a characteristic upstream motif for microRNA gene identification. Rna *10*, 1309-1322.
35. Grad, Y., et al. (2003). Computational and experimental identification of C. elegans microRNAs. Mol Cell *11*, 1253-1263.
36. Bentwich, I., et al. (2005). Identification of hundreds of conserved and nonconserved human microRNAs. Nat Genet *37*, 766-770.
37. Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L., and Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. Genome Res *14*, 1902-1910.

38. Baskerville, S., and Bartel, D.P. (2005). Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. Rna *11*, 241-247.
39. Wienholds, E., et al. (2005). MicroRNA expression in zebrafish embryonic development. Science *309*, 310-311.
40. Calin, G.A., et al. (2004). Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A *101*, 2999-3004.
41. Sempere, L.F., Sokol, N.S., Dubrovsky, E.B., Berger, E.M., and Ambros, V. (2003). Temporal regulation of microRNA expression in Drosophila melanogaster mediated by hormonal signals and broad-Complex gene activity. Dev Biol *259*, 9-18.
42. Johnson, S.M., Lin, S.Y., and Slack, F.J. (2003). The time of appearance of the C. elegans let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. Dev Biol *259*, 364-379.
43. Poy, M.N., et al. (2004). A pancreatic islet-specific microRNA regulates insulin secretion. Nature *432*, 226-230.
44. Chen, C.Z., Li, L., Lodish, H.F., and Bartel, D.P. (2004). MicroRNAs modulate hematopoietic lineage differentiation. Science *303*, 83-86.
45. Sempere, L.F., et al. (2004). Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation. Genome Biol *5*, R13.
46. Babak, T., Zhang, W., Morris, Q., Blencowe, B.J., and Hughes, T.R. (2004). Probing microRNAs with microarrays: tissue specificity and functional inference. Rna *10*, 1813-1819.
47. Barad, O., et al. (2004). MicroRNA expression detected by oligonucleotide microarrays: system establishment and expression profiling in human tissues. Genome Res *14*, 2486-2494.
48. Thomson, J.M., Parker, J., Perou, C.M., and Hammond, S.M. (2004). A custom microarray platform for analysis of microRNA gene expression. Nat Methods *1*, 1-7.
49. Krichevsky, A.M., King, K.S., Donahue, C.P., Khrapko, K., and Kosik, K.S. (2003). A microRNA array reveals extensive regulation of microRNAs during brain development. Rna *9*, 1274-1281.
50. Lee, Y., et al. (2004). MicroRNA genes are transcribed by RNA polymerase II. Embo J *23*, 4051-4060.
51. Cai, X., Hagedorn, C.H., and Cullen, B.R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. Rna *10*, 1957-1966.
52. Smalheiser, N.R., and Torvik, V.I. (2004). A population-based statistical approach identifies parameters characteristic of human microRNA-mRNA interactions. BMC Bioinformatics *5*, 139.
53. Lee, Y., Jeon, K., Lee, J.T., Kim, S., and Kim, V.N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. Embo J *21*, 4663-4670.
54. Lee, Y., et al. (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415-419.
55. Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. Nature *432*, 231-235.
56. Gregory, R.I., et al. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature *432*, 235-240.
57. Han, J., et al. (2004). The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev *18*, 3016-3027.
58. Landthaler, M., Yalcin, A., and Tuschl, T. (2004). The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis. Curr Biol *14*, 2162-2167.
59. Filippov, V., Solovyev, V., Filippova, M., and Gill, S.S. (2000). A novel type of RNase III family proteins in eukaryotes. Gene *245*, 213-221.
60. Wu, H., Xu, H., Miraglia, L.J., and Crooke, S.T. (2000). Human RNase III is a 160-kDa protein involved in preribosomal RNA processing. J Biol Chem *275*, 36957-36965.
61. Basyuk, E., Suavet, F., Doglio, A., Bordonne, R., and Bertrand, E. (2003). Human let-7 stem-loop precursors harbor features of RNase III cleavage products. Nucleic Acids Res *31*, 6593-6597.
62. Zeng, Y., Yi, R., and Cullen, B.R. (2005). Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. Embo J *24*, 138-148.
63. Zeng, Y., and Cullen, B.R. (2005). Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. J Biol Chem *280*, 27595-27603.
64. Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear export of microRNA precursors. Science *303*, 95-98.
65. Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev *17*, 3011-3016.
66. Bohnsack, M.T., Czaplinski, K., and Gorlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. Rna *10*, 185-191.
67. Hutvagner, G., et al. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science *293*, 834-838.
68. Grishok, A., et al. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell *106*, 23-34.
69. Ketting, R.F., et al. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev. *15*, 2654-2659.
70. Zamore, P.D., Tuschl, T., Sharp, P.A., and Bartel, D.P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell *101*, 25-33.
71. Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature *409*, 363-366.
72. Elbashir, S.M., Lendeckel, W., and Tuschl, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes. Dev. *15*, 188-200.
73. Bernstein, E., et al. (2003). Dicer is essential for mouse development. Nat Genet *35*, 215-217.
74. Wienholds, E., Koudijs, M.J., van Eeden, F.J., Cuppen, E., and Plasterk, R.H. (2003). The microRNA-producing enzyme Dicer1 is essential for zebrafish development. Nat Genet *35*, 217-218.
75. Knight, S.W., and Bass, B.L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. Science *293*, 2269-2271.

76. Lee, Y.S., et al. (2004). Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell *117*, 69-81.
77. Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004). Single processing center models for human Dicer and bacterial RNase III. Cell *118*, 57-68.
78. Chendrimada, T.P., et al. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature *436*, 740-744.
79. Zhong, J., Peters, A.H., Lee, K., and Braun, R.E. (1999). A double-stranded RNA binding protein required for activation of repressed messages in mammalian germ cells. Nat Genet *22*, 171-174.
80. Tabara, H., Yigit, E., Siomi, H., and Mello, C.C. (2002). The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in C. elegans. Cell *109*, 861-871.
81. Hatfield, S.D., et al. (2005). Stem cell division is regulated by the microRNA pathway. Nature *435*, 974-978.
82. Forstemann, K., et al. (2005). Normal microRNA Maturation and Germ-Line Stem Cell Maintenance Requires Loquacious, a Double-Stranded RNA-Binding Domain Protein. PLoS Biol *3*, e236.
83. Saito, K., Ishizuka, A., Siomi, H., and Siomi, M.C. (2005). Processing of Pre-microRNAs by the Dicer-1-Loquacious Complex in Drosophila Cells. PLoS Biol *3*, e235.
84. Jiang, F., et al. (2005). Dicer-1 and R3D1-L catalyze microRNA maturation in Drosophila. Genes Dev *19*, 1674-1679.
85. Liu, Q., et al. (2003). R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. Science *301*, 1921-1925.
86. Khvorova, A., Reynolds, A., and Jayasena, S.D. (2003). Functional siRNAs and miRNAs exhibit strand bias. Cell *115*, 209-216.
87. Schwarz, D.S., et al. (2003). Asymmetry in the assembly of the RNAi enzyme complex. Cell *115*, 199-208.
88. Tomari, Y., Matranga, C., Haley, B., Martinez, N., and Zamore, P.D. (2004). A protein sensor for siRNA asymmetry. Science *306*, 1377-1380.
89. Nykanen, A., Haley, B., and Zamore, P.D. (2001). ATP requirements and small interfering RNA structure in the RNA interference pathway. Cell *107*, 309-321.
90. Chiu, Y.L., and Rana, T.M. (2002). RNAi in human cells: basic structural and functional features of small interfering RNA. Mol Cell *10*, 549-561.
91. Chiu, Y.L., et al. (2005). Dissecting RNA-Interference Pathway with Small Molecules. Chem Biol *12*, 643-648.
92. Hutvagner, G., and Zamore, P.D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. Science *297*, 2056-2060.
93. Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. Nature *404*, 293-296.
94. Pham, J.W., Pellino, J.L., Lee, Y.S., Carthew, R.W., and Sontheimer, E.J. (2004). A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in Drosophila. Cell *117*, 83-94.
95. Doench, J.G., Petersen, C.P., and Sharp, P.A. (2003). siRNAs can function as miRNAs. Genes Dev *17*, 438-442.
96. Zeng, Y., Yi, R., and Cullen, B.R. (2003). MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. Proc Natl Acad Sci U S A *100*, 9779-9784.
97. Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R., and Hannon, G.J. (2001). Argonaute2, a link between genetic and biochemical analyses of RNAi. Science *293*, 1146-1150.
98. Tabara, H., et al. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell *99*, 123-132.
99. Cogoni, C., and Macino, G. (1997). Isolation of quelling-defective (qde) mutants impaired in posttranscriptional transgene-induced gene silencing in Neurospora crassa. Proc Natl Acad Sci U S A *94*, 10233-10238.
100. Grishok, A., Tabara, H., and Mello, C.C. (2000). Genetic requirements for inheritance of RNAi in C. elegans. Science *287*, 2494-2497.
101. Fagard, M., Boutet, S., Morel, J.B., Bellini, C., and Vaucheret, H. (2000). AGO1, QDE-2, and RDE-1 are related proteins required for post- transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals. Proc Natl Acad Sci U S A *97*, 11650-11654.
102. Pal-Bhadra, M., Bhadra, U., and Birchler, J.A. (2002). RNAi related mechanisms affect both transcriptional and posttranscriptional transgene silencing in Drosophila. Mol. Cell *9*, 315-327.
103. Mochizuki, K., Fine, N.A., Fujisawa, T., and Gorovsky, M.A. (2002). Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena. Cell *110*, 689-699.
104. Carmell, M.A., Xuan, Z., Zhang, M.Q., and Hannon, G.J. (2002). The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes Dev *16*, 2733-2742.
105. Sasaki, T., Shiohama, A., Minoshima, S., and Shimizu, N. (2003). Identification of eight members of the Argonaute family in the human genome small star, filled. Genomics *82*, 323-330.
106. Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R., and Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell *110*, 563-574.
107. Meister, G., et al. (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Mol Cell *15*, 185-197.
108. Liu, J., et al. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. Science *305*, 1437-1441.
109. Song, J.J., Smith, S.K., Hannon, G.J., and Joshua-Tor, L. (2004). Crystal structure of Argonaute and its implications for RISC slicer activity. Science *305*, 1434-1437.
110. Rand, T.A., Ginalski, K., Grishin, N.V., and Wang, X. (2004). Biochemical identification of Argonaute 2 as the sole protein required for RNA-induced silencing complex activity. Proc Natl Acad Sci U S A *101*, 14385-14389.
111. Rivas, F.V., et al. (2005). Purified Argonaute2 and an siRNA form recombinant human RISC. Nat Struct Mol Biol *12*, 340-349.
112. Zilberman, D., Cao, X., and Jacobsen, S.E. (2003). ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science *299*, 716-719.

113. Pal-Bhadra, M., et al. (2004). Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. Science *303*, 669-672.
114. Williams, R.W., and Rubin, G.M. (2002). ARGONAUTE1 is required for efficient RNA interference in Drosophila embryos. Proc Natl Acad Sci U S A *99*, 6889-6894.
115. Okamura, K., Ishizuka, A., Siomi, H., and Siomi, M.C. (2004). Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. Genes Dev *18*, 1655-1666.
116. Kataoka, Y., Takeichi, M., and Uemura, T. (2001). Developmental roles and molecular characterization of a Drosophila homologue of Arabidopsis Argonaute1, the founder of a novel gene superfamily. Genes Cells *6*, 313-325.
117. Ma, J.B., Ye, K., and Patel, D.J. (2004). Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. Nature *429*, 318-322.
118. Song, J.J., et al. (2003). The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. Nat Struct Biol *10*, 1026-1032.
119. Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2003). Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. Nature *426*, 465-469.
120. Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2004). Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. Nat Struct Mol Biol *11*, 576-577.
121. Yan, K.S., et al. (2003). Structure and conserved RNA binding of the PAZ domain. Nature *426*, 468-474.
122. Parker, J.S., Roe, S.M., and Barford, D. (2004). Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. Embo J *23*, 4727-4737.
123. Ma, J.B., et al. (2005). Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein. Nature *434*, 666-670.
124. Parker, J.S., Roe, S.M., and Barford, D. (2005). Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. Nature *434*, 663-666.
125. Doi, N., et al. (2003). Short-interfering-RNA-mediated gene silencing in mammalian cells requires Dicer and eIF2C translation initiation factors. Curr Biol *13*, 41-46.
126. Tahbaz, N., et al. (2004). Characterization of the interactions between mammalian PAZ PIWI domain proteins and Dicer. EMBO Rep *5*, 189-194.
127. Jin, P., et al. (2004). Biochemical and genetic interaction between the fragile X mental retardation protein and the microRNA pathway. Nat Neurosci *7*, 113-117.
128. Caudy, A.A., et al. (2003). A micrococcal nuclease homologue in RNAi effector complexes. Nature *425*, 411-414.
129. Ishizuka, A., Siomi, M.C., and Siomi, H. (2002). A Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins. Genes Dev *16*, 2497-2508.
130. Caudy, A.A., Myers, M., Hannon, G.J., and Hammond, S.M. (2002). Fragile X-related protein and VIG associate with the RNA interference machinery. Genes Dev *16*, 2491-2496.
131. Scadden, A.D. (2005). The RISC subunit Tudor-SN binds to hyper-edited double-stranded RNA and promotes its cleavage. Nat Struct Mol Biol *12*, 489-496.
132. Tomari, Y., et al. (2004). RISC assembly defects in the Drosophila RNAi mutant armitage. Cell *116*, 831-841.
133. Doench, J.G., and Sharp, P.A. (2004). Specificity of microRNA target selection in translational repression. Genes Dev *18*, 504-511.
134. Kiriakidou, M., et al. (2004). A combined computational-experimental approach predicts human microRNA targets. Genes Dev *18*, 1165-1178.
135. Kloosterman, W.P., Wienholds, E., Ketting, R.F., and Plasterk, R.H. (2004). Substrate requirements for let-7 function in the developing zebrafish embryo. Nucleic Acids Res *32*, 6284-6291.
136. Brennecke, J., Stark, A., Russell, R.B., and Cohen, S.M. (2005). Principles of MicroRNA-Target Recognition. PLoS Biol *3*, e85.
137. Mallory, A.C., et al. (2004). MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region. Embo J *23*, 3356-3364.
138. Haley, B., and Zamore, P.D. (2004). Kinetic analysis of the RNAi enzyme complex. Nat Struct Mol Biol *11*, 599-606.
139. Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P., and Burge, C.B. (2003). Prediction of mammalian microRNA targets. Cell *115*, 787-798.
140. Lai, E.C. (2002). Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. Nat Genet *30*, 363-364.
141. Stark, A., Brennecke, J., Russell, R.B., and Cohen, S.M. (2003). Identification of *Drosophila* MicroRNA Targets. PLOS Biol. *1*, 1-13.
142. Saxena, S., Jonsson, Z.O., and Dutta, A. (2003). Small RNAs with imperfect match to endogenous mRNA repress translation. Implications for off-target activity of small inhibitory RNA in mammalian cells. J Biol Chem *278*, 44312-44319.
143. Zeng, Y., Wagner, E.J., and Cullen, B.R. (2002). Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. Mol Cell *9*, 1327-1333.
144. Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell *75*, 855-862.
145. Olsen, P.H., and Ambros, V. (1999). The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. Dev Biol *216*, 671-680.
146. Seggerson, K., Tang, L., and Moss, E.G. (2002). Two genetic circuits repress the Caenorhabditis elegans heterochronic gene lin-28 after translation initiation. Dev Biol *243*, 215-225.
147. Pillai, R.S., et al. (2005). Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science *309*, 1573-1576.
148. Schwarz, D.S., Hutvagner, G., Haley, B., and Zamore, P.D. (2002). Evidence that siRNAs function as guides, not primers, in the Drosophila and human RNAi pathways. Mol Cell *10*, 537-548.

149. Schwarz, D.S., Tomari, Y., and Zamore, P.D. (2004). The RNA-induced silencing complex is a Mg2+-dependent endonuclease. Curr Biol *14*, 787-791.
150. Martinez, J., and Tuschl, T. (2004). RISC is a 5' phosphomonoester-producing RNA endonuclease. Genes Dev *18*, 975-980.
151. Elbashir, S.M., Martinez, J., Patkaniowska, A., Lendeckel, W., and Tuschl, T. (2001). Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. Embo J *20*, 6877-6888.
152. Shen, B., and Goodman, H.M. (2004). Uridine addition after microRNA-directed cleavage. Science *306*, 997.
153. Orban, T.I., and Izaurralde, E. (2005). Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. Rna *11*, 459-469.
154. Souret, F.F., Kastenmayer, J.P., and Green, P.J. (2004). AtXRN4 degrades mRNA in Arabidopsis and its substrates include selected miRNA targets. Mol Cell *15*, 173-183.
155. Newbury, S., and Woollard, A. (2004). The 5'-3' exoribonuclease xrn-1 is essential for ventral epithelial enclosure during C. elegans embryogenesis. Rna *10*, 59-65.
156. Gazzani, S., Lawrenson, T., Woodward, C., Headon, D., and Sablowski, R. (2004). A link between mRNA turnover and RNA interference in Arabidopsis. Science *306*, 1046-1048.
157. Yekta, S., Shih, I.H., and Bartel, D.P. (2004). MicroRNA-directed cleavage of HOXB8 mRNA. Science *304*, 594-596.
158. Lim, L.P., et al. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature *433*, 769-773.
159. Bagga, S., et al. (2005). Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. Cell *122*, 553-563.
160. Sheth, U., and Parker, R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. Science *300*, 805-808.
161. Cougot, N., Babajko, S., and Seraphin, B. (2004). Cytoplasmic foci are sites of mRNA decay in human cells. J Cell Biol *165*, 31-40.
162. Brengues, M., Teixeira, D., and Parker, R. (2005). Movement of Eukaryotic mRNAs Between Polysomes and Cytoplasmic Processing Bodies. Science.
163. Sen, G.L., and Blau, H.M. (2005). Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies. Nat Cell Biol *7*, 633-636.
164. Liu, J., Valencia-Sanchez, M.A., Hannon, G.J., and Parker, R. (2005). MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. Nat Cell Biol *7*, 719-723.
165. Ding, L., Spencer, A., Morita, K., and Han, M. (2005). The developmental timing regulator AIN-1 interacts with miRISCs and may target the argonaute protein ALG-1 to cytoplasmic P bodies in C. elegans. Mol Cell *19*, 437-447.
166. Jing, Q., et al. (2005). Involvement of MicroRNA in AU-Rich Element-Mediated mRNA Instability. Cell *120*, 623-634.
167. Meister, G., Landthaler, M., Dorsett, Y., and Tuschl, T. (2004). Sequence-specific inhibition of microRNA- and siRNA-induced RNA silencing. Rna *10*, 544-550.
168. Hutvagner, G., Simard, M.J., Mello, C.C., and Zamore, P.D. (2004). Sequence-Specific Inhibition of Small RNA Function. PLoS Biol *2*, E98.
169. Lecellier, C.H., et al. (2005). A cellular microRNA mediates antiviral defense in human cells. Science *308*, 557-560.
170. Lee, Y.S., Kim, H.K., Chung, S., Kim, K.S., and Dutta, A. (2005). Depletion of human micro-RNA miR-125b reveals that it is critical for the proliferation of differentiated cells but not for the down-regulation of putative targets during differentiation. J Biol Chem *280*, 16635-16641.
171. Lai, E.C. (2004). Predicting and validating microRNA targets. Genome Biol *5*, 115.
172. Bentwich, I. (2005). Prediction and validation of miRNAs and their targets. FEBS Lett, In press.
173. John, B., et al. (2004). Human MicroRNA targets. PLoS Biol *2*, e363.
174. Krek, A., et al. (2005). Combinatorial microRNA target predictions. Nat Genet *37*, 495-500.
175. Vella, M.C., Choi, E.Y., Lin, S.Y., Reinert, K., and Slack, F.J. (2004). The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR. Genes Dev *18*, 132-137.
176. Moss, E.G., Lee, R.C., and Ambros, V. (1997). The cold shock domain protein LIN-28 controls developmental timing in C. elegans and is regulated by the lin-4 RNA. Cell *88*, 637-646.
177. Lin, S.Y., et al. (2003). The C elegans hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target. Dev Cell *4*, 639-650.
178. Abrahante, J.E., et al. (2003). The Caenorhabditis elegans hunchback-like gene lin-57/hbl-1 controls developmental time and is regulated by microRNAs. Dev Cell *4*, 625-637.
179. Grosshans, H., Johnson, T., Reinert, K.L., Gerstein, M., and Slack, F.J. (2005). The Temporal Patterning MicroRNA let-7 Regulates Several Transcription Factors at the Larval to Adult Transition in C. elegans. Dev Cell *8*, 321-330.
180. Johnson, S.M., et al. (2005). RAS Is Regulated by the let-7 MicroRNA Family. Cell *120*, 635-647.
181. Chang, S., Johnston, R.J., Jr., Frokjaer-Jensen, C., Lockery, S., and Hobert, O. (2004). MicroRNAs act sequentially and asymmetrically to control chemosensory laterality in the nematode. Nature *430*, 785-789.
182. Boutla, A., Delidakis, C., and Tabler, M. (2003). Developmental defects by antisense-mediated inactivation of micro-RNAs 2 and 13 in Drosophila and the identification of putative target genes. Nucleic Acids Res *31*, 4973-4980.
183. Lai, E.C., Tam, B., and Rubin, G.M. (2005). Pervasive regulation of Drosophila Notch target genes by GY-box-, Brd-box-, and K-box-class microRNAs. Genes Dev *19*, 1067-1080.
184. Kanellopoulou, C., et al. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev *19*, 489-501.
185. Giraldez, A.J., et al. (2005). MicroRNAs regulate brain morphogenesis in zebrafish. Science *308*, 833-838.
186. Fukagawa, T., et al. (2004). Dicer is essential for formation of the heterochromatin structure in vertebrate cells. Nat Cell Biol *6*, 784-791.

187. Mansfield, J.H., et al. (2004). MicroRNA-responsive 'sensor' transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. Nat Genet *36*, 1079-1083.
188. Zhao, Y., Samal, E., and Srivastava, D. (2005). Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. Nature *436*, 214-220.
189. Esau, C., et al. (2004). MicroRNA-143 regulates adipocyte differentiation. J Biol Chem *279*, 52361-52365.
190. Seitz, H., et al. (2003). Imprinted microRNA genes transcribed antisense to a reciprocally imprinted retrotransposon-like gene. Nat Genet *34*, 261-262.
191. Seitz, H., et al. (2004). A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. Genome Res *14*, 1741-1748.
192. Chang, J., et al. (2004). miR-122, a Mammalian Liver-Specific microRNA, is Processed from hcr mRNA and May Downregulate the High Affinity Cationic Amino Acid Transporter CAT-1. RNA Biology *1*, 106-113.
193. Yu, Z., Raabe, T., and Hecht, N.B. (2005). MicroRNA Mirn122a Reduces Expression of the Posttranscriptionally Regulated Germ Cell Transition Protein 2 (Tnp2) Messenger RNA (mRNA) by mRNA Cleavage. Biol Reprod *73*, 427-433.
194. Cheng, A.M., Byrom, M.W., Shelton, J., and Ford, L.P. (2005). Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis. Nucleic Acids Res *33*, 1290-1297.
195. Lu, J., et al. (2005). MicroRNA expression profiles classify human cancers. Nature *435*, 834-838.
196. Bartel, D.P., and Chen, C.Z. (2004). Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. Nat Rev Genet *5*, 396-400.
197. Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.
198. Karube, Y., et al. (2005). Reduced expression of Dicer associated with poor prognosis in lung cancer patients. Cancer Sci *96*, 111-115.
199. Calin, G.A., et al. (2002). Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A *99*, 15524-15529.
200. Michael, M.Z., SM, O.C., van Holst Pellekaan, N.G., Young, G.P., and James, R.J. (2003). Reduced accumulation of specific microRNAs in colorectal neoplasia. Mol Cancer Res *1*, 882-891.
201. Metzler, M., Wilda, M., Busch, K., Viehmann, S., and Borkhardt, A. (2004). High expression of precursor microRNA-155/BIC RNA in children with Burkitt lymphoma. Genes Chromosomes Cancer *39*, 167-169.
202. Eis, P.S., et al. (2005). Accumulation of miR-155 and BIC RNA in human B cell lymphomas. Proc Natl Acad Sci U S A *102*, 3627-3632.
203. Takamizawa, J., et al. (2004). Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. Cancer Res *64*, 3753-3756.
204. He, L., et al. (2005). A microRNA polycistron as a potential human oncogene. Nature *435*, 828-833.
205. O'Donnell, K.A., Wentzel, E.A., Zeller, K.I., Dang, C.V., and Mendell, J.T. (2005). c-Myc-regulated microRNAs modulate E2F1 expression. Nature *435*, 839-843.
206. Cai, X., et al. (2005). Kaposi's sarcoma-associated herpesvirus expresses an array of viral microRNAs in latently infected cells. Proc Natl Acad Sci U S A *102*, 5570-5575.
207. Sullivan, C.S., Grundhoff, A.T., Tevethia, S., Pipas, J.M., and Ganem, D. (2005). SV40-encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. Nature *435*, 682-686.
208. Omoto, S., et al. (2004). HIV-1 nef suppression by virally encoded microRNA. Retrovirology *1*, 44.

Chapter 2

Target-selected inactivation of the zebrafish *rag1* gene

Reprinted from Wienholds *et al*. (2002), *Science* 297:99-102

been shown to inhibit the Met-aminopeptidase MetAP2 (*17, 18*). The effect of these antiangiogenic reagents may stem from inhibition of the N-terminal Met-Cys cleavage in a normally short-lived regulator of angiogenesis that is targeted by the N-end rule pathway through its N-terminal Cys residue.

The oxidation (and subsequent arginylation) of N-terminal Cys may compete with its other known modifications, including acetylation and palmitoylation. N-end rule substrates with the arginylation-dependent destabilizing N-terminal residues (Asn, Gln, Asp, Glu, and Cys) (Fig. 1A) can also be produced through cleavages anywhere in a protein's polypeptide chain. For example, the conditional cleavage of a subunit of the mammalian cohesin complex at the metaphase-anaphase transition is predicted to produce a putative N-end rule substrate whose degradation would require N-terminal arginylation (*8, 19*).

HIF1α, a subunit of hypoxia-inducible factor 1 (HIF1) that functions as a key regulator of angiogenesis, is a conditionally short-lived protein. The degron of HIF1α, recognized by a distinct Ub-dependent proteolytic pathway, is activated through the oxygen-dependent hydroxylation of a specific Pro residue (*20, 21*). By analogy to prolyl-4-hydroxylases that regulate the degron of HIF1α, the currently unknown enzyme that oxidizes N-terminal Cys may also function as an oxygen sensor. If so, the formation and maintenance of the cardiovascular system may involve a battery of distinct, conditionally short-lived regulators such as HIF1 and the currently unknown substrate of the N-end rule pathway that bears N-terminal Cys.

References and Notes

1. A. Bachmair, D. Finley, A. Varshavsky, *Science* **234**, 179 (1986).
2. A. Varshavsky, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12142 (1996).
3. Y. T. Kwon, Z.-X. Xia, I. V. Davydov, S. H. Lecker, A. Varshavsky, *Mol. Cell. Biol.* **21**, 8007 (2001).
4. A. Hershko, A. Ciechanover, A. Varshavsky, *Nature Med.* **10**, 1073 (2000).
5. Y. T. Kwon, A. S. Kashina, A. Varshavsky, *Mol. Cell. Biol.* **19**, 182 (1999).
6. G. C. Turner, F. Du, A. Varshavsky, *Nature* **405**, 579 (2000).
7. Y. T. Kwon *et al.*, *Mol. Cell. Biol.* **20**, 4135 (2000).
8. H. Rao, F. Uhlmann, K. Nasmyth, A. Varshavsky, *Nature* **410**, 955 (2001).
9. Y. T. Kwon *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7898 (1998).
10. D. K. Gonda *et al.*, *J. Biol. Chem.* **264**, 16700 (1989).
11. I. V. Davydov, A. Varshavsky, *J. Biol. Chem.* **275**, 22931 (2000).
12. Materials and methods are available as supplementary material on *Science* Online.
13. T. Tetaz *et al.*, *Biochem. Int.* **22**, 561 (1990).
14. T. Yagi, H. Kagamiyama, M. Nozaki, *Biochem. Biophys. Res. Commun.* **90**, 447 (1979).
15. J. Li, C. M. Pickart, *Biochemistry* **34**, 15829 (1995).
16. H. Kaiji, G. D. Novelli, A. Kaiji, *Biochim. Biophys. Acta* **76**, 474 (1963).
17. E. C. Griffith *et al.*, *Chem. Biol.* **4**, 461 (1997).
18. N. Sin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6099 (1997).
19. S. Hauf, I. C. Waizenegger, J.-M. Peters, *Science* **293**, 1320 (2001).
20. M. Ivan *et al.*, *Science* **242**, 464 (2001).
21. P. Jaakkola *et al.*, *Science* **242**, 468 (2001).
22. Single-letter abbreviations for amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
23. G. C. Turner, A. Varshavsky, *Science* **289**, 2117 (2000).

24. We thank D. Anderson, D. Shin, and Y. Mukouyama for helpful discussions; C. Brower for comments on the manuscript; S. Pease, B. Kennedy, and L. Sandoval for expert care of mice; G. Hathaway for sequencing of RGS4; N. Dinh and M. Young for mass spectrometry; and J. K. Yoon for pLacF. Supported by NIH grant GM31530 and a Kirsch Foundation grant to A.V.

Supporting Online Material

www.sciencemag.org/cgi/content/full/297/5578/96/DC1
Materials and Methods
Figs. S1 to S5
References

3 January 2002; accepted 5 June 2002

Target-Selected Inactivation of the Zebrafish *rag1* Gene

Erno Wienholds,[1] Stefan Schulte-Merker,[2] Brigitte Walderich,[2] Ronald H. A. Plasterk[1]*

The zebrafish has become a favorite organism for genetic analysis of vertebrate development, but methods for generating mutants by reverse genetic approaches have been lacking. We report a method to obtain stable mutants of a gene based on knowledge of the gene sequence only. Parental fish were mutagenized with *N*-ethyl-*N*-nitrosourea; in 2679 F_1 fish, the *rag1* gene was analyzed for heterozygous mutations by resequencing. In total, we found 15 mutations: 9 resulted in amino acid substitutions and 1 resulted in a premature stop codon. This truncation mutant was found to be homozygous viable and defective in V(D)J joining. Although presumably immune deficient, these homozygous *rag1* mutant fish are able to reach adulthood and are fertile. As sperm samples from all 2679 F_1 fish were collected and cryopreserved, we have in principle generated a mutant library from which mutants of most zebrafish genes can be isolated.

Forward genetic screens in the zebrafish have identified thousands of mutants defective in many kinds of biological processes (*1*, *2*). Cloning of the affected genes is accelerated by both an extensive and rapidly growing genetic linkage map (*3*–*5*) and the zebrafish genome sequencing project. However, until now a major drawback in zebrafish genetic analysis has been the inability to specifically study gene function through reverse genetics. It has been demonstrated that protein levels in embryos can be temporally reduced by antisense morpholino oligonucleotides (*6*), but this effect is transient and is not applicable to later stages of development.

In the mouse, reverse genetics was made possible by homologous recombination in embryonic stem cells, eventually giving rise to germ line transmission of the mutant alleles (*7*, *8*). Recently, it has been shown that in zebrafish short-term embryonic stem cell cultures are able to produce germ line chimeras (*9*), but this has not yet been extended to targeted gene inactivation. Therefore, we have taken another approach: target-selected mutagenesis. This involves random mutagenesis, followed by targeted screening for induced mutations at the genomic DNA level. For example, in *Caenorhabditis elegans*, we and others (*10*, *11*) have successfully performed ethylmethane sulfonate mutagenesis and have screened animals with a polymerase chain reaction (PCR) strategy for deletions in hundreds of target genes. In the case of zebrafish, we have now combined *N*-ethyl-*N*-nitrosourea (ENU) mutagenesis with high-throughput resequencing of the target gene.

Random ENU mutagenesis in zebrafish is widely used in forward screens. Nonmosaic mutants are typically found at tester loci at an average frequency of 1 in 650 mutagenized genomes. Recessive embryonic lethal phenotypes occur at a similarly high rate, namely, an average loss-of-function frequency per locus of 1 in 1600 mutagenized genomes (*1*, *2*). Therefore, we constructed a library of 2679 randomly ENU mutagenized F_1 males as outlined in Fig. 1, making use of the Tübingen 2000 screen fish. Genomic DNA was isolated, and testis samples were cryopreserved (*12*), generating a permanent library that could be screened for heterozygous mutations. In principle, we have isolated sufficient DNA samples to screen for mutations in most zebrafish genes, and the library is comprehensive enough that most genes should be represented by at least one null allele.

To test this approach, we screened for mutations in the *rag1* gene (*13*) of each individual

[1]Hubrecht Laboratory, Center for Biomedical Genetics, Uppsalalaan 8, 3584 CT, Utrecht, Netherlands. [2]ARTEMIS Pharmaceuticals GmbH/ Exelixis Deutschland, Spemannstrasse 35, 72076 Tübingen, Germany.

*To whom correspondence should be addressed. E-mail: plasterk@niob.knaw.nl.

F_1 fish by DNA resequencing. By nested PCR, we amplified two fragments that included exons 2 and 3 of the coding region (Fig. 2A). Subsequent sequence analysis with five different primers (Fig. 2A; Table 1) was done on a 96-capillary ABI3700 DNA analyzer (*12*). Heterozygosity detection was done with the programs Phred (*14*), Phrap (*15*), and PolyPhred (*16*), and was visualized with Consed (*17*).

Single nucleotide polymorphisms (SNPs) were encountered at a high frequency; they could be distinguished from ENU-induced mutations because they were detected in multiple fish. As shown in Fig. 2B, we detected two haplotypes that differ from the published sequence of *rag1* at about 1 in 116 nucleotides; five of these SNPs result in coding changes in the *rag1* gene. The two haplotypes indicate that the fish used for this screen, the so-called Tübingen strain, is not homogeneous at this locus. Other loci also have been found to show heterogeneity within this library of F_1 fish (*18*).

After we filtered the SNPs, we detected a total of 15 ENU-induced mutations (Table 1). Each of these mutations was verified by independent PCR and resequencing. Because most mutation spectra are based on the detection bias of the forward screen, we can use these data to get an initial unbiased picture of the mutation spectrum of ENU in the zebrafish. As shown in Fig. 2C, the distribution of the mutations found in *rag1* is similar to those found in previous forward screens. The mutation rate is 15 mutations in 7,268,581 base pairs among 2679 fish (Table 1), which corresponds to a frequency of 2.1×10^{-6} per base pair.

Of the 15 mutations in *rag1*, 3 are silent, 2 are in intronic sequences (some introns were partially covered), and 10 result in changes in the coding sequence, of which 9

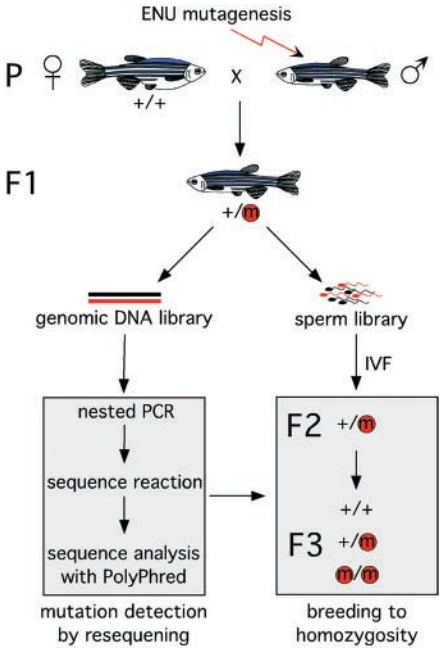


Fig. 1. Overview of target-selected mutagenesis in zebrafish. Ninety-nine adult male zebrafish were mutagenized by three to five consecutive treatments with 3 mM ENU, in accordance with (*32*). The mutagenized fish were crossed with wild-type females to give a nonmosaic F_1 generation of fish. Sperm was isolated and cryopreserved from 2679 fertile F_1 males. Genomic DNA was isolated, arrayed in PCR plates, and screened for mutations by nested PCR amplification of the target gene and subsequent DNA sequence analysis. After a particular mutation was identified, in vitro fertilizations (IVF) were performed to recover the F_2 line carrying the mutation (*12*). Finally, mutations can be bred to homozygosity and analyzed for phenotypes.

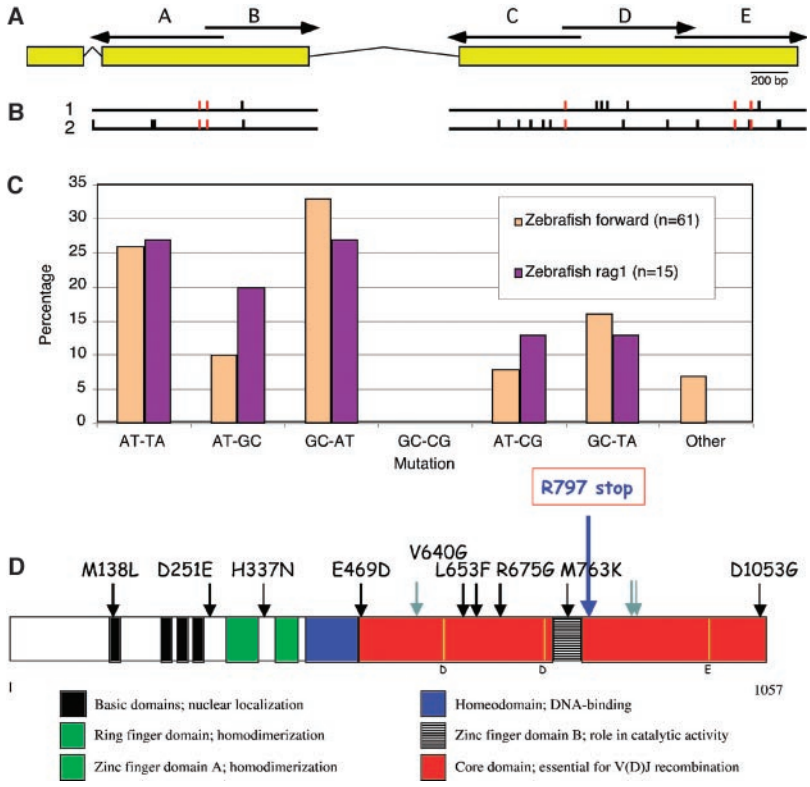


Fig. 2. Resequencing the *rag1* gene. (**A**) Genomic organization of the *rag1* gene. Coding regions are represented as yellow boxes. Regions A to E indicate segments screened by resequencing. (**B**) Distribution of SNPs in the two haplotypes of the *rag1* gene found in this study compared with the published sequence (*13*). Black bars indicate a SNP in one haplotype; red bars indicate a SNP in both haplotypes. (**C**) Spectrum of the ENU-induced mutations found in this screen compared with the mutations detected in previous forward screens. (**D**) Mutations changing the coding region of the *rag1* gene. RAG1 domain structure is adapted from (*33*). The DDE motif characteristic of recombinases (*34*) is indicated below the structure. Amino acid changes are indicated with black arrows above the structure: D, Asp; E, Glu; F, Phe; G, Gly; H, His; K, Lys; L, Leu; M, Met; N, Asn; R, Arg; V, Val. Purple arrow indicates stop codon. Gray arrows indicate silent mutations.

Table 1. *Rag1* resequencing statistics. Fragments correspond to the fragments indicated in Fig. 2A. All sequencing reactions were analyzed with PolyPhred (version 3.5 beta) with a ratio of 0.65, a background ratio of 0.15, and an average Phred score of 25. The total number of base pairs (bp) screened is corrected for fragment overlap and primer sequences. Failed sequences were excluded. Average lengths include standard errors. One mutation was found in fragment D, as well as in the overlapping part of fragment E.

Fragment	No. of reactions	Avg. length (bp)	No. of bp screened	No. of mutations
		Exon 2 (1118 bp)		
A	2,564	629 ± 2.2	1,613,420	4
B	2,615	602 ± 1.6	1,574,606	2
		Exon 3 (1748 bp)		
C	2,603	619 ± 1.5	1,611,149	4
D	2,053	581 ± 1.5	1,193,172	3
E	2,640	656 ± 1.2	1,730,797	3
Total	12,475	619 ± 0.8	7,723,144	16
Total corrected			7,268,581	15

are missense mutations and 1 is a premature stop (Fig. 2D). The latter is in the middle of the catalytic domain and, therefore, is expected to be a complete null allele. We genotyped the F_2 progeny of the fish that carry this allele (*rag1*^{t26683}) by isolating genomic DNA from fin clips, nested PCR, and sequencing. Crossing two F_2 heterozygotes gave Mendelian segregation of the mutant allele. The homozygous F_3 mutant embryos show *rag1* messenger RNA (mRNA) expression in the thymus (*19*) as seen by whole-mount in situ hybridizations (*18*). Adult mutant F_3 fish are viable and fertile in our standard nonsterile aquarium facilities.

A functional RAG1 protein is required for V(D)J recombination (*20*). Therefore, we examined functionality of the truncated RAG1 protein in mutant adults by a PCR assay (*21*) that checks for V(D)J rearrangements of the immunoglobulin heavy-chain genes (*22*) as outlined in Fig. 3A. Whereas heterozygous *rag1*^{t26683} mutant fish show joining of four different classes of heavy-chain genes, homozygous *rag1*^{t26683} mutant fish were found to be deficient in V(D)J joining (Fig. 3B). Apparently, as in other vertebrates, there is only one functional *rag1* gene; loss of function at this locus results in a complete block of immunoglobulin gene assembly and, presumably, in immunodeficiency.

Some of the missense mutations are at conserved amino acid positions; therefore, it is possible for these to affect gene function. Three other alleles have been tested for loss of function as heteroalleles over the null allele. These were not defective in V(D)J joining (Fig. 3B). Although we have assayed only for a loss-of-function phenotype, production of an allelic series is a major advantage of this method over knockout strategies that rely on homologous disruption or deletions. In some cases, complete loss of function may be lethal, and missense mutations may have informative viable phenotypes.

The clonal immune system was first acquired in evolution by teleost fish. Here we describe experimental inactivation of the immune system in a teleost fish. These presumably immunodeficient fish survive to adulthood without obvious signs of infectious disease in a nonsterile environment. Whether homozygous mutant lines can be bred without acquiring diseases at a higher rate than control cases remains to be seen. The *rag1*^{t26683} mutant is now available for detailed analysis of the role of the clonal immune system in resistance to infections and possibly in other processes.

We expect that the first-generation zebrafish mutant library described here will be followed by others, if only because the mutation efficiency can possibly be improved and the size of the library increased. Nevertheless, the current library has recently been used successfully to detect mutations in other genes (*18*), and new screens are ongoing. Once the DNA and the testes library were established, the time and reagent costs for the *rag1* screen were about 2 months and $12,500, respectively (*12*); these are currently being reduced by further automation and alternative mutation detection methods. With establishment of this technology, the zebrafish is now the second vertebrate model organism in which reverse genetics can be applied on a large scale.

Target-selected mutagenesis has also been successfully developed in the two major invertebrate model organisms, *Drosophila* (*23*–*25*) and *C. elegans* (*26*, *27*), as well as in *Arabidopsis* (*28*, *29*) and, recently, the mouse (*30*, *31*). The approach described here for zebrafish differs in that random mutagenesis is followed directly by straightforward, high-throughput resequencing of the target gene. This approach is insensitive to the presence of frequent SNPs in outbred organisms, because these are filtered out in the data analysis step. Therefore, this approach could be applicable for any organism on the condition that it can be bred and mutagenized efficiently. The advantage of multiple alleles, including partial loss-of-function mutants, may make it the preferred approach even for organisms for which gene knockout strategies already exist, such as *Drosophila* and *C. elegans*.

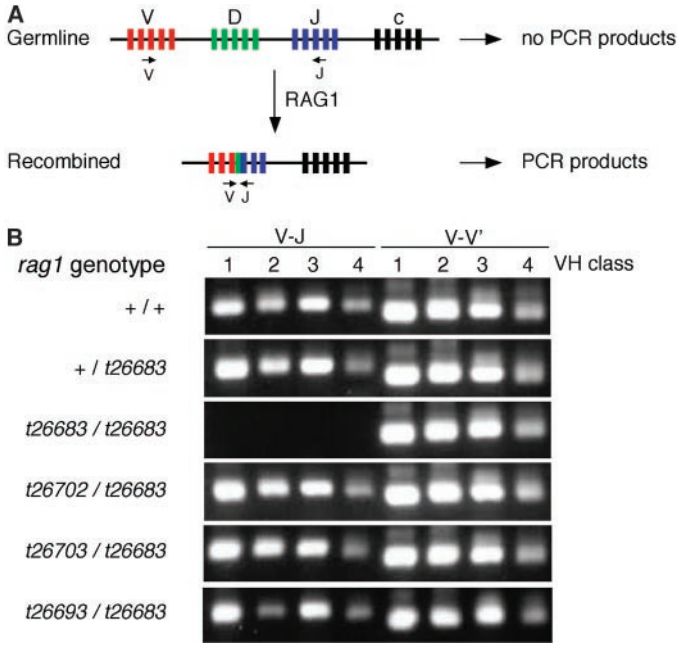


Fig. 3. Functionality of different RAG1 mutants. (**A**) Outline of the PCR assay used to test for V(D)J recombination of the immunoglobulin heavy-chain gene (*12*). PCR primers are represented as arrows. Semi-nested PCRs with an additional internal primer in the V segments were done to obtain a bright signal. C, constant domain. (**B**) V(D)J recombination in tail tissue from adult fish carrying different *rag1* alleles. The mutations tested are as follows: R797stop (*rag1*^{t26683}), D251E (*rag1*^{t26702}), L653F (*rag1*^{t26703}), and M138L (*rag1*^{t26693}). V-J, PCRs with primers in V segments and J segments. V-V', PCRs with primers in V segments only. VH, variable heavy chain.

References and Notes

1. P. Haffter *et al.*, *Development* **123**, 1 (1996).
2. W. Driever *et al.*, *Development* **123**, 37 (1996).
3. N. Hukriede *et al.*, *Genome Res.* **11**, 2127 (2001).
4. R. Geisler *et al.*, *Nature Genet.* **23**, 86 (1999).
5. I. G. Woods *et al.*, *Genome Res.* **10**, 1903 (2000).
6. A. Nasevicius, S. C. Ekker, *Nature Genet.* **26**, 216 (2000).
7. M. R. Capecchi, *Science* **244**, 1288 (1989).
8. S. Thompson, A. R. Clarke, A. M. Pow, M. L. Hooper, D. W. Melton, *Cell* **56**, 313 (1989).
9. C. Ma, L. Fan, R. Ganassin, N. Bols, P. Collodi, *Proc. Natl. Acad. Sci. U.S.A* **98**, 2461 (2001).
10. G. Jansen *et al.*, *Nature Genet.* **21**, 414 (1999).
11. L. X. Liu *et al.*, *Genome Res.* **9**, 859 (1999).
12. Materials and methods are available as supporting material on *Science* Online.
13. C. E. Willett, J. J. Cherry, L. A. Steiner, *Immunogenetics* **45**, 394 (1997).
14. B. Ewing, P. Green, *Genome Res.* **8**, 186 (1998).
15. D. Gordon, C. Desmarais, P. Green, *Genome Res.* **11**, 614 (2001).
16. D. A. Nickerson, V. O. Tobe, S. L. Taylor, *Nucleic Acids Res.* **25**, 2745 (1997).
17. D. Gordon, C. Abajian, P. Green, *Genome Res.* **8**, 195 (1998).
18. E. Wienholds, S. Schulte-Merker, B, Walderich, R. H. A. Plasterk, data not shown.
19. C. E. Willett, A. G. Zapata, N. Hopkins, L. A. Steiner, *Dev. Biol.* **182**, 331 (1997).
20. D. G. Schatz, M. A. Oettinger, D. Baltimore, *Cell* **59**, 1035 (1989).
21. N. Danilova and L. Steiner, personal communication.
22. N. Danilova, V. S. Hohman, E. H. Kim, L. A. Steiner, *Immunogenetics* **52**, 81 (2000).
23. D. G. Ballinger, S. Benzer, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9402 (1989).
24. K. Kaiser, S. F. Goodwin, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1686 (1990).
25. A. Bentley, B. MacLennan, J. Calvo, C. R. Dearolf, *Genetics* **156**, 1169 (2000).
26. R. R. Zwaal, A. Broeks, J. van Meurs, J. T. Groenen, R. H. Plasterk, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 7431 (1993).
27. G. Jansen, E. Hazendonk, K. L. Thijssen, R. H. Plasterk, *Nature Genet.* **17**, 119 (1997).
28. S. Parinov, V. Sundaresan, *Curr. Opin. Biotechnol.* **11**, 157 (2000).
29. C. M. McCallum, L. Comai, E. A. Greene, S. Henikoff, *Nature Biotechnol.* **18**, 455 (2000).
30. D. R. Beier, *Mamm. Genome* **11**, 294 (2000).
31. E. L. Coghill *et al.*, *Nature Genet.* **19**, 19 (2002).
32. F. J. van Eeden, M. Granato, J. Odenthal, P. Haffter, *Methods Cell Biol.* **60**, 21 (1999).
33. L. D. Notarangelo, A. Villa, K. Schwarz, *Curr. Opin. Immunol.* **11**, 435 (1999).
34. J. Kulkosky, K. S. Jones, R. A. Katz, J. P. Mack, A. M. Skalka, *Mol. Cell Biol.* **12**, 2331 (1992).

35. We thank R. Nordin, I. Santana-Stamm, H. Geiger,

D. Labusch, C. Weiler, and C. Weiler for help with testis freezing, in vitro fertilization and fin clipping, and the Tubingen 2000 screen consortium; S. Wicks, F. Verbeek, and E. Berezikov for help with bioinformatics; D. Jongejan-Zivkovic, F. van Eeden, E. Cuppen, and G. Stott for stimulating discussions and help with preparation of the manuscript; and L. Steiner for sharing unpublished results and providing primer sequences. R.H.A.P. is supported by a SPINOZA grant from the Netherlands Organization for Scientific Research.

Supporting Online Material
www.sciencemag.org/cgi/content/full/297/5578/99/DC1
Materials and Methods

12 March 2002; accepted 28 May 2002

Sustained Loss of a Neoplastic Phenotype by Brief Inactivation of *MYC*

Meenakshi Jain,[1]* Constadina Arvanitis,[1]* Kenneth Chu,[2] William Dewey,[2] Edith Leonhardt,[2] Maxine Trinh,[2] Christopher D. Sundberg,[1] J. Michael Bishop,[3] Dean W. Felsher[1]†

Pharmacological inactivation of oncogenes is being investigated as a possible therapeutic strategy for cancer. One potential drawback is that cessation of such therapy may allow reactivation of the oncogene and tumor regrowth. We used a conditional transgenic mouse model for *MYC*-induced tumorigenesis to demonstrate that brief inactivation of *MYC* results in the sustained regression of tumors and the differentiation of osteogenic sarcoma cells into mature osteocytes. Subsequent reactivation of *MYC* did not restore the cells' malignant properties but instead induced apoptosis. Thus, brief *MYC* inactivation appears to cause epigenetic changes in tumor cells that render them insensitive to *MYC*-induced tumorigenesis. These results raise the possibility that transient inactivation of *MYC* may be an effective therapy for certain cancers.

Activation of oncogenes plays an important role in tumorigenesis (*1*). Strategies that inactivate oncogenes for the treatment of cancer are in development; however, such approaches may be limited by the toxicity caused by the prolonged inactivation of the associated proto-oncogene. Moreover, cessation of the pharmacologic inactivation of an oncogene may result in tumor regrowth. To determine whether brief oncogene inactivation can produce sustained tumor regression, we used the tetracycline regulatory system to conditionally regulate *MYC* expression in transgenic mice. We previously described transgenic mice that conditionally express *MYC* in their lymphocytes (*2*). About 1% of these mice develop osteogenic sarcomas, and these tumors expressed abundant levels of MYC, presumably because the EμSRα enhancer causes *MYC* expression in immature osteoblasts. Consistent with this, *MYC* is commonly overexpressed in human and rodent osteogenic sarcomas (*3*–*11*).

The tumors in our transgenic model share some features with human osteogenic sarcoma (*12*–*14*). They present as invasive masses in the skeleton; they are associated with disorganized bone matrix; and they readily metastasize (fig. S1) (*15*). These properties were maintained as the tumors were adapted to in vitro growth and were inoculated into syngeneic hosts (*15*). To investigate the effects of *MYC* inactivation, we administered doxycycline (dox) treatment to mice with transplanted osteogenic sarcoma cells or primary transgenic tumors. After dox treatment in vivo, osteogenic sarcomas stopped expressing the *MYC* transgene, differentiated into mature bone, and exhibited sustained tumor regression (fig. S2). Similarly, primary transgenic tumors regressed and differentiated into bone (fig. S4). After dox treatment in vitro, the tumor cells exhibited a reduced growth rate, assumed a flattened morphology, lost alkaline phosphatase activity, and continued to express osteopontin (*15*) (Fig. 1). These phenotypic features are associated with the differentiation of immature osteoblasts into mature osteocytes (*7*, *16*–*18*). We conclude that *MYC* inactivation causes osteogenic sarcoma cells to differentiate into mature osteocytes.

To examine the effects of *MYC* inactivation and reactivation in individual tumor cells, we cultured osteogenic sarcoma cells

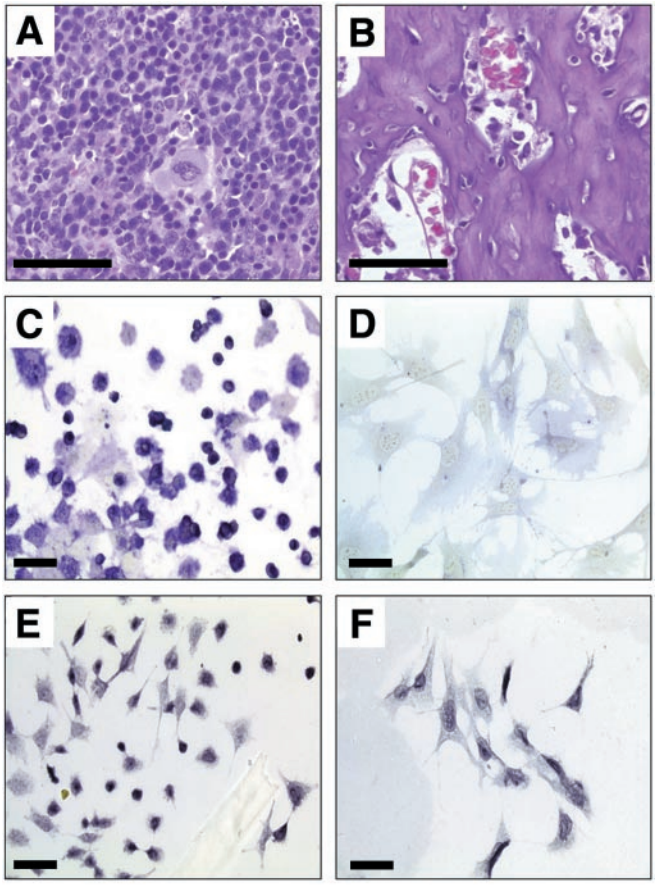


Fig. 1. Inactivation of *MYC* causes regression and differentiation of tumor cells. *MYC* inactivation resulted in the differentiation of (**A**) osteogenic sarcomas into (**B**) mature osteoid. Alkaline phosphatase activity (**C**) before and (**D**) after dox treatment. Osteopontin expression (**E**) before and (**F**) after dox treatment. Representative data from one of five experiments. At least five mice were injected per experiment. Similar results were seen for two other transplanted tumors and two independent primary transgenic tumors (*15*). Bars, 50 μm.

[1]Division of Oncology, Departments of Medicine and Pathology, Stanford University, Stanford, CA 94305–5151, USA. [2]Department of Radiation Oncology, University of California, San Francisco, CA 94143–0806, USA. [3]G. W. Hooper Foundation, University of California, San Francisco, CA 94143–0552, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: dfelsher@leland.stanford.edu

Supporting Online Material

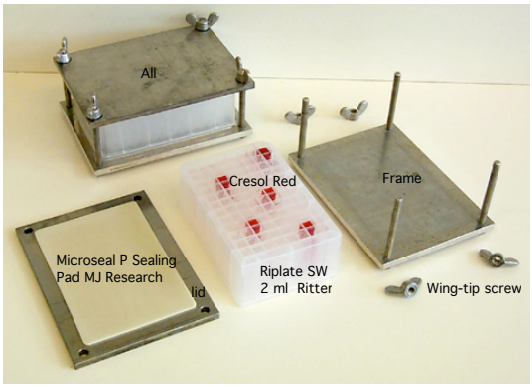
Materials and Methods

Protocol: Genomic DNA isolation

Cut of a piece of tail or muscle of approximately 1.0 cm. Cut up into smaller pieces and put in 1.0 ml lysis buffer in 96 well blocks/plates or any other 96 well format. Put indicator for leakage in some wells (e.g. cresol red). Seal wells with caps and/or pads and heat to 55°C in a waterbath. Tighten the lid as strong as possible. Incubate o/n in 55°C stove with continuous rotation/inverting. Spin down debris: 6000g for 15 min. Transfer most of supernatant to new well/plate. Add 700 µl isopropanol to each well and mix by inverting 10 times. Centrifuge at 6000g for 10-15 min. Discard supernatant (By turning upside down). Wash once with 400 µl 70% ethanol (6000g for 5 min). Discard supernatant. Dissolve pellet in 200 µl H2O or T10E0.1 (pH7.5). (Determine concentration on gel before performing a PCR reaction.) Dilute stock 50 times with H2O. Use 2.5 µl for (nested) PCR.

Lysis Buffer	Final Conc	Stock Conc	**0.5**	**1**	**5**	**10**	**50**	**100**	**500**	**1000**
H20 milliQ			0.410	0.82	4.10	8.2	41.0	82	410	820
Tris HCl pH 8-8.5	100 mM	1 M	0.050	0.10	0.50	1.0	5.0	10	50	100
NaCl	200 mM	5 M	0.020	0.04	0.20	0.4	2.0	4	20	40
SDS	0.2%	10%	0.010	0.02	0.10	0.2	1.0	2	10	20
EDTA	5 mM	0.5 M	0.005	0.01	0.05	0.1	0.5	1	5	10
Proteinase K[1]	100 ug/ml	10 mg/ml	0.005	0.01	0.05	0.1	0.5	1	5	10

[1] add fresh from -20°C



Protocol: Cryopreservation of sperm from isolated zebrafish testis

Materials:

- polystyrene box (at least 30 x 40 x 30 cm) for dry ice
- ice bucket
- transport bucket for liquid nitrogen for 1 – 3 l nitrogen
- 50 ml Falcon tubes
- 2 ml cryovials with screw caps
- 10 ml capillaries 3 cm long (Brand oder Hirschmann)
- Eppendorf Tubes 1,5 ml, sterile
- 5 cm Petri dishes
- 10 ml, 200 ml and 1 ml micropipettes
- Cristal Tips (Eppendorf), 200 ml Tips and 1000 ml Tips
- mouth piece and capillary adaptor for a hose, to fill the capillaries
- tweezers: 2 x pointed, 1 x blunt
- small microscopy scissors
- 30 cm tweezer to sort the cryovials into the nitrogen
- vortex
- binocular
- timer

Solutions:

- MESAB (0,4% ethyl-m-aminobenzoate methanesulfonate + 1% Na_2HPO4 x 2 H_2O, pH 7,2). For usage: 6 ml MESAB-solution added to in 100 ml fishwater.
- Ginzburg Ringer Solution (GRS): 6,5 g NaCl, 0,25 g KCl, 0,4 g $CaCl_2$ x 2 H_2O in 1 l dd H_2O add 0,2 g $NaHCO_3$. Freeze in 4,5 ml portions at – 20°
- Freezing medium: 4,5 ml Ginzburg Ringer Solution + 750 mg milkpowder, fatfree (Carnation milk, Nestle) + 0,75 ml Methanol: The milkpowder has to be resolved into the GRS completely. Therefore vortexing for 15 min is needed. Then add Methanol. Only to on the day of preparation. Keep cold on ice!

Procedure:

Preparation:

- prepare freezing medium
- fill crushed dry ice into the polystyrene box (at least 10 cm height) and put 6 Falcon tubes deep up to the screw cap into the dry ice. Don´t take the Falcons out of the ice during the experiment!
- mark cryovials and cool them in the ice bucket on water ice
- mark the Eppendorf cups, fill 20 ml of icecold freezing medium into the cups and keep them on ice.

Sperm isolation:

- All equipment and material used have to be dry as water activates the sperm!
- anaestize the male fish in the MESAB solution until the gills stop to move
- take the fish out of the solution and dry it very carefully with a tissue paper
- put the fish into a petri dish and decapitate it
- open the body cavity 2 cm to the anal fin
- open up the body cavity and grip the testis with the pointed tweezer below the swim bladder and pull it out.
- transfer the testis into the corresponding Eppendorf cup. Release the sperm by gently flushing and pipetting the testis with the help of a cristal tip. Remove the last 3 mm of the tip before usage!

Sperm freezing:

- fill the sperm solution into capillaries: 2 x 10 ml or 4 x 5 ml.
- transfer the capillaries into the corresponding cryovials on ice. Be careful by fixing the screw cap: put the screw cap gently on the vials, turn them upside down, so that the capillaries fall into the screw cap and fix the cap.
- transfer the cryovials into the falcon tubes, which are buried in the dry ice. Please don´t take the Falcons out of the dry ice, only unscrew them! 2 cryovials can be kept in one Falcon tube, but make sure that they are not over each other, but next to each other. Beginners should only put one cryovial per Falcon tube. Freeze the vials 30 min on dry ice.
- transfer the cryovials into liquid nitrogen (transport bucket)
- sort the cryovials into the nitrogen container.

Protocol: In vitro fertilization

Materials:

- transport bucket for liquid nitrogen for 1 – 3 l nitrogen
- 2 ml cryovials with screw caps filled with capillaries containing frozen sperm
- 5 cm Petri dishes
- 200 ml and 1 ml micropipettes
- 200 ml Tips and 1000 ml Tips
- mouth piece and capillary adaptor for a hose, to fill the capillaries
- small spatula
- 30 cm tweezer to sort the cryovials into the nitrogen
- vortex
- binocular
- timer

Solutions:

- MESAB (0,4% ethyl-m-aminobenzoate methanesulfonate + 1% Na_2HPO4 x 2 H_2O, pH 7,2). For usage add 4 ml MESAB-Lösung to 100 ml fishwater.
- Hanks final solution (Zebrafish book, S. 10.8 + 10.9):
- Stock 1: 8,0 g NaCl, 0,4 g KCl in 100 ml dd H_2O
- Stock 2: 0,358 g Na_2HPO_4, 0,6 g KH_2PO_4 in 100 ml dd H_2O
- Stock 4: 1,44 g $CaCl_2$ x 2 H_20 in 100 ml dd H_2O
- Stock 5: 2,46 g $MgSO_4$ x 7 H_2O in 100 ml dd H_2O
- Stock 1, 2, 4 und 5 freeze at –20°C in 1 ml or 0,1 ml portions
- Stock 6: 0,35 g $NaHCO_3$ in 10 ml dd H_2O: don´t store! use freshly!
- Hanks Premix: 1 ml Stock 1 + 100 ml Stock 2 + 100 ml Stock 4 + 100 ml Stock 5 fill to 10 ml using 8,6 ml dd H_2O
- Hanks Final: 9,9 ml Hanks Premix + 100 ml Stock 6
- Embryo medium E3 (Haffter et al., 1996)

Procedure:

Preparation:

- On the day before IVF female fish (4 – 12 month old) are seperated from male fish and put into a separate mouse cage. Don´t feed them until the squeezing procedure is over.
- The fish need to be squeezed at least one time before the eggs should be used for IVF

In vitro fertilization:

- *Please make sure that you need not to interrupt between the IVF steps, as the squeezed eggs dry out very quickly and may not survive! All equipment and materials used have to be dry as water activates the sperm.*
- transfer the sperm samples in cryovials in liquid nitrogen to your working place
- anaesthize a female fish with MESAB until the gills stop to move
- take female out of the MESAB solution and dry it very carefully with tissue paper, put the fish into a Petri dish
- Squeeze the eggs by carafully pressing the belly of the female towards the analfin.
- Move the eggs by using the spatula into the Petri dish. Eggs which are ok look yellowish and keep together. Eggs looking cloudy or dirty and which are less then 50 should not be used for IVF.
- Put the squeezed females into a separate mousecage with fishwater
- pipet a drop of 70 ml Hanks final solution next to the eggs
- use the 30 cm tweezer to get the cryovials out of the liquid nitrogen. Remove one capillary and thaw it
- blow the content of the capillary carefully into the Hanks final drop and mix with the capillary. Mix the suspension carefully with the eggs and incubate for 30 sec.
- add 750 ml fishwater to the eggs and incubate 2 min
- add 9 ml of fishwater and incubate for 6 h at 28°C
- sort the fertilized eggs into E3 – Embryo – Medium and calculate the fertilization rate.

Nested PCR

In order to get high purity PCR products suited for direct sequencing without additional purification steps nested PCR reactions are performed. The first PCR reaction is a standard reaction. In the nested PCR reaction the amount of primers and dNTPs are optimized for maximal yield with minimal input. This should be done for each primer pair used. Conditions given below are standard conditions for amplification of an approximately 500 bp fragment. All PCR reactions are performed in 384 well plates. A small volume (< 0.2 ul) from the first PCR reaction is used as template in the nested PCR reaction. This volume is transferred with 384 well replicators.

1st PCR

Amounts			PCR program		
Component	1 sample	384w plate	step	Temperature	Time (s)
DNA	2.499	999.6	1	94	120
primer1 (10 µM)	0.15	60	2	92	20
primer2 (10 µM)	0.15	60	3	58	30
buffer (10 X)	0.75	300	4	72	30
MgCl2 (50 mM)	0.225	90	5	go to step 2, 29 times	
dNTPs (10 mM)	0.15	60	6	72	180
Taq (5U/µl)	0.03	12	7	15	forever
H20	3.546	1418.4			
Total	7.5	3000			

Nested PCR

Amounts			PCR program		
Component	1 sample	384w plate	step	Temperature	Time (s)
DNA	0	0	1	92	20
primer1 (10 µM)	0.03	12	2	58	30
primer2 (10 µM)	0.03	12	3	72	30
buffer (10 X)	0.3	120	4	go to step 1, 29 times	
MgCl2 (50 mM)	0.09	36	5	72	180
dNTPs (10 mM)	0.012	4.8	6	15	forever
Taq (5U/µl)	0.024	9.6			
H20	2.514	1005.6			
Total	3	1200			

DNA sequencing

Sequencing reactions:
In principle, standard sequencing reactions are performed, except for the fact that everything is scaled down and optimized for each amplicon. The amount of BigDye Terminator used is in the range of 0.25 – 0.35 ul per reaction, dependent on the amlicon. Template for sequencing is diluted PCR product from the nested PCR (11 times diluted).

Sequence Reaction					
Amounts			PCR program		
Component	1 sample	384w plate	step	Temperature	Time (s)
PCR product (11 x diluted)	1	390	1	95	10
Primer (10 μM)	0.5	195	2	50	5
BigDye Terminator mix	0.3	117	3	60	150
Sequence dilution buffer (2.5 x)	3.7	1443	4	go to step 1, 49 times	
DMSO (50%)	1	390	5	4	forever
H_2O	3.5	1365			
Total	10	3900			

2.5x sequence dilution buffer: 200mM Tris-HCl, pH9.0 containing 10 mM $MgCl_2$. Thermal ramp is 1°C/sec.

Purification of sequencing reactions:
- Sequence samples are purified using G50 superfine minicolumns in Millipore multiscreen 96-well filtration plates. This is done to manufacture's instructions with slight modifications:
- Load dry Sephadex G-50 Superfine into all 96 wells of a MultiScreen HV plate as follows:
  - Add Sephadex G-50 Superfine to the column loader.
  - Remove excess of resin from the top of the column loader with the scraper
  - Place Multiscreen HV plate upside-down on top of the column loader
  - Invert both MultiScreen HV plate and column loader
  - Tap on top of the column loader to release the resin
- Add 300 μl MQ H_2O to each well to swell resin. Incubate at room temperature for 3 hours.
- Place a centrifuge alignment frame on top of a standard 96-well microtiter plate, then place the HV plate on the assembly and centrifuge at 910 x g for 5 min at 15°C to pack the columns.
- Wash the columns once with 150 μl MQ H_2O and centrifuge at 910 x g for 5 min at 15°C.
- Pipette the samples on top of the columns.
- Place the HV plate on top of a 96 well PCR plate and centrifuge at 910 x g for 5 min. at 15°C.
- Change the orientation of the plates, in the buckets, 180 degrees and centrifuge again for 5 min.
- Pipette the purified samples into a 384 well PCR plate suited for the ABI 3700 DNA Analyzer.
- Cover plate with aluminium foil tape (3M Scotch 431, 75 mm). The plate is now ready for direct loading on the 3700 or you can store your samples at -20°C. Don't throw away the MultiScreen HV filterplate. This plate can be used again so clean it thoroughly!

Sequence Analysis on ABI 3700 DNA Analyzer:
Standard conditions are used, except that running time and sheet flow volume are adapted for each desired read length.

VDJ-recombination PCR-assay

To test for VDJ-recombination (Fig. 3 in manuscript), approximately 400 ng genomic DNA, isolated from whole tails, was used as starting template. V-J PCR reactions were done for 4 different classes of immunoglobulin heavy genes (VH). Semi-nested PCR reactions were done in order to obtain bright signals. V-V' PCR-reactions were done to show the presence of intact V segments for the 4 different VH classes. The PCR conditions used in the first and nested PCR are standard with an annealing temperature of 52 ºC and 58 ºC respectively. Primers used for this assay are listed in the table below.

VH class	PCR	Primers first reaction		Primers nested reaction	
1	V-J	VHL1	GATGGACGTGTTACAATTTGG	VHFR1	CCTCCTCAGACTCTGTGGTGA
		FR4	GTTCCYTTHCCCCAGTAGTCAAA	FR4	GTTCCYTTHCCCCAGTAGTCAAA
	V-V'	VHL1	GATGGACGTGTTACAATTTGG	VHFR1	CCTCCTCAGACTCTGTGGTGA
		VHFR3	CGTGCACAGTAATAAACAGCT	VHFR3	CGTGCACAGTAATAAACAGCT
2	V-J	5RL	TTGTAACATGACCATGAATATT	5RFR1	CGATTAGATCAGTCACCTTCT
		FR4	GTTCCYTTHCCCCAGTAGTCAAA	FR4	GTTCCYTTHCCCCAGTAGTCAAA
	V-V'	5RL	TTGTAACATGACCATGAATATT	5RFR1	CGATTAGATCAGTCACCTTCT
		5RFR3r	CGAGCACAGTAATAAACAGCA	5RFR3r	CGAGCACAGTAATAAACAGCA
3	V-J	vh3L	CATGACAATGGATATTGTGTCC	vh3fr1	CTCTGTTGGTGTCAAACACTG
		FR4	GTTCCYTTHCCCCAGTAGTCAAA	FR4	GTTCCYTTHCCCCAGTAGTCAAA
	V-V'	vh3L	CATGACAATGGATATTGTGTCC	vh3fr1	CTCTGTTGGTGTCAAACACTG
		vh3fr3	CTTGCACAATAATATACAGCAG	vh3fr3	CTTGCACAATAATATACAGCAG
4	V-J	vh103L	CAAGATGAAGAATGCTCTCTG	vh103fr1	TGTCAAAGTATGGAGTCGA
		FR4	GTTCCYTTHCCCCAGTAGTCAAA	FR4	GTTCCYTTHCCCCAGTAGTCAAA
	V-V'	vh103L	CAAGATGAAGAATGCTCTCTG	vh103fr1	TGTCAAAGTATGGAGTCGA
		vh103r	TGCACAGTAATACACGGCTGA	vh103r	TGCACAGTAATACACGGCTGA

Chapter 3

Efficient target-selected mutagenesis in zebrafish

Reprinted from Wienholds *et al*. (2003), *Genome Res* 13:2700-2707

Efficient Target-Selected Mutagenesis in Zebrafish

Erno Wienholds, Freek van Eeden, Marit Kosters, Josine Mudde, Ronald H.A. Plasterk, and Edwin Cuppen[1]

Hubrecht Laboratory, The Netherlands Institute for Developmental Biology, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

One of the most powerful methods available to assign function to a gene is to inactivate or knockout the gene. Recently, we described the first target-selected knockout in zebrafish. Here, we report on the further improvements of this procedure, resulting in a highly efficient and easy method to do target-selected mutagenesis in zebrafish. A library of 4608 ENU-mutagenized F_1 animals was generated and kept as a living stock. The DNA of these animals was screened for mutations in 16 genes by use of CEL-I-mediated heteroduplex cleavage (TILLING) and subsequent resequencing. In total, 255 mutations were identified, of which 14 resulted in a premature stop codon, 7 in a splice donor/acceptor site mutation, and 119 in an amino acid change. By this method, we potentially knocked out 13 different genes in a few months time. Furthermore, we show that TILLING can be used to detect the full spectrum of ENU-induced mutations in a vertebrate genome with the presence of many naturally occurring polymorphisms.

[Supplemental material is available online at www.genome.org.]

Over the years, the zebrafish has proven to be an excellent vertebrate model organism for studying many aspects of human biology and disease. The development of the embryo outside of the mother and the full transparency of the embryo, make it well suited for studying, for example, early developmental processes and organ formation. More importantly, however, due to its small size, large number of offspring, and relative short generation time, the zebrafish is well suited for genetic studies. At present, many dominant and recessive forward genetic screens have resulted in thousands of phenotypic mutants (Driever et al. 1996; Haffter et al. 1996), for which the genes involved are now being cloned.

Currently, the genome sequence of the zebrafish is being elucidated and a good draft version is expected soon. At that moment, we will know nearly the complete set of genes encoded by the zebrafish genome, leaving us with the question of what the function is for most of these genes. Although the forward genetic studies will undoubtedly result in extremely valuable data to this end, cloning of the responsible genes is still laborious. Therefore, reverse genetics techniques or knockout technology may become increasingly important tools for revealing gene functions.

A few years ago, RNA interference seemed to be a promising new approach for specifically inactivating genes in zebrafish (Wargelius et al. 1999; Li et al. 2000), but success rates in zebrafish were low and varying, with reports of major nonspecific effects on embryonic development reported as well (Oates et al. 2000; Zhao et al. 2001). The most widely used reverse genetics method in zebrafish is undoubtedly the use of morpholinos that are targeted to the translation initiation site of a specific transcript, thereby inhibiting its translation (Nasevicius and Ekker 2000). The disadvantage of this method is that it is a transient method, requiring repeated injection for each experiment, and only suited for early developmental stages. In addition, nonspecific side effects, like widespread cell death and neuronal degeneration, have been reported (for review, see Heasman 2002).

Attempts to implement a method for generating permanent knockouts similarly (as routinely done for the mouse) have not been successful yet for zebrafish. Such an approach requires homologous recombination in pluripotent embryonic stem (ES) cells and subsequent generation of chimeric embryos. Although ES-like cells for zebrafish have been described (Sun et al. 1995), and the generation of some chimeras from wild-type cells has been reported (Ma et al. 2001), no targeted knockouts have been obtained using this approach. An alternative to this might be to clone zebrafish by nuclear transfer of genetically modified cultured cells (Lee et al. 2002). However, this has also not resulted in a targeted knockout.

Recently, we have shown that reverse genetic analysis of gene function in zebrafish is possible by target-selected mutagenesis (Wienholds et al. 2002); random mutagenesis, followed by screening for mutations in target genes. This general approach, using different types of chemical mutagens and mutation-detection methods, has proven to be successful in *Caenorhabditis elegans* (Jansen et al. 1997), *Drosophila* (Bentley et al. 2000), plants (*Arabidopsis* and *Lotus*) (McCallum et al. 2000; Perry et al. 2003), mouse (Beier 2000; Coghill et al. 2002), and rat (Zan et al. 2003). For target-selected mutagenesis in zebrafish, male zebrafish are mutagenized using *N*-ethyl-*N*-nitrosourea (ENU), identical to that done for forward genetic screens. These males are used to generate a large population of F_1 animals that consequently harbor many random heterozygous mutations in their genomes. Next, DNA from these animals is analyzed for the occurrence of mutations in a specific gene of interest. Proof of principle for the zebrafish was obtained by screening a library of DNA samples and matching frozen sperm samples derived from 2679 individual fish. By DNA resequencing, 15 mutations were identified in the *rag1* gene. One of these introduced a premature stop codon in the essential core domain of the recombinase, resulting in a complete loss-of-function phenotype, as demonstrated by the lack of recombination of *V(D)J* segments at the immunoglobulin locus in animals that were homozygous for this mutation (Wienholds et al. 2002).

Here, we describe major modifications to the original method, now allowing efficient generation of knockouts in zebrafish. Firstly, a larger library of ENU-mutagenized animals was constructed and kept alive during the screening for mutations. Secondly, as resequencing is both laborious and expensive, we implemented an alternative method for detection of mutations—TILLING (targeting induced local lesions in genomes; McCallum et al. 2000). This method is based on enzymatic cleavage of het-

[1]Corresponding author.
E-MAIL ecuppen@niob.knaw.nl; FAX 31-30-2516554.
Article and publication are at http://www.genome.org/cgi/doi/10.1101/gr.1725103. Article published online before print in November 2003.

eroduplex DNA using the plant endonuclease CEL-I (Oleykowski et al. 1998). TILLING is used successfully for a similar objective in *Arabidopsis* (Colbert et al. 2001; Till et al. 2003). We report the identification of 255 mutations, including 21 that are most likely to result in a loss-of-function of 13 different genes. We thus generated 13 different potential knockout fish in a few months.

METHODS

Zebrafish Housing and Mutagenesis

Forty adult TL male zebrafish (4-months-old) were mutagenized by six consecutive treatments with 3.0 mM ENU as described (van Eeden et al. 1999). Of these, 28 surviving and fertile fish were outcrossed with heterozygous *albino* (*alb/+*) females, derived from a hybrid München/AB background for the single locus test, or with TL females to generate F_1 progeny for the library. To construct a library of 4608 mutagenized fish, both healthy-looking F_1 males and females were finclipped (tail) and grouped in 384 pools of 12 fish per tank at the age of from 6 to 8 mo. To recover the fish carrying a specific mutation, all 12 fish from a positive pool were finclipped again, housed separately, and genotyped. Subsequently, carriers were outcrossed against AB or München/AB hybrid fish.

Genomic DNA Isolation and DNA Library Construction

Genomic DNA for the library was isolated from finclips in deep 96-well plates (1 mL capacity per well) for the library and single tubes for the genotyping. Freshly cut finclips were transferred directly to plates/tubes that are kept on dry ice. Fins were lysed by overnight incubation at 55°C in 400 µL pre-warmed lysis buffer (100 mM Tris-HCl at pH 8–8.5, 200 mM NaCl, 0.2% SDS, 5 mM EDTA, and 100 µg/mL proteinase K). DNA was precipitated by adding 300 µL of isopropanol and centrifugation at >6000*g*, and washed using 70% ethanol. Finally, pellets were dissolved in 1 mL of water. For the DNA library, 5-µL aliquots were arrayed in 384-well PCR plates using a 96-channel pipettor (HYDRA-96, Robbins Scientific), covered with aluminum foil tape (3M) and stored at −20°C.

CEL-I-Mediated Library Screening

CEL-I enzyme was isolated from celery according to Oleykowski et al. (1998) and Yang et al. (2000), with minor modifications (see our Web site for a detailed protocol, http://cuppen.niob.knaw.nl). The enzyme activity for each batch of CEL-I was determined experimentally using a dilution series on control samples. Screening for ENU-induced mutations was done using CEL-I mediated heteroduplex cleavage, analogous as described for *Arabidopsis* (Colbert et al. 2001), but with several adaptations, as described below. All pipeting steps were done on a Genesis Workstation 200 (Tecan) and Microlab 2200 (Hamilton), or using multichannel pipets. Target genes were amplified by a nested PCR approach in 384-well plates. In the first PCR with gene-specific primers, a touchdown cycling program was used (94°C for 60 sec; 30 cycles of 94°C for 20 sec, 65°C for 30 sec with a decrement of 0.5°C per cycle, and 72°C for 60 sec; followed by 10 cycles of 94°C for 20 sec, 58°C for 30 sec, and 72°C for 60 sec, and an additional extension step of 72°C for 180 sec; GeneAmp9700, Applied Biosystems). PCR samples contained 5 µL of genomic DNA isolated from finclips, 0.2 µM forward (f1) and 0.2 µM reverse (r1) primer, 200 µM of each dNTP, 25 mM Tricine, 7.0% Glycerol (m/v), 1.6% DMSO (m/v), 2 mM $MgCl_2$, 85 mM NH_4Acetate (pH8.7), and 0.2 U Taq DNA polymerase in a total volume of 10 µL.

After the first PCR reactions, the samples were diluted with 20 µL of water, and 1 µL was used as template for the second, nested PCR reaction. This reaction contained a mixture of gene-specific forward (M13F-f2, 0.08 µM) and reverse (M13R-r2, 0.04 µM) primers that contain universal M13 adaptor sequences at their 5′ end, and the two corresponding universal M13F (5′-TGTAAAACGACGGCCAGT; 0.1 µM) and M13R (5′-AGGAAACAGCTATGACCAT; 0.1 µM) primers labeled with fluorescent dyes (IR Dye 700 and IR Dye 800, respectively) for detection. In addition, the PCR samples contained 200 µM of each dNTP, 20 mM Tris-Hcl (pH8.4), 50 mM KCl, 1.5 mM $MgCl_2$, and 0.1 U Taq DNA polymerase in a total volume of 5 µL. Standard cycling conditions were used for the nested PCR reactions (30 cycles of 94°C for 20 sec, 58°C for 30 sec, and 72°C for 60 sec, followed by an additional extension step of 72°C for 180 sec).

Directly following the nested PCR, heteroduplex formation was done by incubating at 99°C for 10 min, and 70 cycles of 70°C for 20 sec with a decrement of 0.3°C per cycle. Next, 1.25 µL of aliquots of four individual PCR reactions were pooled (total volume of 5 µL) and incubated with 0.01 µL of CEL-I enzyme solution in a total volume of 15 µL (buffered in 10 mM Hepes at pH7.0, 10 mM $MgSO_4$, 10 mM KCl, 0.002% Triton X-100, 0.2 µg/mL BSA) at 45°C for 15 min. CEL-I reactions were stopped by adding 5 µL of 75 mM EDTA. Fragments were purified using Sephadex G50 (medium coarse) minicolumns in 96-well filter plates (Multiscreen HV; Millipore) and eluted into plates prefilled with 5 µL of formamide loading buffer (37% [v/v] de-ionized formamide, 4 mM EDTA [pH8.0], 90 µg/mL bromophenol blue) per well. Samples were concentrated to about 1 µL by heating at 85°C for 45–60 min without cover. A total of 0.4 µL was applied to a 96-lane membrane comb (The Gel Company) and loaded on 25 cm of denaturing 6% polyacrylamide gels on LI-COR 4200 DNA analyzers. Raw TIFF-images produced by the analyzers were manipulated using Adobe Photoshop, and potential mutations were detected and scored manually.

Genotyping by Sequencing

For genotyping, PCR reactions were performed using the same conditions as for the CEL-I-mediated screening. In the nested PCR, only gene-specific primers (M13F-f2 and M13R-r2) were used at 0.2 µM. Nested PCR products were diluted with 30 µL of water, and 1 µL was used as template for the sequencing reactions. Sequencing reactions contained 0.5 µL of DYEnamic ET Terminator (Amersham Pharmacia Biotech), 3.5 µL of ET Terminator dilution buffer (Amersham Pharmacia Biotech), and 0.5 µM of M13F or M13R primer in a total volume of 10 µL. Cycling conditions were as recommended by the manufacturer. Sequencing products were purified using Sephadex G50 (superfine coarse) minicolumns, and analyzed on a 96-capillary 3700 DNA analyzer (Applied Biosystems).

RESULTS

Generation of a Living Library of Mutagenized F_1 Fish

The success rate of both forward and reverse mutagenesis screens is mainly dependent on two factors, mutagenesis efficiency and the number of individuals screened. To find loss-of-function mutations in target genes, both factors should be as high as possible. In our first target-selected mutagenesis screen, we were successful in finding a knockout in the zebrafish *rag1* gene (Wienholds et al. 2002). However, the library of mutagenized F_1 fish we used only comprised 2679 individuals and had a molecular mutation frequency of about one mutation per 450,000 bp (Wienholds et al. 2002). These small numbers make the odds for finding at least one loss-of-function mutation in any average target gene relatively small. Therefore, we generated a novel and larger library. Founder animals were heavily mutagenized with ENU and 4608

healthy F_1 male and female fish were raised and used to construct the library (Fig. 1). In contrast to the previous library, we did not generate a permanent resource by cryopreservation of sperm or testis samples. Instead, we maintained the library as a living resource that can be used for a limited time. There are several reasons for this. First of all, the cryopreservation of testis or sperm samples is labor intensive, taking several months to complete, whereas the construction of a living library only took 2 wk to complete. Secondly, the success-rate for the recovery of fish carrying mutations from frozen testis samples by in vitro fertilization (IVF) turned out to be highly variable, in some cases resulting in loss of alleles (E. Wienholds, unpubl.). To minimize the space requirements of a living library, the fish were grouped in 384 pools of 12. After identification of an interesting mutation, all the fish of the positive pool were finclipped and genotyped again to recover the carrier of this mutation.

Single locus tests were done to calculate the mutation frequency of the mutagenized founder fish. In total, we found 11 albinos in 3427 mutagenized genomes screened. This is comparable with previous efficient mutagenesis screens (Mullins et al. 1994). Preliminary data from a small-scale forward F_3 screen with a subset of the animals also suggests that the mutagenesis efficiency for this new library is high (F. van Eeden, unpubl.). All fish of the library were screened for mutations in several target genes. The average per-base mutation frequency for these genes was 1 in 235,000 bp (Table 1). This is approximately twofold higher than for the previous library.

Amplicon Selection and Primer Design

As no fully assembled zebrafish genome is available yet, PCR amplicons for genes of interest were designed from cDNA or EST

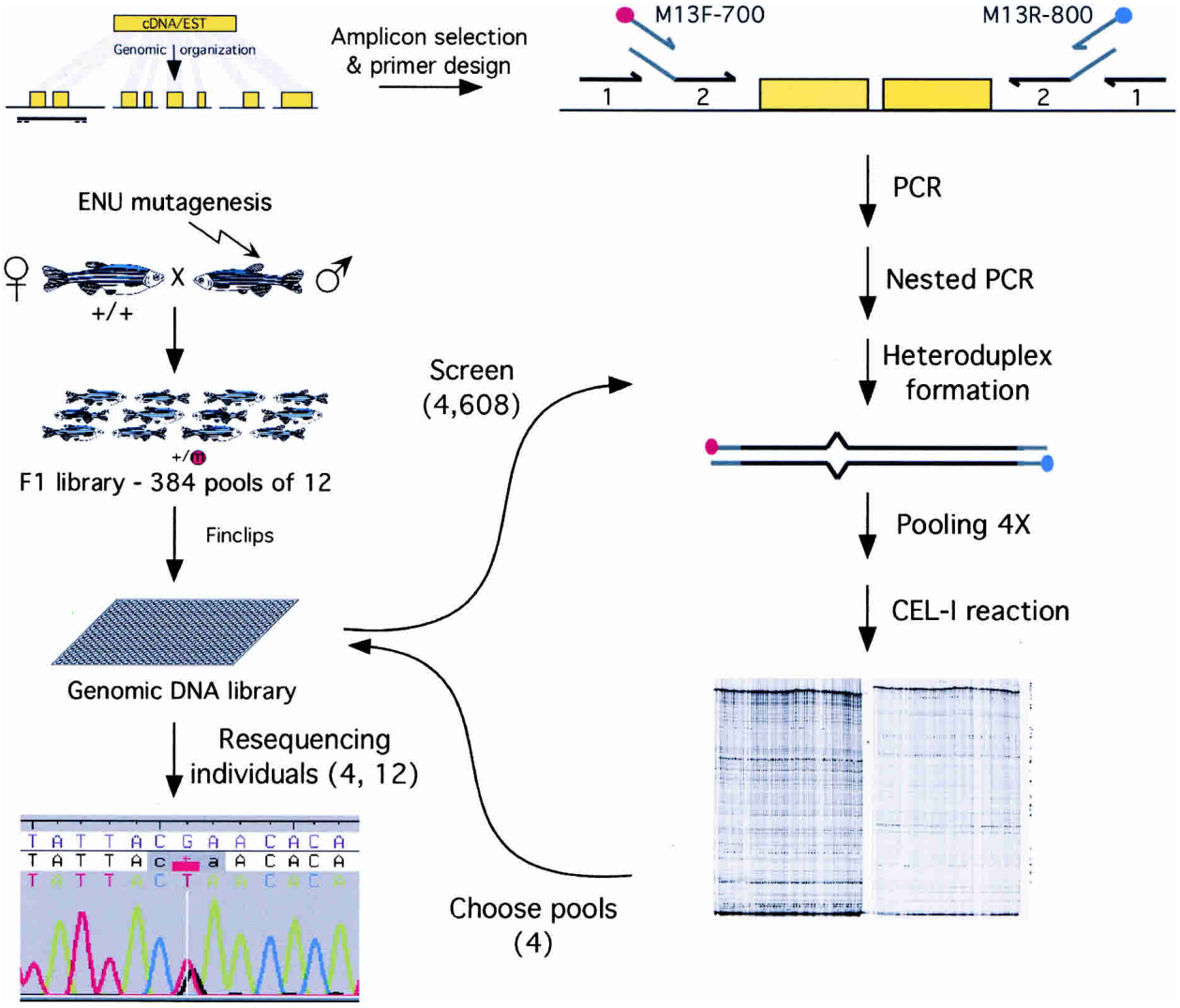


Figure 1 Efficient target-selected mutagenesis in zebrafish. Forty male zebrafish were mutagenized with ENU and outcrossed with wild-type females to generate a library of 4608 mutagenized F_1 fish. Both males and females were finclipped and grouped in 384 pools of 12 fish per fish tank. DNA was isolated from the finclips and arrayed in twelve 384-well PCR plates. Amplicons of target genes were selected from local gene assemblies using GENOTRACE and primer picking software. Targets were amplified by PCR with gene-specific primers (1), followed by a nested PCR with internal gene-specific primers containing universal adaptor sequences (2), in combination with IR DYE-labeled universal M13 primers. Heteroduplexes were formed and samples were pooled fourfold. Pooled samples were incubated with CEL-I enzyme, and fragments were analyzed by denaturing polyacrylamide gel electrophoresis. Steps from PCR amplification to CEL-I incubation can be done in a completely unattended robotic setup. The four samples represented in a positive pool were reamplified from genomic DNA and subsequently resequenced. Fish carrying interesting mutations were recovered from pools of the F_1 library by finclipping and resequencing each of the 12 fish of a pool.

Table 1. Summary of Amplicons Screened and Mutations Found

Gene		Amplicon		Screening		No. of mutations							
Name	Cds (bp)	Name	Length/cds[c] (bp)	Success (%)	Screened (Mb)	Total	Recovered	Nonsense	Splice	Missense	Silent	Noncoding	Mutation frequency (bp/mutation)
gene1	690	A	569/264	88	2.3	16	6	—	—	—	1	5	384,553
gene2	1434	A	777/381	74	2.6	28	15	—	1	6	2	6	176,634
gene3	1821	A	639/327	80	2.4	22	18	1	2	7	3	5	130,867
gene4	4860[a]	A	911/879	81	3.4	21	14	1	—	8	5	—	242,878
gene5	1239	A	812/327	76	2.8	24	18	2	1	2	3	10	157,983
gene6	951	A	541/436	33[d]	0.8	6	6	1	—	2	3	—	137,111
gene7	2271[b]	A	590/264	68	1.8	24	6	—	1	2	1	2	308,122
gene8	3240[b]	A	442/270	79	1.6	14	6	—	1	2	2	1	268,170
gene9	1122	A	841/279	78	3.0	34	20	—	—	7	2	11	151,138
		B	448/230	81	1.7	9	7	—	1	2	2	2	238,879
gene10	1410	B	664/395	64	2.0	8	4	1	—	3	—	—	489,554
gene11	4110	A	720/720	87	2.9	27	11	—	—	6	5	—	262,405
gene12	480	A	496/469	87	2.0	13	6	—	—	3	3	—	331,407
gene13	2151	A	816/443	75	2.8	17	9	1	—	4	1	3	313,344
gene14	8532[b]	A	953/953	83	3.6	28	16	—	—	14	2	—	227,805
		B	873/873	80	3.2	22	13	1	—	7	5	—	247,556
		C	813/813	91	3.4	20	10	—	—	5	5	—	340,914
		G	921/881	87	3.7	22	11	1	—	6	4	—	335,659
gene15	6912[b]	A	955/955	89	4.4	25	18	2	—	12	4	—	244,480
gene16	6524	A	798/775	90	3.3	26	17	2	—	9	6	—	194,674
		C	752/418	93	3.2	32	16	1	—	8	2	5	201,416
		E	736/526	81	2.7	20	8	—	—	4	2	2	343,388
Average	2984		730/540	79	2.7	20.8	11.6	0.6	0.3	5.4	2.9	2.4	234,625
Total	47,747		16,067/11,878		59.8	458	255	14	7	119	63	52	

[a]Length of complete coding sequence/cds is based on *Drosophila melanogaster* homolog.
[b]Length of complete cds is based on *Homo sapiens* ortholog.
[c]Length of nested PCR amplicon, excluding primer sequences; length of cds in amplicon.
[d]Library not completely screened for this amplicon; false positives not counted.

sequences. These sequences were used as input for GENOTRACE (Berezikov et al. 2002; http://genotrace.niob.knaw.nl) to determine the genomic organization by use of whole-genome sequencing trace archives and retrieve noncoding sequences flanking the exons. These noncoding sequences were used for primer design to amplify the coding sequences (Fig. 1). Preferably, amplicons were chosen such that the coding sequences from the 5′ end of the gene were amplified, maximizing the chance that a premature stop codon results in a full loss-of-function. Nested sets of primers were designed automatically using a specially designed PRIMER3-based (Rozen and Skaletsky 2000) Web application (http://primers.niob.knaw.nl). Using this setup, we have designed nested sets of primers for 29 amplicons, of which 22 turned out to be suitable for screening the library. Failure of the other seven amplicons is most likely due to the repetitive nature of the zebrafish genome and the large amounts of single nucleotide polymorphisms (SNPs; ~1 in 200 bp; data not shown). The size of the amplicons that did work varied from 442 to 955 bp, with an average of 730 bp, including 450 bp of coding sequence (Table 1). For most target genes, we designed and screened only one amplicon, resulting in only 25% of the total coding sequences screened (11,878 bp of 47,747 total, Table 1).

High-Throughput Screening of the Mutant Library by TILLING

DNA from the 4608 F_1 fish was isolated from finclips and arrayed in twelve 384-well PCR plates for automated robotic mutation discovery analysis. Although we initially used resequencing for mutation detection (Wienholds et al. 2002), here we switched to an enzyme-mediated heteroduplex cleavage approach (Oleykowski et al. 1998), because this is both faster and cheaper than sequencing. We adapted the protocol for CEL-I mediated heteroduplex cleavage from the originally described approach for mutation discovery in *Arabidopsis* (TILLING; Colbert et al. 2001) and introduced some minor modifications (Fig. 1). Firstly, we introduced a nested PCR to avoid the influence of the variation in quality and quantity of the genomic DNA isolated from finclips. Secondly, in the second PCR, we use a mixture of gene-specific primers with M13 adapters and fluorescently labeled universal M13 forward and reverse primers to reduce the costs for fluorescently labeled primers. Thirdly, after PCR amplification of individual samples, products are denatured and reannealed to form heteroduplexes, pooled only fourfold instead of eightfold, and subsequently treated with CEL-I. We found that due to the presence of many SNPs in the zebrafish strains that we used, further pooling might decrease the sensitivity and result in an increase of false negatives. The final CEL-I-digested and purified samples resulting from a single PCR amplicon for the whole library were analyzed on 12 denaturing polyacrylamide slab gels with 96 samples, representing 384 animals per gel (e.g., Fig. 2). The amplicons of the four samples of the positive pool were subsequently reamplified from genomic DNA, and the nature of the mutation was determined by resequencing (Fig. 1). We have set up a high-throughput mutation discovery pipeline that consists of a liquid handling robot with four 384-well PCR blocks incorporated, which runs the PCR amplification, pooling, and CEL-I digestion steps completely unattended within 20 h. Purified samples are subsequently analyzed on four LICOR analyzers that can be run at least three times a day. This setup can be operated by a single person, and allows the initial screening of the complete library for one amplicon of maximal 1000 bp within a single day.

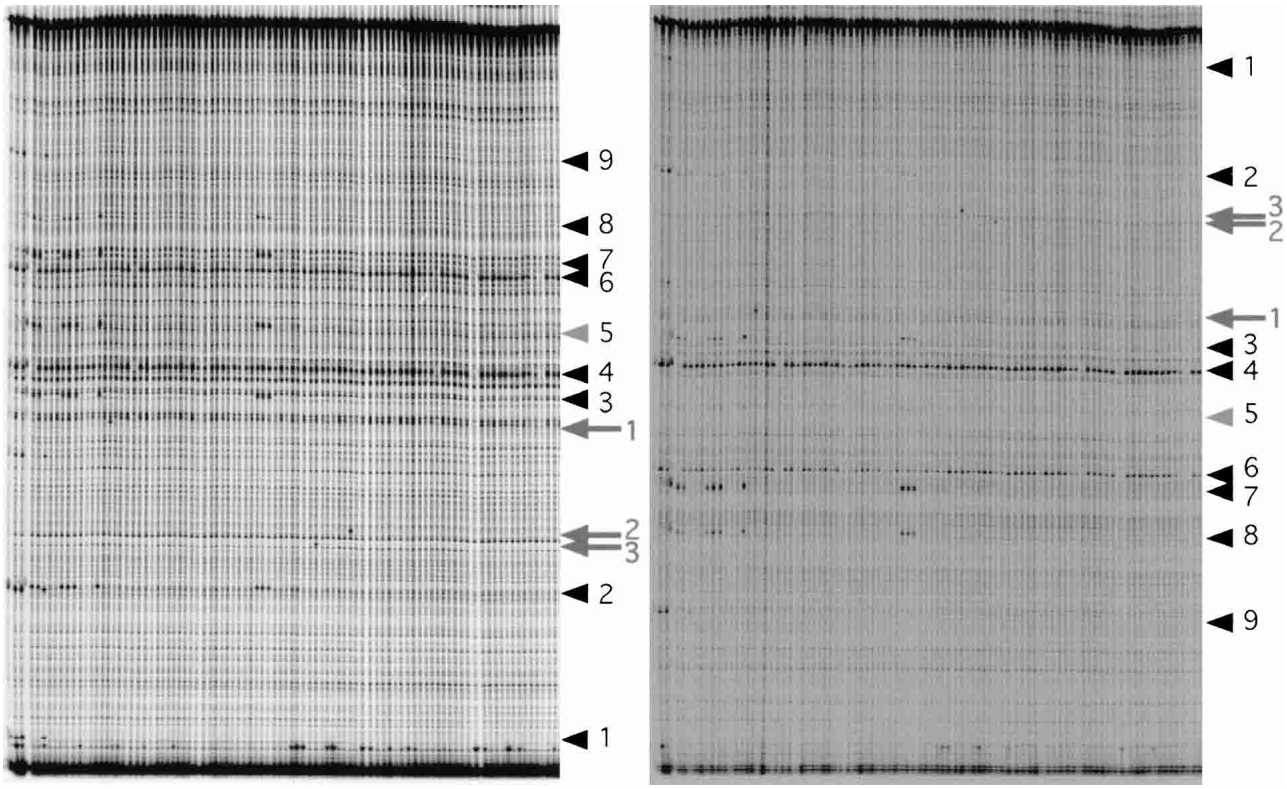


Figure 2 Representative example of a gel used to screen 384 fish for mutations in the *dicer* gene. Each of the 96 lanes of the polyacrylamide gel contains fourfold-pooled CEL-1-digested PCR products of ~830 bp. The IR Dye 700 and IR Dye 800 channels are shown at *left* and *right*, respectively. Arrowheads indicate nine SNPs present in the founder fish. SNP no. 5 is only visible in one channel, and therefore, indicated by a gray arrowhead. Arrows indicate the three potential mutations on this gel (all rank 1).

ENU-Induced Mutations Are Efficiently Identified by TILLING

The TILLING method for mutation discovery has been set up for ethylmethanesulfonate (EMS)-induced point mutations in the relative simple genome of *Arabidopsis* (Colbert et al. 2001). The organization of the zebrafish genome is much more complex than the *Arabidopsis* genome. There are many SNPs and repetitive DNA stretches that might negatively influence the discovery of ENU-induced mutations. In addition, the spectrum of ENU-induced mutations is much broader than for EMS (Fig. 3A). To determine the sensitivity of TILLING in identifying ENU-induced mutations in zebrafish, potential mutations were scored according to the criteria in Table 2 (Example is shown in Supplemental Figure 1, available online at www.genome.org). In total, 59.8 million bp was screened by TILLING (Table 1), in which we identified 435 unique potential mutations. These 429 suspected mutations could be distributed over the six different scoring classes (Table 2). Overall, the confirmation rate after resequencing is 59%, resulting in 255 real mutations. As expected, success rates for confirmation decreased from scoring rank 1 to 5. However, it remained sufficiently high in all classes (minimum of 23%), so that in future screens, classification according to these criteria might not be necessary. Aberrant-looking lanes and potential mutations were classified in rank 6. Confirmation of these mutations also turned out to be quite efficient (38%; Table 2), justifying the inclusion of this type of potential mutations in future screens. Thus, any difference observed in the pooled CEL-I-digested samples is worth further investigation by resequencing.

The confirmation rate varies from gene to gene. We do not find a correlation with the number of SNPs within an amplicon, but we cannot exclude an effect of SNPs in the primer regions, which may result in less-efficient PCR and increased Taq polymerase errors. When variation in product intensity per lane is seen, as for example for genes 7 and 1 (data now shown), indicative for an inefficient amplification of one or more of the samples in this pool, confirmation rates decrease.

The CEL-I enzyme has a slight preference for certain types of heteroduplexes (Oleykowski et al. 1998). TILLING has been proven to work nicely for the limited spectrum of ethylmethanesulfonate (EMS)-induced mutations (mainly GC to AT transitions, Colbert et al. 2001; Greene et al. 2003). We inspected whether it also efficiently recognized the full spectrum of ENU-induced mutations. We compared the spectrum found in this screen with the spectrum found by resequencing of the *rag1* and two other genes (Wienholds et al. 2002; E. Wienholds and R.H.A. Plasterk, unpubl.; Fig. 3A). Overall, the spectra are quite similar. Differences do not reflect the CEL-I preferences and are likely statistical variations due to the small number of mutations (25) found in the resequencing screens. This indicates that we do not miss a certain class of mutations. The mutation spectrum is also very similar to the mutation spectrum of the ENU-induced mutations found in forward genetic screens (Knapik 2000; http://zfin.org; Fig. 3A). This indicates that we also do not miss a certain class of mutations that might give a phenotypic change. In conclusion, we think that TILLING is sensitive enough to identify mutations in all classes of the ENU mutation spectrum.

Potential Knockouts by Target-Selected Mutagenesis

We screened for mutations in 16 different target genes, and in most cases, only one amplicon per gene. Due to the genomic

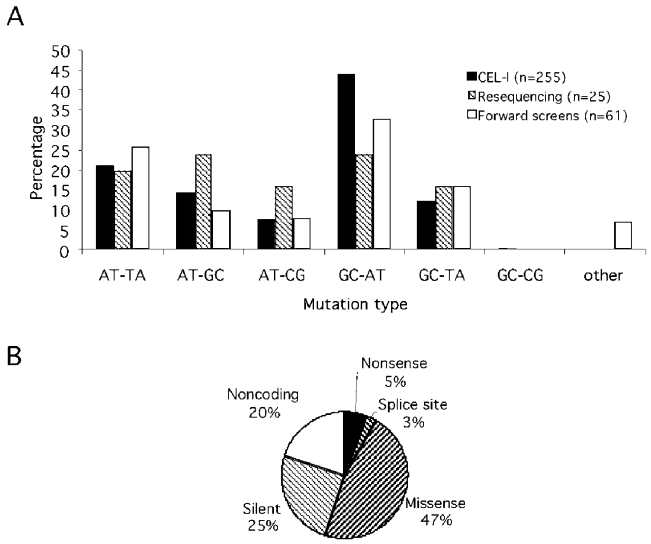


Figure 3 Characterization of the ENU-induced mutations. (*A*) ENU mutation spectrum of this screen is similar to that of resequenced genes (Wienholds et al. 2002; E. Wienholds and R.H.A. Plasterk, unpubl.) and forward screens (Knapik 2000; http://zfin.org). Single nucleotide changes are indicated, and small deletions/insertions and chromosomal rearrangments are summarized under others. (*B*) Distribution of the mutations at the coding level.

organization of most target genes, the amplicons also contained, on average, 26% noncoding sequences (4189 bp of 16,067 bp total; Table 1), like introns and untranslated regions. However, of the 255 confirmed mutations (Table 1; Fig. 3B), 52 (20%) were found in these noncoding sequences. ENU mutagenesis is expected to be random for coding and noncoding sequences. The discrepancy observed here (6%) is most likely caused by the fact that most of the noncoding sequences are located near the amplicons ends, in which mutations are more difficult to detect (Greene et al. 2003). Seven of these noncoding mutations (3%) affected splice donor or acceptor consensus sequences, and are likely to result in splice pattern alterations. Of the 196 mutations found in the coding regions, 63 (25%) are silent, 119 (47%) are missense, resulting in a variety of amino acid changes, and 14 (5%) are nonsense. These results are in full agreement with the changes that can be calculated from codon usage in zebrafish (Nakamura et al. 2000) combined with the molecular ENU spectrum found in our screen (data not shown). The mutations most likely to result in complete loss-of-function of the genes are the nonsense and splice donor/acceptor site mutations. In total, we found 21 of these types of mutations in 13 of the 16 target genes screened.

DISCUSSION

Recently, we have shown proof of principle that target-selected mutagenesis is possible in the zebrafish by screening ENU-mutagenized animals for point mutations (Wienholds et al. 2002). In this study, we implemented several changes to the original procedure to make it even more efficient, resulting in perhaps the method of choice for making knockouts in zebrafish.

First of all, we constructed a larger library of heavily mutagenized F_1 fish. These fish were kept alive during the screening process. To minimize the space requirements, the fish were grouped in 384 pools of 12. As a consequence, the fish of individual pools had to be finclipped and genotyped again before we could recover the carrier of an interesting mutation. However, this can be done in a single day, which is diminishable in the timescale for creating a homozygous knockout. The carrier can be used immediately for large-scale outcrosses, gaining enough fish in the next generation to perform linkage analysis. Both males and females have been used to construct the library. This led to an approximately twofold reduction in the initial animals needed if compared with a library of frozen sperm or testis samples. In addition, interesting mutations found in opposite sexes can be crossed together straight away. For one of the target genes, we found two different, independent nonsense alleles, both disrupting most of the protein. Because these different mutations were found in a male and female, respectively, we were able to do phenotypic analysis of transheterozygotic embryos (without interference of homozygous background mutations) within 1 wk after recovery of the carriers.

A major holdback for creating a living library is that the library can only be screened for a limited time. This starts at the moment that fish can be finclipped (~2–3 mo) up to their fertile lifetime (~1.5 yr). This implies a scheduled mutagenesis and library construction each 1.5 yr. Because construction of the library only took us 2 wk to complete (mutagenesis not included), this is a fast alternative for constructing a permanent (cryopreserved) library, providing the availability of sufficient space. The latter can be addressed by the construction of several small libraries (e.g., 384 animals) at a regular interval. These libraries should then be screened for as many target genes as possible before a new library is ready for screening. Recovery of mutant animals from frozen sperm of the first library turned out to be very variable. In some cases, we were not able to recover a fish at all and lost the mutant allele (E. Wienholds, unpubl.). Keeping the library alive might circumvent this problem. However, diseases and other factors might cause fish to die during the screen-

Table 2. Mutation Quality Classification

Class	Scoring criteria	No. of mutations observed by TILLING	No. of mutations confirmed by sequencing
Rank 1	Strong signal in channel 1, strong signal in channel 2	110	97 (88%)
Rank 2	Strong signal in channel 1, weak signal in channel 2	74	50 (68%)
Rank 3	Strong signal in channel 1, no signal channel 2	66	40 (61%)
Rank 4	Weak signal in channel 1, weak signal in channel 2	40	16 (40%)
Rank 5	Weak signal in channel 1, no signal channel 2	47	11 (23%)
Rank 6[a]	Aberrant	92	35 (38%)
Total[b]		435	255 (59%)

[a]Lanes in gel that look different than normal, contain multiple signals, have a signal at the same height as a background signal, etc.
[b]Including six unclassified mutations from *gene6*.
Gel files as shown in Fig. 2 were inspected manually and a rank was assigned to each of the potential mutations as observed by TILLING. Potential mutations were confirmed by resequencing.

ing process. We did see a substantial fraction of the library dying (~25% over a half-year time period) in one part of the library. This was most likely caused by overcrowding and insufficient fresh water inflow, presumably resulting in toxification of the water. Improvement of housing conditions should abolish this problem in future libraries.

In the original method for target-selected mutagenesis in zebrafish, we identified mutations by resequencing (Wienholds et al. 2002). To both speed up and reduce the costs of target-selected mutagenesis in zebrafish, we incorporated the TILLING method (Colbert et al. 2001) for initial mutation detection. Subsequently, potential mutations were then confirmed by resequencing. We set up a high-throughput robotic pipeline that allows the screening of the complete library by TILLING for mutations in one amplicon within 1 d by a single person. To even further reduce the costs, nested PCR reactions with gene-specific primers coupled to universal adaptors and fluorescence-labeled universal primers were performed. Compared with sequencing, this reduced the costs ~10-fold. Pooling samples more than fourfold (like eightfold in *Arabidopsis*; Colbert et al. 2001) might further speed up the screening process and reduce the costs. However, due to the presence of many SNPs and repetitive DNA stretches, it might be more difficult to detect the mutations, and therefore, the number of missed mutations might increase considerably. Although we used a high-throughput robotics setup for finding mutations in target genes by TILLING in a short time, this approach can also be adapted for use on a smaller scale by use of manual pipeting.

TILLING has originally been set up for the identification of EMS-induced mutations in *Arabidopsis* (Colbert et al. 2001). We found that ENU-induced mutations in zebrafish are also efficiently recognized. From the 435 potential mutations found by TILLING, 255 (59%) were confirmed by resequencing. There is a correlation with signal strength on the polyacrylamide gels and the ability to reconfirm the mutation. But the recovery is sufficiently high, even for the lowest quality class (23%), so that in future screens, any indication of a mutation being present should be followed up. The spectrum of ENU-induced mutations found by TILLING in zebrafish is very similar to the spectrum found by resequencing and forward screens. This indicates that TILLING can be used reliably to detect all classes of mutations introduced by ENU mutagenesis. Only the GC to AT class of transversions is considerably larger in our screen. This may be due to the low quality of DNA for some parts of the library (data not shown), resulting in cytosine deamination (Hofreiter et al. 2001). Although it is unlikely to identify such mutations twice in two completely independent assays (TILLING and resequencing), these types of mistakes end up in our final list of mutations, because this list is based upon the mutations that were mainly confirmed only once in the resequencing phase. To exclude this type of mistake from the list, samples have to be reamplified and sequenced once more. Other small differences with respect to the forward screens could reflect the bias for phenotypic alterations in forward screens. For example, nonsense codons (most likely to have a phenotypic effect) can never be found in the class of AT to GC transitions, but are found mainly in the AT to TA and GC to TA transversions and GC to TA transitions.

In the first target-selected mutagenesis screen in zebrafish, we screened extensively for mutations only in one target gene, the *rag1* gene, and found one nonsense mutation. This turned out to be a loss-of-function allele (Wienholds et al. 2002). Here, we screened for mutations in 16 different target genes. The average molecular mutation frequency for these genes was almost twofold higher (1 in 235,00 bp) than for the *rag1* gene in our previous screen (1 in 450,000 bp). Taking differences in mutagenesis efficiency into account, these results indicate that the success rate for finding mutations using TILLING is at least comparable or maybe even higher than for resequencing. Together with a larger library, this increased the chance of finding detrimental mutations considerably. In 13 of the 16 target genes we screened, we found 21 nonsense and/or splice acceptor/donor site mutations. These mutations are likely to result in loss-of-function or knockouts of the genes. On average, we screened only 25% of the coding regions of these target genes (Table 1). This implies that we could have found many additional nonsense and splice acceptor/donor site mutations if we had screened the complete coding sequences for a specific gene. Here, we focused on amplicons that were most suited for screening by TILLING and for which the PCR reactions worked under standard conditions (22 out of 29 amplicons). Despite the fact that most of these amplicons contained small exons, we were still able to find nonsense and splice acceptor/donor mutations in these exons. This indicates that it was worthwhile to screen these small exons for mutations.

Not only the nonsense and splice acceptor/donor site mutations are likely to be loss-of-function alleles, also several of the 199 missense mutations might be loss-of-function. Similar to forward genetic screens, in which about equal numbers of missense and nonsense mutations are responsible for the observed phenotypes (39% and 48%, respectively; Knapik 2000; http://zfin.org), ~10–20 missense mutations in our set may be expected to result in a phenotypic difference. Furthermore, an allelic series of the remaining mutations may have more subtle, hypomorphic effects on the gene function, and help in the elucidation of specific protein (subdomain) function.

In summary, we have set up a fast, easy, and cheap high-throughput pipeline for making knockouts in zebrafish by target-selected mutagenesis. We show that this setup is well-suited for the detection of the full mutation spectrum induced by ENU in a vertebrate genome, and that the method is sensitive enough to detect novel mutations in the background of natural occurring SNPs, which is important when working with outbred animals such as zebrafish. We created 13 potential knockout animals in a few months. In addition, by use of this method, informative missense mutations may well be retrieved that can be very valuable when complete loss-of-function results in (embryonic) lethality. Furthermore, such mutations may be useful in protein-domain function studies. Taken together, we think that the target-selected mutagenesis approach described here is the method of choice for the generation of zebrafish knockouts.

ACKNOWLEDGMENTS

We thank Steve Henikoff and Brad Till for the valuable help needed for setting up the TILLING method in our lab, Marco Koudijs, Adam Hurlstone, Anna-Pavlina Haramis, Marjo den Broeder, and Astrid Keijser for help with library construction and screening, and our collaborators for sharing unpublished sequences. This work was supported financially by a NWO genomics grant.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Beier, D.R. 2000. Sequence-based analysis of mutagenized mice. *Mamm. Genome* **11:** 594–597.

Bentley, A., MacLennan, B., Calvo, J., and Dearolf, C.R. 2000. Targeted recovery of mutations in *Drosophila*. *Genetics* **156:** 1169–1173.

Berezikov, E., Plasterk, R.H., and Cuppen, E. 2002. GENOTRACE: cDNA-based local GENOme assembly from TRACE archives. *Bioinformatics* **18:** 1396–1397.

Coghill, E.L., Hugill, A., Parkinson, N., Davison, C., Glenister, P., Clements, S., Hunter, J., Cox, R.D., and Brown, S.D. 2002. A gene-driven approach to the identification of ENU mutants in the mouse. *Nat. Genet.* **30:** 255–256.

Colbert, T., Till, B.J., Tompa, R., Reynolds, S., Steine, M.N., Yeung, A.T., McCallum, C.M., Comai, L., and Henikoff, S. 2001. High-throughput screening for induced point mutations. *Plant Physiol.* **126:** 480–484.

Driever, W., Solnica-Krezel, L., Schier, A.F., Neuhauss, S.C., Malicki, J., Stemple, D.L., Stainier, D.Y., Zwartkruis, F., Abdelilah, S., Rangini, Z., et al. 1996. A genetic screen for mutations affecting embryogenesis in zebrafish. *Development* **123:** 37–46.

Greene, E.A., Codomo, C.A., Taylor, N.E., Henikoff, J.G., Till, B.J., Reynolds, S.H., Enns, L.C., Burtner, C., Johnson, J.E., Odden, A.R., et al. 2003. Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in *Arabidopsis*. *Genetics* **164:** 731–740.

Haffter, P., Granato, M., Brand, M., Mullins, M.C., Hammerschmidt, M., Kane, D.A., Odenthal, J., van Eeden, F.J., Jiang, Y.J., Heisenberg, C.P., et al. 1996. The identification of genes with unique and essential functions in the development of the zebrafish, *Danio rerio*. *Development* **123:** 1–36.

Heasman, J. 2002. Morpholino oligos: Making sense of antisense? *Dev. Biol.* **243:** 209–214.

Hofreiter, M., Jaenicke, V., Serre, D., Haeseler Av, A., and Paabo, S. 2001. DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Res.* **29:** 4793–4799.

Jansen, G., Hazendonk, E., Thijssen, K.L., and Plasterk, R.H. 1997. Reverse genetics by chemical mutagenesis in *Caenorhabditis elegans*. *Nat. Genet.* **17:** 119–121.

Knapik, E.W. 2000. ENU mutagenesis in zebrafish–from genes to complex diseases. *Mamm. Genome* **11:** 511–519.

Lee, K.Y., Huang, H., Ju, B., Yang, Z., and Lin, S. 2002. Cloned zebrafish by nuclear transfer from long-term-cultured cells. *Nat. Biotechnol.* **20:** 795–799.

Li, Y.X., Farrell, M.J., Liu, R., Mohanty, N., and Kirby, M.L. 2000. Double-stranded RNA injection produces null phenotypes in zebrafish. *Dev. Biol.* **217:** 394–405.

Ma, C., Fan, L., Ganassin, R., Bols, N., and Collodi, P. 2001. Production of zebrafish germ-line chimeras from embryo cell cultures. *Proc. Natl. Acad. Sci.* **98:** 2461–2466.

McCallum, C.M., Comai, L., Greene, E.A., and Henikoff, S. 2000. Targeted screening for induced mutations. *Nat. Biotechnol.* **18:** 455–457.

Mullins, M.C., Hammerschmidt, M., Haffter, P., and Nusslein-Volhard, C. 1994. Large-scale mutagenesis in the zebrafish: In search of genes controlling development in a vertebrate. *Curr. Biol.* **4:** 189–202.

Nakamura, Y., Gojobori, T., and Ikemura, T. 2000. Codon usage tabulated from international DNA sequence databases: Status for the year 2000. *Nucleic Acids Res.* **28:** 292.

Nasevicius, A. and Ekker, S.C. 2000. Effective targeted gene 'knockdown' in zebrafish. *Nat. Genet.* **26:** 216–220.

Oates, A.C., Bruce, A.E., and Ho, R.K. 2000. Too much interference: Injection of double-stranded RNA has nonspecific effects in the zebrafish embryo. *Dev. Biol.* **224:** 20–28.

Oleykowski, C.A., Bronson Mullins, C.R., Godwin, A.K., and Yeung, A.T. 1998. Mutation detection using a novel plant endonuclease. *Nucleic Acids Res.* **26:** 4597–4602.

Perry, J.A., Wang, T.L., Welham, T.J., Gardner, S., Pike, J.M., Yoshida, S., and Parniske, M. 2003. A TILLING reverse genetics tool and a Web-accessible collection of mutants of the legume *Lotus japonicus*. *Plant Physiol.* **131:** 866–871.

Rozen, S. and Skaletsky, H. 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* **132:** 365–386.

Sun, L., Bradford, C.S., Ghosh, C., Collodi, P., and Barnes, D.W. 1995. ES-like cell cultures derived from early zebrafish embryos. *Mol. Mar. Biol. Biotechnol.* **4:** 193–199.

Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C., Odden, A.R., Young, K., Taylor, N.E., et al. 2003. Large-scale discovery of induced point mutations with high-throughput TILLING. *Genome Res.* **13:** 524–530.

van Eeden, F.J., Granato, M., Odenthal, J., and Haffter, P. 1999. Developmental mutant screens in the zebrafish. *Methods Cell Biol.* **60:** 21–41.

Wargelius, A., Ellingsen, S., and Fjose, A. 1999. Double-stranded RNA induces specific developmental defects in zebrafish embryos. *Biochem. Biophys. Res. Commun.* **263:** 156–161.

Wienholds, E., Schulte-Merker, S., Walderich, B., and Plasterk, R.H. 2002. Target-selected inactivation of the zebrafish *rag1* gene. *Science* **297:** 99–102.

Yang, B., Wen, X., Kodali, N.S., Oleykowski, C.A., Miller, C.G., Kulinski, J., Besack, D., Yeung, J.A., Kowalski, D., and Yeung, A.T. 2000. Purification, cloning, and characterization of the CEL I nuclease. *Biochemistry* **39:** 3533–3541.

Zan, Y., Haag, J.D., Chen, K.S., Shepel, L.A., Wigington, D., Wang, Y.R., Hu, R., Lopez-Guajardo, C.C., Brose, H.L., Porter, K.I., et al. 2003. Production of knockout rats using ENU mutagenesis and a yeast-based screening assay. *Nat. Biotechnol.* **21:** 645–651.

Zhao, Z., Cao, Y., Li, M., and Meng, A. 2001. Double-stranded RNA injection produces nonspecific defects in zebrafish. *Dev. Biol.* **229:** 215–223.

WEB SITE REFERENCES

http://cuppen.niob.knaw.nl; Detailed protocols for TILLING procedures and CEL-I isolation.

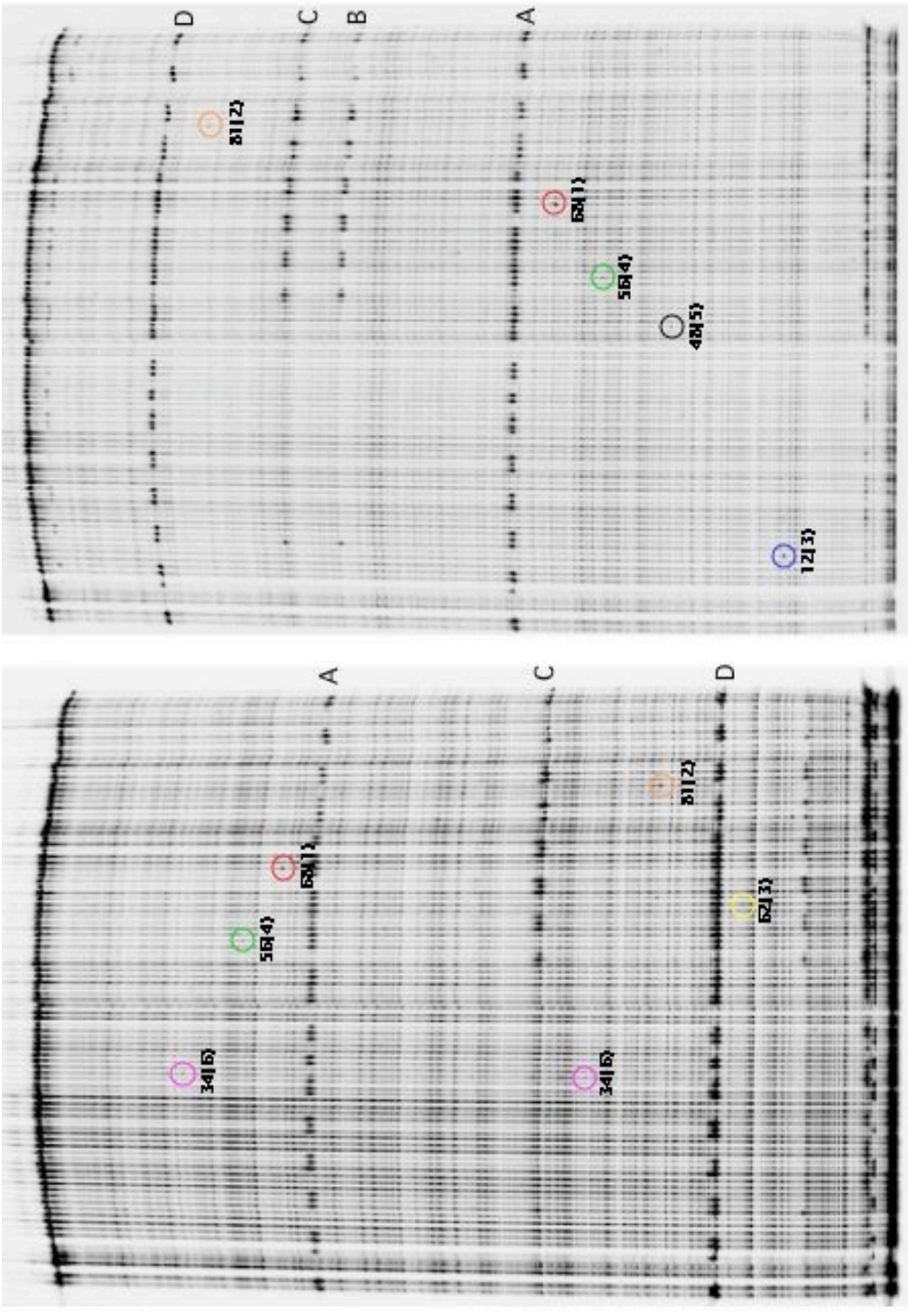
http://genotrace.niob.knaw.nl; Interface to GENOTRACE for local genomic sequence assemblies from raw genome sequencing traces using EST or cDNA sequences as input.

http://primers.niob.knaw.nl; Design of sets of nested (tailed) primers for TILLING.

http://zfin.org; The zebrafish information network.

Received July 4, 2003; accepted in revised form September 10, 2003.

Supporting Online Material



Supplementary figure 1: LiCOR image showing raw TILLING results and ranking criteria. The left and right panel show the results from the 700 nm and 800 nm fluorescence, respectively. Potential mutations are indicated by a circle. The numbers below the circle indicate the lane number and the scoring rank (between brackets). SNPs are indicated on the side (capital letters A-D). Lane 12, rank 3, strong signal in one channel (800), mutation confirmed by sequencing. Lane 34, rank 6 (aberrant): two signals are found in this lane, but only in one channel, not confirmed. Lane 48, rank 5, weak signal in one channel (800), not confirmed. Lane 56, rank 4, weak signals in both channels, confirmed. Lane 62, rank 3, strong signal in one channel (700), not confirmed. Lane 68, rank 1, strong signal in both channels, confirmed, Lane 81, rank 2, strong signal in one channel (700), weak in the other channel (800), not confirmed.

Chapter 4

Target-selected gene inactivation in zebrafish

Reprinted from Wienholds and Plasterk (2004), *Methods Cell Biol* 77:69-90

Target-Selected Gene Inactivation in Zebrafish

1. Introduction

1.1. Reverse Genetics in a ‘‘Forward’’ Genetic Model Organism

The zebrafish was developed as a vertebrate genetic model organism (Kimmel, 1989; Streisinger *et al.*, 1981) because of properties that allow forward genetic screens: small size, relative short generation time, and many offspring. The logistics of breeding allow mutant screens of such size that most genes whose loss results in a mutant phenotype can be hit multiple times (Mullins *et al.*, 1994; Solnica-Krezel *et al.*, 1994). This can produce a virtually complete, and saturated inventory of genes, specifically involved in different stages of embryonic development. Large-scale forward mutagenesis screens have identified thousands of mutants defective in a variety of embryological processes (Driever *et al.*, 1996; Haffter *et al.*, 1996) and more sophisticated forward mutagenesis screens, such as modifier screens, are being conducted to find more genes involved in embryonic development.

An alternative way to link gene and function is by reverse genetics. Here, a specific targeted gene is inactivated and the consequences are studied. The vertebrate model at present commonly used for reverse genetics is the mouse. However, there is also a growing need for reverse genetics in the fish for a variety of reasons: (1) With the genome being sequenced, researchers recognize zebrafish orthologs of interesting genes from other species and want to obtain mutants to study the functions. (2) Genes initially resulting from a forward mutant hunt in the zebrafish might deserve more detailed attention; for example, the mutant originally isolated might be a missense mutant, and it can be of interest to obtain a known null mutant. (3) Large-scale studies in zebrafish on gene expression (using microarrays or *in situ* hybridizations) or screens for morpholino effects might identify genes of interest, and again the question is what the mutant phenotype will be.

1.2. Knockouts and Knockdowns

Strategies to reduce or completely knock out gene function fall into two classes: those that aim to knock down gene expression (RNAi and morpholinos) and those that aim at isolation of a mutation in the gene. Both approaches have advantages. In zebrafish, knockdowns are mostly done by morpholinos (Nasevicius *et al.*, 2000); given the success of this approach, RNAi with short interfering RNAs (siRNAs; Elbashir *et al.*, 2001) has not yet been thoroughly tested in zebrafish. Such knockdowns have great advantages: they are quick and relatively easy, can be done at a large scale, and (except for the costs of morpholinos) are cheap. Nevertheless, there are also advantages in having genuine mutations. They are permanent, they can be combined with other mutations, and mutant fish can be used as starting point for subsequent screens for modifier genes. Also, in principle, one can obtain mutations that are guaranteed nulls, whereas with knockdowns one can never be sure. In practice, it is often useful to apply both morpholino and mutation studies to the same gene. This chapter further addresses the generation of zebrafish that have mutations in a gene of choice.

2. Gene Targeting Strategies

2.1. Homologous Disruption

The method of choice for gene targeting in mouse is homologous disruption by recombination in embryonic stem (ES) cells (Capecchi, 1989; Thompson *et al.*, 1989). Although this approach is very efficient in mouse, it has not yet successfully been

applied to any other model organism, primarily because of the lack of isolation or cultivation of pluripotent ES cells. In zebrafish, ES-like cells have been isolated (Sun *et al.*, 1995). After cultivation in vitro and transplantation into the developing embryo, these cells have been shown to be able to contribute to the germline (Ma *et al.*, 2001). However, this has not yet resulted in a targeted knockout. An alternative to the use of ES cells is to perform homologous recombination in somatic cell lines followed by nuclear transfer into oocytes. In sheep and pigs, this has resulted in gene targeting (Dai *et al.*, 2002; McCreath *et al.*, 2000). In zebrafish, it is now also possible to perform nuclear transfer of genetically modified cultured cells (Lee *et al.*, 2002), but a knockout has not yet been reported.

In *Drosophila*, gene targeting by homologous recombination has been performed successfully *in vivo*. Extrachromosomal linear DNA fragments recombined with the target gene (Rong *et al.*, 2000, 2001). Similarly, in *Arabidopsis*, homologous disruption also occurred on transformation (Miao *et al.*, 1995). In principle, one can envisage that microinjection of DNA fragments into zebrafish oocytes or early embryos might provide a similar method for gene targeting. Presumably, the limiting factor is not so much the frequency of such homologous recombination events (which is also often low in mouse ES cells) as the absolute numbers of injected eggs. Coinjection of proteins involved in homologous recombination (Cui *et al.*, 2003) or genetic backgrounds that enhance homologous recombination (de Wind *et al.*, 1995; Hanada *et al.*, 1997) might reduce the number of embryos considerably. Although the methodologies described next might mean that homologous disruption is not required for simple gene knockouts, it would still be a valuable method to develop (e.g., for specific knockins or gene fusions).

2.2. Target-Selected Mutagenesis

Other approaches currently available for making gene knockouts are, strictly speaking, not gene targeting, because lesions in the genes of interest are not generated in a targeted manner but randomly throughout the genome. Therefore, it is more precise to describe them as 'target-selected' gene inactivations: the mutagenesis per se is random, but mutations are sought in a targeted manner in a gene of interest by analysis of the genomic DNA (Fig. 1). Target-selected mutagenesis has successfully been used in various organisms.

The first class of target-selected gene inactivation is insertional mutagenesis. Here, the mutagen is a transposon or virus, which is inserted randomly throughout the genome, thereby disrupting genes. A (pooled) library of mutagenized animals is screened for insertions in target genes. By PCR with insert-specific primers and gene-specific primers, insertions in target genes can be discovered in a (pooled) library of mutagenized animals. This relatively easy method has been successful in *Drosophila* (Ballinger *et al.*, 1989; Kaiser *et al.*, 1990), *C. elegans* (Zwaal *et al.*, 1993), and plants (Das *et al.*, 1995; Koes *et al.*, 1995; Meissner *et al.*, 1999). Alternatively, the insert is used as a starting point for adaptor-mediated or inverse PCR. The sequences flanking the inserts are determined and mapped to the genome and (predicted) genes by, for example, BLAST (Altschul *et al.*, 1990). By determining the flanking sequences of a comprehensive number of insertions, disruptions can be identified in virtually any of the (predicted) target genes. This has been proven feasible in *C. elegans* (Korswagen *et al.*, 1996; Martin *et al.*, 2002) and mouse (Zambrowicz *et al.*, 1998) and highly efficient in *Drosophila* (Spradling *et al.*, 1999) and *Arabidopsis* (Alonso *et al.*, 2003; Parinov *et al.*, 1999; Sessions *et al.*, 2002; Tissier *et al.*, 1999).

In zebrafish, a large-scale insertional mutagenesis screen has been performed by using retroviral vectors as mutagen (Amsterdam *et al.*, 1999; Golling *et al.*, 2002). Here, the goal is a forward screen followed by rapid cloning of the affected genes. However, these same founder fish might also be used for the target-selected mutagenesis approaches described previously. The founder fish harbor approxi-

mately 1,000,000 insertions and 1 embryonic lethal mutation is expected per 70–100 insertions (Amsterdam *et al.*, 1999). Thus, in this collection, at least 10,000 insertions are expected to disrupt genes essential for embryonic development (five times the number of genes expected to be essential for embryonic development). In principle, insertions in these and other (nonessential) genes can be found by the methods described previously. However, the mutagenesis efficiency is such that high numbers of fish (30,000–100,000; Amsterdam, 2003) have to be archived to find at least one insertion per gene. It remains to be seen whether this is feasible in practice. An alternative insertion mutagen is a transposon (Davidson *et al.*, 2003; Raz *et al.*, 1998), although insertion frequencies are not optimal.

The second class of target-selected mutagenesis is chemical mutagenesis. Chemicals can introduce various kinds of mutations. In *C. elegans*, chemical mutagenesis is performed using *N*-ethyl-*N*-nitrosourea (ENU), ethylmethanesulfonate (EMS), or trimethylpsoralen (TMP) in order to generate small deletions. These deletions can be detected by a simple PCR method (Jansen *et al.*, 1997; Liu *et al.*, 1999). Individuals carrying deletions in any gene can then be recovered from a pooled library of millions of animals. In zebrafish, ENU is also widely used as mutagen in forward screens. Screening these animals for small deletions is possible in theory (Lekven *et al.*, 2000), but the number of animals needed to hit any gene at least once is probably unfeasable (at least 100,000 if 10% of the mutations are small deletions as in *C. elegans*; De Stasio *et al.*, 1997). In addition, small deletions have seldom been recovered after ENU mutagenesis of mouse spermatogonial germ cells and ES cells (Chen *et al.*, 2000; Douglas *et al.*, 1995) and zebrafish premeiotic germ cells in forward screens (Fig. 2B), indicating that using ENU for creation of small deletions is not practical. Alternatively, one could use EMS or TMP, but mutagenesis frequencies are too low and only large deletions have been recovered so far (Lekven *et al.*, 2000; Mullins *et al.*, 1994; Solnica-Krezel *et al.*, 1994). The most powerful and widely used mutagens for mutagenesis screens are ENU and EMS, which cause predominantly point mutations. Point mutations are more difficult to detect than insertions and small deletions, but the mutagenesis efficiency is so much higher that only a limited number of animals are needed to reach saturation (typically a few thousand). For organisms such as *C. elegans* and *Drosophila*, the ability to raise a larger number of animals might not be a limiting factor, but for vertebrates it definitely is. Recent methods for the identification of point mutations have improved considerably in speed and cost. One can now extensively screen a limited number of animals and still recover knockout alleles of genes of interest. Therefore target-selected mutagenesis using mutagens causing point mutations, such as ENU and EMS, is likely to be the method of choice for creating knockouts. Various target-selected mutation detection methods have already been successfully applied to *Arabidopsis* (McCallum *et al.*, 2000), lotus (Perry *et al.*, 2003), *Drosophila* (Bentley *et al.*, 2000), *C. elegans* (R. Plasterk and E. Cuppen, unpublished data), mouse (Beier, 2000; Coghill *et al.*, 2002), rat (Smits *et al.*, 2004; Zan *et al.*, 2003), and zebrafish (Wienholds *et al.*, 2002, 2003b).

3. Target-Selected Mutagenesis in Zebrafish

At present, the only working method for making knockouts in zebrafish is chemical mutagenesis followed by targeted screening for point mutations (Fig. 1). The first gene in zebrafish to be knocked out was *rag1* (Wienholds *et al.*, 2002). In this case, and in other approaches described hereafter, the mutagen was the standard chemical mutagen ENU. Because ENU is the same mutagen that is being used for forward genetic screens, one can use one mutagenesis regime for forward as well as reverse genetic goals. For finding mutations in the *rag1* gene, we screened the F1 fish from the Tübingen 2000 mutagenesis screen and found a premature stop codon by straightforward DNA resequencing of amplicons containing most of the coding re-

gions of the gene. This turned out to be a genuine null allele. Since then, a faster and cheaper method has been developed to prescreen fish for the presence of mutations within an amplicon and then sequence only those amplicons known to contain a mutation. The method is called TILLING, and was initially developed by the laboratory of Steve Henikoff for *Arabidopsis* (Colbert *et al.*, 2001) and applied to the zebrafish (Fig. 4) by us and others (Wienholds *et al.*, 2003b). Noted that there are also potential alternative approaches for this prescreen step, such as a method based on phage Mu transposition described recently (Yanagihara *et al.*, 2002), and possibly other methods. To date, the only methods that have generated the target-selected zebrafish mutants are resequencing and TILLING, and the latter is described here in more detail.

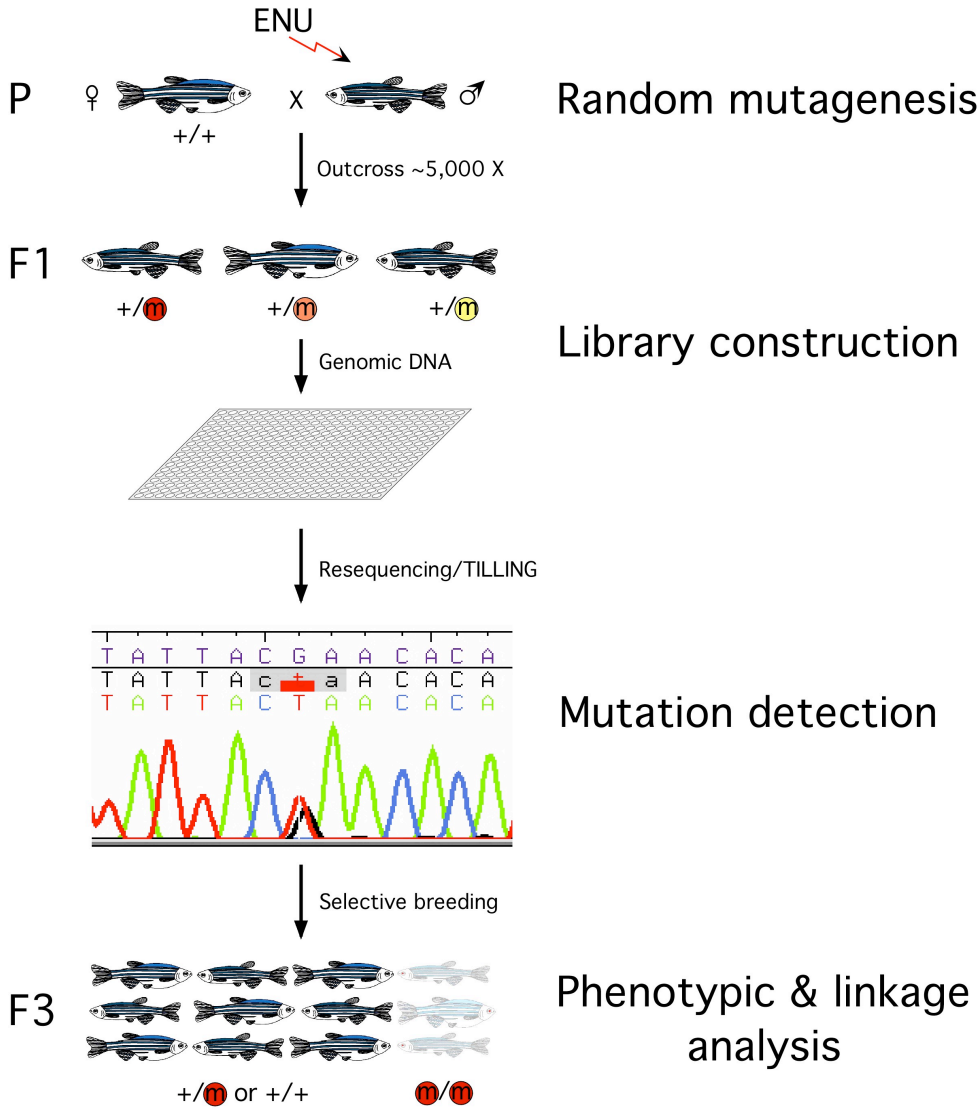


Fig. 1 Target-selected mutagenesis in zebrafish. Approximately 100 adult male zebrafish are randomly mutagenized with *N*-Ethyl-*N*-nitrosourea (ENU) and outcrossed against wild-type females. A library of approximately 5000 F1 animals is constructed, in principle having independent mutations. Genomic DNA of these F1 animals is isolated and arrayed in 384-well PCR plates, suitable for robotic handling. The DNA is screened for mutations in target genes by resequencing or TILLING. Animals with interesting mutations are recovered from the library [reidentified from a pool of living F1 fish or recovered by in vitro fertilization (IVF) with frozen sperm] and outcrossed against wild-type fish or incrossed with other mutants. Finally, the mutations are homozygosed and animals are analyzed for phenotypes and linkage to the mutation.

3.1. Mutagenesis

The success rate of ENU mutagenesis screens depends on several factors. First, the mutagenesis must be optimal. The better the mutagenesis, the more chance to obtain mutants in a given set of animals. Typically, young adult male zebrafish are mutagenized with 4–6 consecutive treatments of ENU (van Eeden *et al.*, 1999), and after a few weeks they are crossed with wild-type females to obtain an F1 generation of fish carrying nonmosaic mutations (Fig. 1). Increasing ENU dose or the number of treatments might result in higher mutation frequencies, but also in decreased survival (Solnica-Krezel *et al.*, 1994), reduced fertility, and increased chance of clonal mutants. Prior to outcrossing, the mutagenesis efficiency is monitored with a specific locus test. Mutagenized founders are crossed with tester females carrying known homozygous mutations that can be scored easily, for example, pigmentation mutants *albino* or *sparse*. Depending on the marker gene assayed, a good mutagenesis typically gives a specific locus rate of 0.2–0.3% (Mullins *et al.*, 1994; Solnica-Krezel *et al.*, 1994). The second factor influencing the possibility of recovering mutants is the number of animals screened. Screening more animals will lead to better chance of finding at least one mutant allele. The third factor influencing the possibility of finding mutant alleles is the size of the target gene. The larger the gene, the more chance of finding potential loss-of-function alleles. In forward screens, the mean mutation frequency per gene is approximately 1 in 1000 genomes. Thus, for an average gene one would have to screen about 1000 animals to get a potential loss-of-function allele. Screening smaller genes requires more animals.

To date, more than a hundred ENU-induced mutations have been cloned in forward genetic screens. The spectrum of these mutations (Fig. 2A) is biased, because each of these mutations has a phenotype, mostly loss-of-function. About half (46%) introduce a nonsense codon; 15% alter a splice site, resulting in insertions and deletions in the mRNA; and 34% are missense mutations (Fig. 2B). The ENU spectrum of 280 mutations identified in reverse genetics screens is unbiased.

However, it is similar to that of forward screens (Fig. 2A). Of the mutations found in noncoding regions, only the ones in splice sites (2.5%, Fig. 2B) will, with great certainty, alter protein function. The influence of the mutations found in coding regions can be calculated from the codon usage in zebrafish (Nakamura *et al.*, 2000, http://www.kazusa.or.jp/codon/) and the observed spectrum of ENU-induced mutations in unbiased (thus reverse) genetic screens. Of these, 29% are expected to be silent, 66% missense, and 5.0% nonsense. (The last is predicted to be 5.6% if the ENU spectrum from forward genetic screens is used.) The number of nonsense mutations found in reverse screens thus far (Fig. 2B) is even somewhat higher (6.9%), which could reflect statistical variation or local codon usage differences. Splice-site and nonsense mutations are the most likely candidates to be loss of function. However, the pool of missense mutations is also likely to contain functional changes. In forward screens, the ratio between missense and nonsense mutations is 0.7 (37/52). Extrapolation to reverse screens suggests that 3.5% (0.7 times 5%) of the missense mutations found in coding regions also give rise to a phenotypic chance. The challenge is to predict which missense mutations might give rise to a functional change of the protein (Chasman *et al.*, 2001; Ng *et al.*, 2001; Sunyaev *et al.*, 2001).

Together, up to 11% of the mutations (nonsense, splice site, and missense) are expected to have a (partial) loss-of-function phenotype or other influence on protein function. The probability of finding at least one such allele in different-sized target fragments in a library of mutagenized animals, with a per-base mutation frequency of approximately 1 in 250,000 bp (Wienholds *et al.*, 2003b), is shown in Fig. 2C. For example, to have an 80% chance of finding at least one nonsense or splice site mutant in a 1-kb gene, one needs to screen approximately 6000 animals.

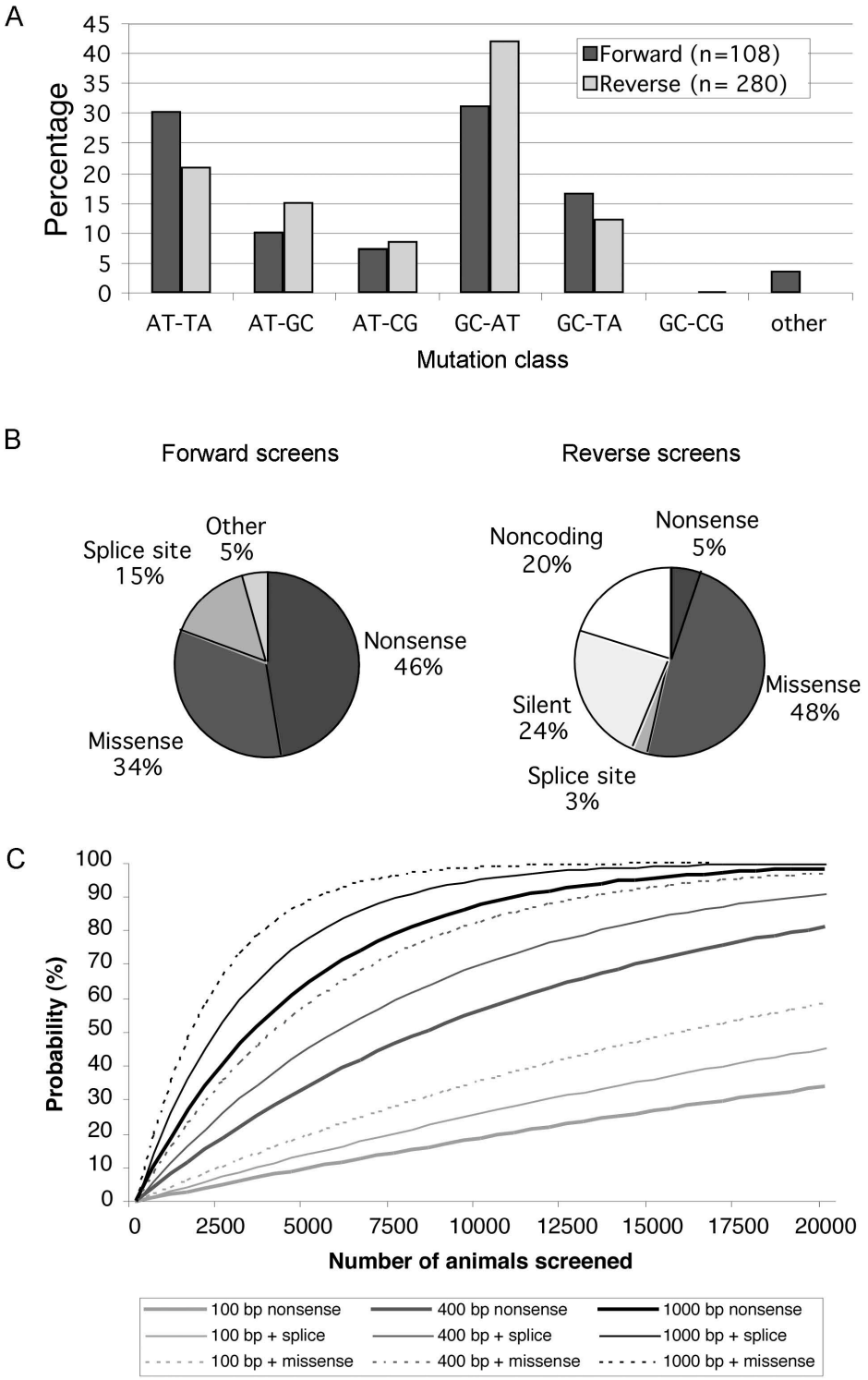


Fig. 2 ENU-induced mutations in zebrafish. (A) Mutations spectra of ENU-induced mutations in forward and reverse genetic screens. (B) Consequence of mutations at coding level. Other mutations in forward screens are insertions and deletions. The ratio of mutations, strictly in the coding regions in reverse screens, is 6.9% nonsense, 62.2% missense, and 30.9% silent. (C) Probabilities for finding at least one potential loss-of-function mutation in ENU-mutagenized F1 zebrafish. Binomial distribution probabilities are calculated for three different coding sequence fragment sizes (100, 400, and 1000 bp) with a mutagenesis efficiency of 1 in 250,000 bp (0.00004 mutations per base). Nonsense, splice site, and missense mutations that potentially result a phenotypic change represent approximately 5%, and 2.5%, and 3.5% of the mutations, respectively. The last two are accumulated to the class of nonsense mutations (nonsense and splice site mutations; nonsense, splice site and missense mutations).

3.2. Knockout Libraries

There are two ways to store the library of mutagenized fish. For inactivation of the *rag1* gene and screening of several other genes, we have frozen the sperm of 2679 mutagenized F1 males (Brand *et al.*, 2002; Wienholds *et al.*, 2002). Keeping the library as a frozen stock results in a permanent library that can be screened many times for many different genes over an unlimited period of time. In addition, it saves a lot of space, compared to an aquarium facility. Some drawbacks are that the cryopreservation of sperm samples is quite laborious and mutants need to be recovered by in vitro fertilization (IVF). For recovery of the *rag1* loss-of-function allele this worked well, but we were unable to recover several other mutants by IVF (unpublished observation). However, recent advances in sperm freezing methods (Morris *et al.*, 2003) and IVF indicate that the creation of a reliable, permanently frozen library of sperm is feasible. A fast alternative to freezing sperm is to keep the mutagenized animals alive during the screening process. Such a library is relatively easy to construct. As soon as the outcrossed mutagenized F1 animals are old enough, pieces of the fins are removed for DNA isolations (fin-clipped). The fish are then pooled to minimize the aquarium space required. For the second knockout library, we fin-clipped 4608 fish and grouped them into 384 pools of 12 (Wienholds *et al.*, 2003b). A consequence of pooling is that after screening the complete library, the fish of a positive pool have to be fin-clipped and genotyped again. An advantage of such a living library is that both males and females can be screened. If different interesting mutations are found in the same gene, but in opposite sexes, these can be crossed at once, generating transheterozygous fish. In addition to all background mutations then being heterozygous, this saves a full generation time for analysis. A drawback of living libraries is that they can be screened for only a limited period of time: from the point the fish are fin-clipped until they lose fertility (approximately 1.5 years). If aquarium space is limiting, one can construct several small libraries (e.g., 384 animals) that are extensively screened for many genes (e.g., 100 genes), as is being done for rats (Smits *et al.*, 2004). Of course, a combination of a living and a cryopreserved library is possible. One can start constructing and screening a living library, remove the males, and cryopreserve their sperm.

3.3. Mutation Detection by Resequencing

The most straightforward and reliable method for point mutation detection is DNA sequencing. During the past couple of years, DNA sequencing techniques have improved considerably in throughput and cost. This has resulted in the human genome being sequenced in about 1 year (Lander *et al.*, 2001; Venter *et al.*, 2001).

The first library used for target-selected inactivation of the *rag1* gene required only 2679 animals. To maximize the chance of finding a guaranteed loss-of-function allele in this limited set of animals, we have chosen to screen them by resequencing in a one-to-one manner, minimizing the number of false negatives. The pipeline for this resequencing is simple. First, we PCR-amplified most of the coding regions of the *rag1* gene. Because this resulted in variations in yield, we used a small volume of this PCR as template in a nested PCR with internal primers. After the nested PCR, all samples had equal yield. Next, the PCR products were either sequenced directly or were first purified and then sequenced with internal primers designed to cover approximately 600 bp per read. Sequence samples were purified and analyzed on a 96-capillary sequencer (ABI3700 DNA analyzer). Although initially only the liquid handling steps were performed by robots, we have now automated the entire process, including all the PCR and purification steps.

Because we screened F1 animals, the mutations were heterozygous. Therefore, we could not do a simple alignment to detect these mutations and filter out the differences. We detected heterozygous positions by parsing the trace files in batches of 96 through Phred (Ewing *et al.*, 1998), Phrap (Gordon *et al.*, 2001), and Polyphred (Nickerson *et al.*, 1997) and visualized them with Consed (Gordon *et al.*, 1998). The

batches of 96 samples were inspected simultaneously for heterozygous mutations. New mutations could be discriminated from naturally occurring single nucleotide polymorphisms (SNPs), because these were present in multiple samples whereas the genuine mutations were present only in one sample.

To find a knockout of the *rag1* gene, we analyzed almost 12,500 sequence reactions covering most of the coding regions of the gene. This yielded nine amino acid changes and one premature stop. This nonsense mutation was in the middle of the catalytic domain of the recombinase and therefore expected to be a null. Indeed, fish homozygous for this stop allele were defective in V(D)J recombination and immunodeficient, indicating that this allele was a true loss-of-function. In addition to this *rag1* loss-of-function allele, we found dominant-negative mutations in *p53* (S. Berghmans, R. Murphey *et al.*, manuscript in preparation) and several potential loss-of-function mutations in *tie2* (A. M. Kuechler *et al.*, manuscript in preparation). Screening these genes took us approximately 1 month for each gene and costs are estimated around U.S. $12,500. More recently, the method of choice for finding mutations is TILLING (Section 3.4). However, new technological advances in sequencing techniques and automation will probably reduce both time and cost for finding mutants by direct resequencing significantly, so that in the future resequencing might again be the method of choice.

3.4. Mutation Detection by TILLING)

TILLING (targeting induced local lesions in genomes) was originally developed for *Arabidopsis* and used denaturing high-performance liquid chromatography (DHPLC) to detect EMS-induced mutations (McCallum *et al.*, 2000). The implementation of enzyme-mediated mismatch recognition (Oleykowski *et al.*, 1998) allowed a high throughput detection of mutations (Colbert *et al.*, 2001). This led to the *Arabidopsis* TILLING Project (ATP) (Till *et al.*, 2003), which has isolated more than 100 potential knockouts within 1 year. Because TILLING is approximately 10 times faster and cheaper than resequencing, we implemented it to prescreen for ENU-induced mutations in zebrafish (Wienholds *et al.*, 2003b).

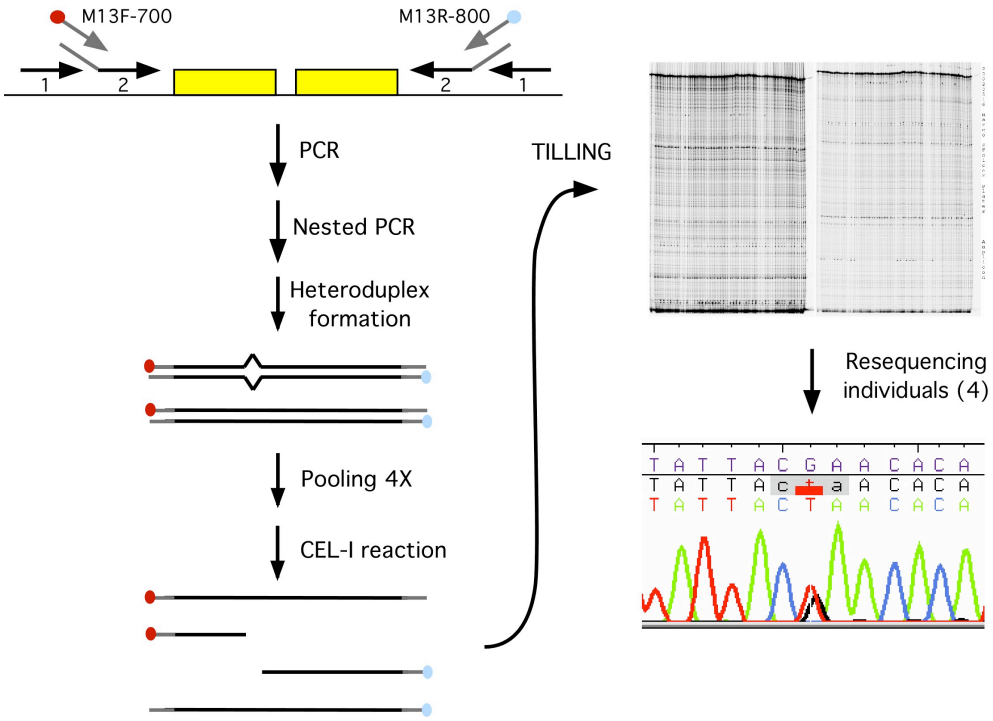


Fig. 3 Mutation detection by TILLING. See text for detailed description of all the steps involved.

The setup for TILLING in zebrafish is similar to that of resequencing. It involves nested PCR followed by mutation detection in the amplified fragments (Fig. 3). First the target is amplified by PCR with gene-specific primers. The maximum length of the fragments that can be analyzed, with good sensitivity, is approximately 1000 bp. The continuous stretches of coding regions or exons in zebrafish (and most other vertebrates) are usually smaller (around 100 bp). Therefore, amplicons are designed to contain several of these smaller exons or, occasionally, single exons. In addition, exons in the 5´ region of the gene are favored for screening, because a nonsense or splice-site mutation might lead to a complete removal of downstream part of the protein. In nested PCR, the concentrations of the amplified fragments are equalized and the fragments are labeled with fluorescent dyes. These labels can be incorporated in two ways. The nested set of gene-specific primers can be directly labeled or the nested set of primers can be tailed with universal adaptor sequences, for which the corresponding labeled universal primers are mixed into the PCR reaction. A major advantage of the latter method is that the same fluorescent labeled primers can be used to label any fragment, thereby reducing costs considerably: unmodified primers are cheaper than fluorescent primers and the two universal fluorescent primers are ordered in large batches. After PCR, heteroduplexes between wild-type and mutant PCR fragments are formed by denaturing and reannealing the PCR fragments. This is necessary to form mismatches, which can be recognized by the mismatch recognition enzyme CEL-1 isolated from celery (Oleykowski *et al.*, 1998; Yang *et al.*, 2000). To increase the throughput, samples are pooled four times. Additional pooling is expected to result in a decreased sensitivity because the complex nature of the vertebrate genome. SNPs present in the target fragments might mask the discovery of new mutations. Pooled fragments are treated with the CEL-1 enzyme.

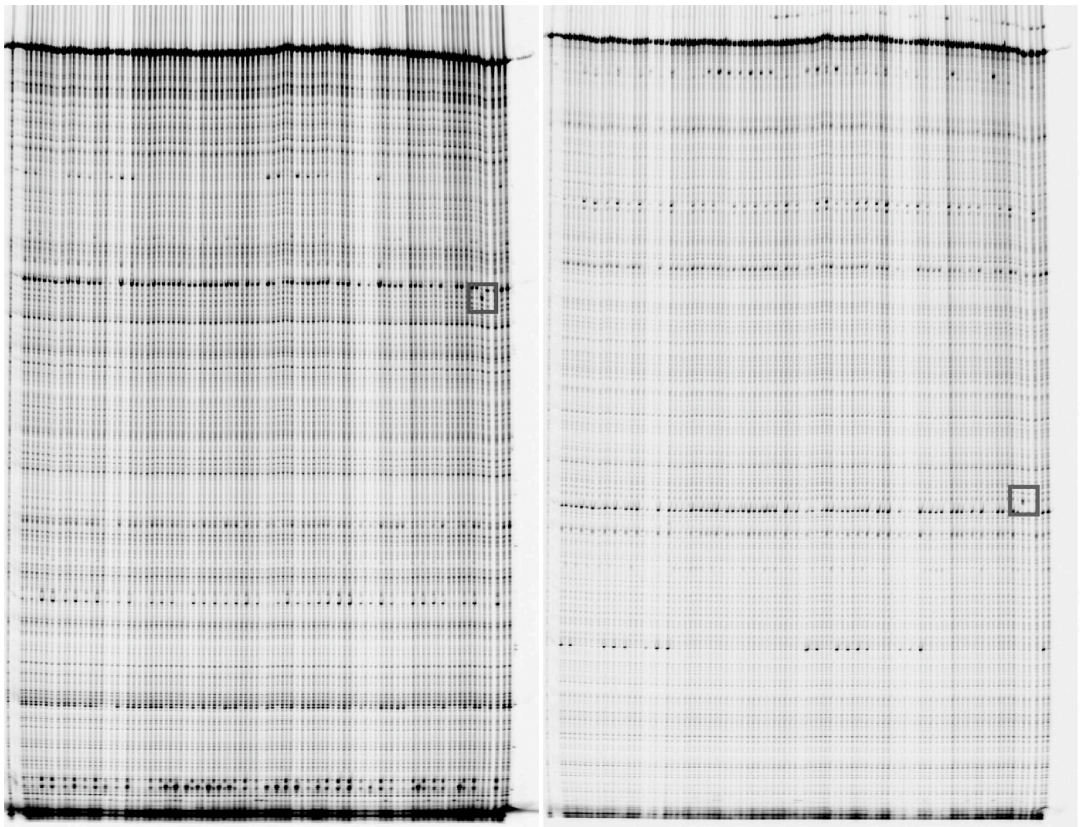


Fig. 4 Example TILLING gel. Each of the 96 lanes contains pooled CEL-1-digested PCR fragments of the dicer1 gene of four different animals. The IR Dye 700 and IR Dye 800 channels are shown on the left and right, respectively. A potential mutation is boxed in both the 700 and 800 IR Dye channels. After sequencing, this mutation turned out to be nonsense and causes a loss-of-function phenotype.

Table 1.
Summary of Genes Screened and Mutations Found by TILLING in Zebrafish

Amplicon			No. of confirmed mutations[b]					
Gene	Name	Length/cds3 (bp)[a]	Total	Nonsense	Splice	Missense	Silent	Noncoding
gene1	A	569/264	6	—	—	—	1	5
gene2	A	777/381	15	—	1	6	2	6
gene3	A	639/327	18	1	2	7	3	5
gene4	A	911/879	14	1	—	8	5	—
gene5	A	812/327	18	2	1	2	3	10
gene6	A	541/436	6	1	—	2	3	—
gene7	A	590/264	6	—	1	2	1	2
gene8	A	442/270	6	—	1	2	2	1
gene9	A	841/279	20	—	—	7	2	11
	B	448/230	7	—	1	2	2	2
gene10	B	664/395	4	1	—	3	—	—
gene11	A	720/720	11	—	—	6	5	—
gene12	A	496/469	6	—	—	3	3	—
gene13	A	816/443	9	1	—	4	1	3
apc[c]	A	953/953	16	—	—	14	2	—
	B	873/873	13	1	—	7	5	—
	C	813/813	10	—	—	5	5	—
	G	921/881	11	1	—	6	4	—
gene15	A	955/955	18	2	—	12	4	—
dicer1[d]	A	798/775	17	2	—	9	6	—
	C	752/418	16	1	—	8	2	5
	E	736/526	8	—	—	4	2	2
Average		730/540	11.6	0.6	0.3	5.4	2.9	2.4
Total		16,067/11,878	255	14	7	119	63	52

[a] Length of nested PCR amplicon, excluding primer sequences; length of coding sequence (cds) in amplicon.
[b] Potential mutations were found by TILLING and confirmed by resequencing.
[c] Hurlstone *et al.* (2003).
[d] Wienholds *et al.* (2003a).
Source: Adapted from Wienholds, E. *et al*. (2003b). Efficient target selected mutagenesis in zebrafish. *Genome Res*. **13**, 2700–2707, with permission.

Fragments are purified using G50 minicolumns or isopropanol precipitation. After purification, the labeled digested fragments are separated and visualized on slab gel sequencers (e.g., see Fig. 4). Each gel can be used to analyze 96 samples at one time. Because the samples are pooled four times, these gels represent 384 animals. The gels are manually processed and inspected with the Photoshop software. In principle, a mutation should give a strong signal in both lanes at reciprocal heights (Fig. 4), but we found that this is not an absolute prerequisite to confirm a mutation; often any signal that is different in one lane is confirmed successfully. The mutation can be confirmed by resequencing the individual four samples of the pooled lane. First, the target amplicon is reamplified from the first PCR to confirm the TILLING results. Next it is amplified from the original genomic DNA again to exclude any PCR artifacts or other mistakes.

The TILLING procedure described can be performed at different scales. We use a sophisticated robotic setup, with a 96-channel pipette and eight integrated 384-well PCR machines and six slab gel sequencers to screen for mutations in a large num-

ber of genes. With this setup, a single person can prescreen a library of 9216 animals by TILLING in 1 to 2 days. The same protocol can be used at a smaller scale if only a small number of genes need to be mutated. All the steps can be done manually, using multichannel pipettes. With a few PCR machines and a single slab gel sequencer, one person should be able to screen the same number of animals for one gene within a week. Recently, we screened a library of 4608 mutagenized F1 animals for mutations in 16 different genes by TILLING (Table I, Wienholds *et al.*, 2003b). We found 458 potential mutations, of which 255 could be confirmed by resequencing. These included 119 missense, 14 nonsense, and 7 splice site mutations, the last two classes found in 13 genes and expected to be loss of function. Thus, we potentially knocked out 13 different genes by using the TILLING setup for mutation detection. One of the genes we screened for mutations is the *dicer1* gene (Fig. 5). Three different amplicons were designed to cover as much of the coding region as possible. In these three fragments, 78 potential mutations were identified. Approximately half (41) of them could be confirmed by resequencing (Table I). Of these, 21 were missense mutations and 3 were nonsense mutations (Fig. 5). All three nonsense mutations displayed similar phenotypes as homozygotes: developmental arrest and failure in microRNA processing (Wienholds *et al.*, 2003a). The influence of the missense mutations on the protein function is currently being investigated.

4. Discussion

4.1. Linking Genotypes and Phenotypes

There is something counterintuitive about this approach: to target a gene one mutagenizes the entire genome. The expected number of genes per haploid genome that is knocked out by a standard ENU mutagenesis is of the order of 10 to 20 (which fits with the expected numbers of genes in a vertebrate genome of about 30,000, and the chance for an average gene to be knocked out in 1 in 1000–2000 animals). So if one recognizes by TILLING that an F1 fish has a mutation in a gene of interest, it will most likely contain multiple other mutations. How are these filtered? How do you know the mutation causes the observed phenotype? These are questions most commonly raised when this approach is presented.

First, the collateral damage is equally big in any forward genetic protocol (in which it does not seem to raise nearly as many eyebrows). In practice, it is hardly ever a problem, for the following reasons:

1. Even in the very first backcross to homozygose a mutation observed in an F1 fish, one recognizes linkage between being homozygous for the mutation and having a given phenotype. It is common that in such backcrosses multiple additional phenotypes are seen, many with Mendelian segregation patterns, but genotyping the offsping will quickly sort out the unlinked mutations.
2. As with any mutagenesis, it is wise to outcross a mutant once or twice more before too much analysis is done. Then only tightly linked mutations might blur the analysis.
3. Fortunately, one often encounters more than one reduction-of-function-or even loss-of-function allele in one screen. These can then be used to address the specificity question in two ways: if two independent mutations in the same gene have the same phenotype, this strongly argues for a causal effect, because it would be highly unlikely that in both cases a genetically linked second mutation causes the same phenotype. We have also used two alleles to further prove causation: one can create heteroallelic animals, in which one allele has one mutation and the homologous allele has the other. Then this fish is heterozygous for all other possibly linked mutations, and if it still has the same phenotype as fish homozygous for each individual mutation, one can be basically sure of a causal effect.

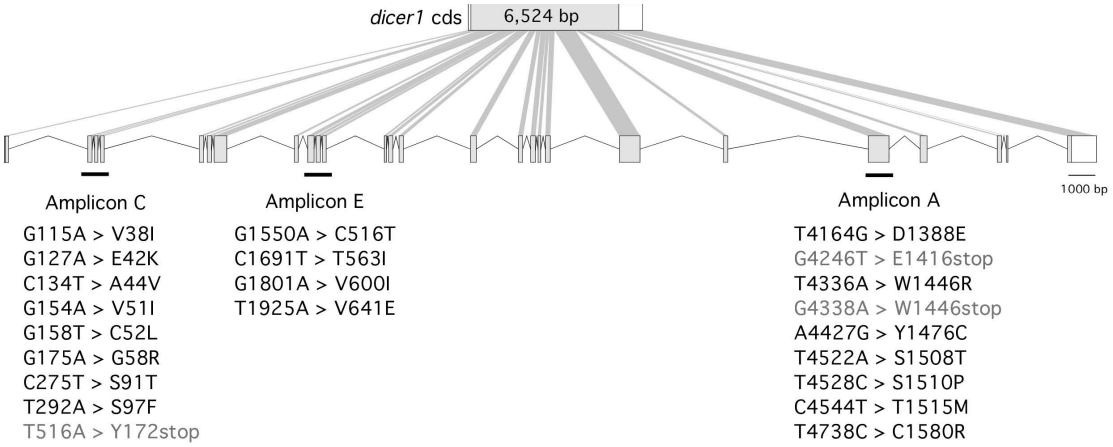


Fig. 5 Target-selected inactivation of the *dicer1* gene. *Dicer1* mRNA and genomic organization are schematically drawn. Exons in the genomic structure are indicated as boxes. The three fragments screened by TILLING are underlined. Nucleotide and amino acid positions are given with regard to the predicted start codon. On homozygosing, the three nonsense mutations were all loss of function (Wienholds et al., 2003a).

4. Ideally, one could rescue a mutation by transgenesis with the wild-type gene. In practice, this is not needed to prove causality, if the first two or three points are followed.

5. A final argument to consider, although with caution, is that one might have good reasons to expect a very specific phenotype. For the *rag1* gene, required for V(D)J joining in mice and humans, it was not unreasonable to expect that the mutant would fail to show V(D)J joining and be immunodeficient; when it had those phenotypes, a causal relation seemed likely. With more common phenotypes (such as lethality), this reasoning does not apply.

4.2. Null Alleles, Weak Alleles, or Silent Alleles

The first goal of most gene targeting work is to obtain a guaranteed null or loss-of-function allele. We prefer to focus on stop mutations, or (less preferred) splice site mutations. With stops the only concern is that an alternative splice removes the part of the exon that contains the stop. Therefore, an extra safeguard might be to focus initial mutant searches on exon domains encoding a conserved portion of a protein, so that any alternative splice that bypasses the mutation will not encode a functional protein. In some cases, knowledge of protein structure might help predict that a missense mutation can reasonably be expected to be a null or strong reduction of function. Again, as mentioned previously; having two different stop mutations with the same phenotype is a strong argument that both are null.

In practice, approximately 1 in 20 mutations in coding regions is a stop; the missense mutations can be highly valuable as well. To enhance possible weak phenotypic effects of such mutations, one can cross them to a null allele and analyze the phenotype of the transheterozygote (it will have half the gene function of the homozygous missense mutant). In theory, one might be able to sort a series of missense mutants into an allelic series of weak to strong mutants. If the null allele is lethal or sterile, and therefore in some cases of limited use for some experimental studies, weaker alleles might be quite useful.

5. Materials and Methods

5.1. Zebrafish Mutagenesis and Libraries

Zebrafish are raised under standard conditions. Adult wild-type (TL) male zebrafish (4 months old) are mutagenized by six consecutive treatments with 3.0 mM ENU as described (van Eeden *et al.*, 1999). For the specific locus test, the surviving and fertile fish are outcrossed with *albino* (*alb*) or *sparse* (*spa*) females. To generate F1 progeny for the library, the mutagenized males are outcrossed against wild-type (TL) females. To prevent clonal mutants, it is recommended to raise a maximum of a few hundred progeny per mutagenized F1 male. For a living library, all healthy looking adult fish are fin-clipped and stored in pools of 12. This can be more if the aquarium facility is limiting. To construct a permanently frozen library, the sperm of the mutagenized males is cryopreserved according to the protocol described by Morris *et al.* (2003).

5.2. Genomic DNA isolation

For a living library, genomic DNA is isolated from finclips in 96 deep-well plates (1.0 ml capacity per well). Freshly cut finclips are directly transferred to plates that are kept on dry ice. Fins are lysed by incubation in 400 µl pre-warmed lysis buffer (100 mM Tris-HCl pH 8-8.5, 200 mM NaCl, 0.2% SDS, 5 mM EDTA and 100 µg/ml proteinase K) at 55°C for at least 3 hrs with occasional vortexing. DNA is precipitated by adding 300 µl isopropanol and centrifugation at >6,000 x g, followed by washing with 70% ethanol. Finally, pellets are dissolved in 500 µl TE. For screening, the DNA is diluted 10 times and 5-µl aliquots are arrayed in 384-well PCR plates using a 96-channel pipet (HYDRA-96, Robbins Scientific). PCR plates are covered with aluminum foil tape (3M) and stored at -20°C. DNA from a permanently frozen library can be isolated similarly from complete tails (approximately 1.0 cm), except that tails are lysed in 1.0 ml lysis buffer in 2.0 ml 96 deep-well plates and precipitated with 700 µl isopropanol. DNA is dissolved in 200 µl TE and diluted 50 fold (in water) prior to screening

5.3. Screening by resequencing

Gene-specific primers are designed to amplify most of the coding regions of target genes by a nested PCR approach. In the first PCR, target genes are amplified in 384 well PCR plates with a standard PCR program (94°C for 60 sec; 30 cycles of 92°C for 20 sec, 58°C for 30 and 72°C for 60 sec and a additional extension step of 72°C for 180 sec; GeneAmp9700, Applied Biosystems). PCR samples contain 5 µl genomic DNA, 0.2 µM forward (f1) and 0.2 µM reverse (r1) primer, 200 µM of each dNTP, 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM $MgCl_2$ and 0.2 U Taq DNA polymerase in a total volume of 10 µl. For the nested PCR, a small volume is transferred by 384-well replicators to a new 5 µl PCR mixture with internal gene-specific primers (f2 and r2). The nested PCR is empirically optimized to have maximal yield and have minimal residual primers, dNTPS and Taq DNA polymerase left after cycling (typically this is 0.1 µM of each primer, 50 µM dNTPs and 0.1 U Taq DNA polymerse). The cycling program is identical to that of the first PCR. Next, PCR fragments are diluted 10-fold with water and from this 1 µl is used as template in a sequencing reaction. Alternatively, PCR fragments can first be purified with DNA binding filterplates (Itoh *et al.*, 1997) according to manufactures' instructions (Whatman). Sequencing reactions contain 0.25 to 0.35 µl BigDye Terminator (Applied Biosystems), 3.75 to 3.65 µl dilution buffer (200 mM Tris-HCl, pH9.0 and 10 mM $MgCl_2$), 5% DMSO and 0.5 µM of one of the nested primers (f2 or r2) in a total volume of 10 µl. Cycling conditions are as recommended by the manufacturer. Sequencing products are purified using Sephadex G50 (superfine coarse) minicolumns and analyzed on a 96-capillary 3700 DNA analyzer (Applied Biosystems) for which the running time is

adjusted to the fragment lengths. Mutations are found by parsing the trace files through Phred (Ewing *et al.*, 1998), Phrap (Gordon *et al.*, 2001) and Polyphred (Nickerson *et al.*, 1997) and are visualized with Consed (Gordon *et al.*, 1998).

5.4. Screening by TILLING

CEL-I enzyme is isolated from celery according to Oleykowski *et al.* (1998) and Yang *et al.* (2000) with minor modifications (for a detailed protocol see: http://www.niob.knaw.nl/ researchpages/cuppen/cel1.html). ENU-induced mutations are screened by using CEL-I-mediated heteroduplex cleavage, as described for *Arabidopsis* (Colbert *et al.*, 2001), but with several adaptations described next. All pipeting steps are done on a Genesis Workstation 200 (Tecan) and Microlab 2200 (Hamilton) or using multichannel pipets. Target genes are amplified by a nested PCR approach in 384 well plates. In the first PCR with gene specific primers, a touchdown cycling program is used (94°C for 60 sec; 30 cycles of 94°C for 20 sec, 65°C for 30 sec with a decrement of 0.5°C per cycle, and 72°C for 60 sec; followed by 10 cycles of 94°C for 20 sec, 58°C for 30 sec and 72°C for 60 sec and an additional extension step of 72°C for 180 sec; GeneAmp9700, Applied Biosystems). PCR samples contain 5 µl genomic DNA, 0.2 µM forward (f1) and 0.2 µM reverse (r1) primer, 200 µM of each dNTP, 25 mM Tricine, 7.0% Glycerol (m/v), 1.6% DMSO (m/v), 2 mM $MgCl_2$, 85 mM NH_4Acetate pH8.7 and 0.2 U Taq DNA polymerase in a total volume of 10 µl.

After the first PCR reactions the samples are diluted with 20 µl water and 1 µl is used as template for the second nested PCR reaction. This reaction contains a mixture of gene-specific forward (M13F-f2, 0.08 µM) and reverse (M13R-r2, 0.04 µM) primers that contain universal M13 adaptor sequences at their 5´ end, and the two corresponding universal M13F (5´-TGTAAAACGACGGCCAGT, 0.12µM) and M13R (5´-AGGAAACAGCTATGACCAT, 0.16µM) primers labeled with fluorescent dyes (IR Dye 700 and IR Dye 800, respectively) for detection. In addition, the PCR samples contain 200 µM of each dNTP, 20 mM Tris-Hcl pH8.4, 50 mM KCl, 1.5 mM $MgCl_2$ and 0.1 U Taq DNA polymerase in a total volume of 5 µl. Standard cycling conditions are used for the nested PCR reactions (30 cycles of 94°C for 20 sec, 58°C for 30 sec and 72°C for 60 sec, followed by an additional extension step of 72°C for 180 sec).

Directly following the nested PCR, heteroduplex formation is done by incubation at 99°C for 10 min and 70 cycles of 70°C for 20 sec with a decrement of 0.3°C per cycle. Next, 1.25 µl aliquots of four individual PCR reactions are pooled (total volume of 5 µl) and incubated with 0.01 µl CEL-I enzyme solution in a total volume of 15 µl (buffered in 10 mM Hepes pH7.0, 10 mM $MgSO_4$, 10 mM KCl, 0.002% Triton X-100, 0.2 µg/ml BSA) at 45°C for 15 min. CEL-I reactions are stopped by adding 5 µl 75 mM EDTA. Fragments are purified using Sephadex G50 (medium coarse) minicolumns in 96-wells filter plates (Multiscreen HV, Millipore) and eluted into plates prefilled with 5 µl formamide loading buffer [37% (v/v) de-ionized formamide, 4 mM EDTA pH8.0, 90 µg/ml bromophenol blue] per well or purified by isopropanol precipitation. Samples are concentrated to about 1 µl by heating at 85°C for 45-60 min without cover. A 0.4 µl sample is applied to a 96-lane membrane comb (The Gel Company) and loaded on 25 cm denaturing 6% polyacrylamide gels on LI-COR 4200 DNA analyzers. Raw TIFF-images produced by the analyzers are manipulated using Adobe Photoshop and potential mutations are detected and scored manually.

Acknowledgments

We thank Ewart de Bruijn for help with collecting ENU data from forward genetic screens, Elizabeth Greene for advice on expectations after ENU mutagenesis, our collaborators for sharing (unpublished) sequences, and Edwin Cuppen and Robin May for useful comments on this manuscript. This work is sponsored by a NWO Genomics grant.

References

Alonso, J. M., et al. (2003). Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science* **301,** 653-7.

Altschul, S. F., et al. (1990). Basic local alignment search tool. *J Mol Biol* **215,** 403-10.

Amsterdam, A. (2003). Insertional mutagenesis in zebrafish. *Dev Dyn* **228,** 523-34.

Amsterdam, A., et al. (1999). A large-scale insertional mutagenesis screen in zebrafish. *Genes Dev* **13,** 2713-24.

Ballinger, D. G., and Benzer, S. (1989). Targeted gene mutations in Drosophila. *Proc Natl Acad Sci U S A* **86,** 9402-6.

Beier, D. R. (2000). Sequence-based analysis of mutagenized mice. *Mamm Genome* **11,** 594-7.

Bentley, A., MacLennan, B., Calvo, J., and Dearolf, C. R. (2000). Targeted recovery of mutations in Drosophila. *Genetics* **156,** 1169-73.

Brand, M., Granato, M., and Nusslein-Volhard, C. (2002). Keeping and raising zebrafish. *In* "Zebrafish" (C. Nusslein-Volhard and D. Ralf, Eds.), Vol. 261, pp. 7-37. Oxford University Press, Oxford.

Capecchi, M. R. (1989). Altering the genome by homologous recombination. *Science* **244,** 1288-92.

Chasman, D., and Adams, R. M. (2001). Predicting the functional consequences of non-synonymous single nucleotide polymorphisms: structure-based assessment of amino acid variation. *J Mol Biol* **307,** 683-706.

Chen, Y., et al. (2000). Genotype-based screen for ENU-induced mutations in mouse embryonic stem cells. *Nat Genet* **24,** 314-7.

Coghill, E. L., et al. (2002). A gene-driven approach to the identification of ENU mutants in the mouse. *Nat Genet* **30,** 255-6.

Colbert, T., et al. (2001). High-throughput screening for induced point mutations. *Plant Physiol* **126,** 480-4.

Cui, Z., et al. (2003). RecA-mediated, targeted mutagenesis in zebrafish. *Mar Biotechnol (NY)* **5,** 174-84.

Dai, Y., et al. (2002). Targeted disruption of the alpha1,3-galactosyltransferase gene in cloned pigs. *Nat Biotechnol* **20,** 251-5.

Das, L., and Martienssen, R. (1995). Site-selected transposon mutagenesis at the hcf106 locus in maize. *Plant Cell* **7,** 287-94.

Davidson, A. E., et al. (2003). Efficient gene delivery and gene expression in zebrafish using the Sleeping Beauty transposon. *Dev Biol* **263,** 191-202.

De Stasio, E., et al. (1997). Characterization of revertants of unc-93(e1500) in Caenorhabditis elegans induced by N-ethyl-N-nitrosourea. *Genetics* **147,** 597-608.

de Wind, N., et al. (1995). Inactivation of the mouse Msh2 gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer. *Cell* **82,** 321-30.

Douglas, G. R., et al. (1995). Temporal and molecular characteristics of mutations induced by ethylnitrosourea in germ cells isolated from seminiferous tubules and in spermatozoa of lacZ transgenic mice. *Proc Natl Acad Sci U S A* **92,** 7485-9.

Driever, W., et al. (1996). A genetic screen for mutations affecting embryogenesis in zebrafish. *Development* **123,** 37-46.

Elbashir, S. M., et al. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411,** 494-8.

Ewing, B., and Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* **8,** 186-94.

Golling, G., et al. (2002). Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development. *Nat Genet* **31,** 135-40.

Gordon, D., Abajian, C., and Green, P. (1998). Consed: a graphical tool for sequence finishing. *Genome Res* **8,** 195-202.

Gordon, D., Desmarais, C., and Green, P. (2001). Automated finishing with autofinish. *Genome Res* **11,** 614-25.

Haffter, P., et al. (1996). The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio. *Development* **123,** 1-36.

Hanada, K., et al. (1997). RecQ DNA helicase is a suppressor of illegitimate recombination in Escherichia coli. *Proc Natl Acad Sci U S A* **94,** 3860-5.

Itoh, M., et al. (1997). Simple and rapid preparation of plasmid template by a filtration method using microtiter filter plates. *Nucleic Acids Res* **25,** 1315-6.

Jansen, G., Hazendonk, E., Thijssen, K. L., and Plasterk, R. H. (1997). Reverse genetics by chemical mutagenesis in Caenorhabditis elegans. *Nat Genet* **17,** 119-21.

Kaiser, K., and Goodwin, S. F. (1990). "Site-selected" transposon mutagenesis of Drosophila. *Proc Natl Acad Sci U S A* **87,** 1686-90.

Kimmel, C. B. (1989). Genetics and early development of zebrafish. *Trends Genet* **5,** 283-8.

Koes, R., et al. (1995). Targeted gene inactivation in petunia by PCR-based selection of transposon insertion mutants. *Proc Natl Acad Sci U S A* **92,** 8149-53.

Korswagen, H. C., Durbin, R. M., Smits, M. T., and Plasterk, R. H. (1996). Transposon Tc1-derived, sequence-tagged sites in Caenorhabditis elegans as markers for gene mapping. *Proc Natl Acad Sci U S A* **93,** 14680-5.

Lander, E. S., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* **409,** 860-921.

Lee, K. Y., et al. (2002). Cloned zebrafish by nuclear transfer from long-term-cultured cells. *Nat Biotechnol* **20,** 795-9.

Lekven, A. C., et al. (2000). Reverse genetics in zebrafish. *Physiol Genomics* **2,** 37-48.

Liu, L. X., et al. (1999). High-throughput isolation of Caenorhabditis elegans deletion mutants. *Genome Res* **9,** 859-67.

Ma, C., et al. (2001). Production of zebrafish germ-line chimeras from embryo cell cultures. *Proc Natl Acad Sci U S A* **98,** 2461-6.

Martin, E., et al. (2002). Identification of 1088 new transposon insertions of Caenorhabditis elegans: a pilot study toward large-scale screens. *Genetics* **162,** 521-4.

McCallum, C. M., Comai, L., Greene, E. A., and Henikoff, S. (2000). Targeted screening for induced mutations. *Nat Biotechnol* **18,** 455-7.

McCreath, K. J., et al. (2000). Production of gene-targeted sheep by nuclear transfer from cultured somatic cells. *Nature* **405,** 1066-9.

Meissner, R. C., et al. (1999). Function search in a large transcription factor gene family in Arabidopsis: assessing the potential of reverse genetics to identify insertional mutations in R2R3 MYB genes. *Plant Cell* **11,** 1827-40.
Miao, Z. H., and Lam, E. (1995). Targeted disruption of the TGA3 locus in Arabidopsis thaliana. *Plant J* **7,** 359-65.
Morris, J. P. t., et al. (2003). Zebrafish sperm cryopreservation with N,N-dimethylacetamide. *Biotechniques* **35,** 956-8, 960, 962 passim.
Mullins, M. C., Hammerschmidt, M., Haffter, P., and Nusslein-Volhard, C. (1994). Large-scale mutagenesis in the zebrafish: in search of genes controlling development in a vertebrate. *Curr Biol* **4,** 189-202.
Nakamura, Y., Gojobori, T., and Ikemura, T. (2000). Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res* **28,** 292.
Nasevicius, A., and Ekker, S. C. (2000). Effective targeted gene 'knockdown' in zebrafish. *Nat Genet* **26,** 216-20.
Ng, P. C., and Henikoff, S. (2001). Predicting deleterious amino acid substitutions. *Genome Res* **11,** 863-74.
Nickerson, D. A., Tobe, V. O., and Taylor, S. L. (1997). PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nucleic Acids Res* **25,** 2745-51.
Oleykowski, C. A., Bronson Mullins, C. R., Godwin, A. K., and Yeung, A. T. (1998). Mutation detection using a novel plant endonuclease. *Nucleic Acids Res* **26,** 4597-602.
Parinov, S., et al. (1999). Analysis of flanking sequences from dissociation insertion lines: a database for reverse genetics in Arabidopsis. *Plant Cell* **11,** 2263-70.
Perry, J. A., et al. (2003). A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume Lotus japonicus. *Plant Physiol* **131,** 866-71.
Raz, E., et al. (1998). Transposition of the nematode Caenorhabditis elegans Tc3 element in the zebrafish Danio rerio. *Curr Biol* **8,** 82-8.
Rong, Y. S., and Golic, K. G. (2000). Gene targeting by homologous recombination in Drosophila. *Science* **288,** 2013-8.
Rong, Y. S., and Golic, K. G. (2001). A targeted gene knockout in Drosophila. *Genetics* **157,** 1307-12.
Sessions, A., et al. (2002). A high-throughput Arabidopsis reverse genetics system. *Plant Cell* **14,** 2985-94.
Smits, B. M. G., Mudde, J., Plasterk, R. H. A., and Cuppen, E. (2003). Target-selected mutagenesis of the rat. *Genomics* **83,** 332-334.
Solnica-Krezel, L., Schier, A. F., and Driever, W. (1994). Efficient recovery of ENU-induced mutations from the zebrafish germline. *Genetics* **136,** 1401-20.
Spradling, A. C., et al. (1999). The Berkeley Drosophila Genome Project gene disruption project: Single P-element insertions mutating 25% of vital Drosophila genes. *Genetics* **153,** 135-77.
Streisinger, G., et al. (1981). Production of clones of homozygous diploid zebra fish (Brachydanio rerio). *Nature* **291,** 293-6.
Sun, L., et al. (1995). ES-like cell cultures derived from early zebrafish embryos. *Mol Mar Biol Biotechnol* **4,** 193-9.
Sunyaev, S., et al. (2001). Prediction of deleterious human alleles. *Hum Mol Genet* **10,** 591-7.
Thompson, S., et al. (1989). Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. *Cell* **56,** 313-21.
Till, B. J., et al. (2003). Large-Scale Discovery of Induced Point Mutations With High-Throughput TILLING. *Genome Res* **13,** 524-30.
Tissier, A. F., et al. (1999). Multiple independent defective suppressor-mutator transposon insertions in Arabidopsis: a tool for functional genomics. *Plant Cell* **11,** 1841-52.
van Eeden, F. J., Granato, M., Odenthal, J., and Haffter, P. (1999). Developmental mutant screens in the zebrafish. *Methods Cell Biol* **60,** 21-41.
Venter, J. C., et al. (2001). The sequence of the human genome. *Science* **291,** 1304-51.
Wienholds, E., et al. (2003a). The microRNA-producing enzyme Dicer1 is essential for zebrafish development. *Nat Genet* **35,** 217-8.
Wienholds, E., Schulte-Merker, S., Walderich, B., and Plasterk, R. H. (2002). Target-selected inactivation of the zebrafish rag1 gene. *Science* **297,** 99-102.
Wienholds, E., et al. (2003b). Efficient Target-Selected Mutagenesis in Zebrafish. *Genome Res*.
Yanagihara, K., and Mizuuchi, K. (2002). Mismatch-targeted transposition of Mu: a new strategy to map genetic polymorphism. *Proc Natl Acad Sci U S A* **99,** 11317-21.
Yang, B., et al. (2000). Purification, cloning, and characterization of the CEL I nuclease. *Biochemistry* **39,** 3533-41.
Zambrowicz, B. P., et al. (1998). Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells. *Nature* **392,** 608-11.
Zan, Y., et al. (2003). Production of knockout rats using ENU mutagenesis and a yeast-based screening assay. *Nat Biotechnol* **21,** 645-51.
Zwaal, R. R., et al. (1993). Target-selected gene inactivation in Caenorhabditis elegans by using a frozen transposon insertion mutant bank. *Proc Natl Acad Sci U S A* **90,** 7431-5.

Chapter 5

The microRNA-producing enzyme Dicer1 is essential for zebrafish development

Reprinted from Wienholds *et al*. (2003), *Nat Genet 35:217-218*

The microRNA-producing enzyme Dicer1 is essential for zebrafish development

Erno Wienholds, Marco J Koudijs, Freek J M van Eeden, Edwin Cuppen & Ronald H A Plasterk

MicroRNAs (miRNAs) are produced by the Dicer1 enzyme; the role of Dicer1 in vertebrate development is unknown. Here we report target-selected inactivation of the *dicer1* gene in zebrafish. We observed an initial build-up of miRNA levels, produced by maternal Dicer1, in homozygous *dicer1* mutants, but miRNA accumulation stopped after a few days. This resulted in developmental arrest around day 10. These results indicate that miRNA-producing Dicer1 is essential for vertebrate development.

MicroRNAs have essential roles in the development of plants[1], nematodes[2] and flies[3]. These miRNAs are produced by the Dicer1 enzyme[4–7], which is conserved from fungi to vertebrates. No genetic analysis of Dicer1 has been done in vertebrates, however, and its role in vertebrate development is not known. We cloned the zebrafish *dicer1* ortholog (**Fig. 1**) and applied a method for target-selected gene inactivation that we recently developed[8]. We identified three different premature stop alleles induced by N-ethyl-N-nitrosourea, all of which probably cause loss of function (**Fig. 1**). The alleles had identical recessive phenotypes (developmental arrest), did not complement each other and cosegregated with the phenotype. Of 49 fish with a mutant phenotype, 41 were *dicer1*$^{-/-}$ (wild-type larvae occasionally arrest in development). Of 112 fish with normal phenotype, none were *dicer1*$^{-/-}$. These data show that the phenotype is caused by disruption of *dicer1*.

Homozygous and trans-heterozygous (having two different null alleles) *dicer1*$^{-/-}$ embryos appeared normal during the first week but at 8 d post fertilization (d.p.f.) had lethargic behavior and developmental growth arrest. Most *dicer1*$^{-/-}$ embryos died after 14–15 d (**Fig. 2a**), and no *dicer1*$^{-/-}$ fish were alive after 3 weeks of development (n = 112). Microscopic analysis indicated no obvious defect in one specific organ but rather a general arrest of growth (**Fig. 2a**).

Morpholino knockdown experiments resulted in an earlier arrest, indicating that maternal *dicer1* mRNA is necessary for embryonic development (**Fig. 2c** and **Supplementary Fig. 1** and **Supplementary Note** online). Strong expression of maternal *dicer1* mRNA and ubiquitous expression up to 2 d.p.f. has been detected by whole-mount *in situ* hybridizations (data not shown). Morpholinos target only mRNA; thus, we cannot rule out the possibility that maternal Dicer1 protein is essential for the earliest stages of embryogenesis.

Dicer1 may be essential in vertebrate development because it is required to make miRNAs. Hundreds of miRNAs are expected to be present in the fish[9,10], and many of them may act in different organs and stages of development. There was a swift build-up of miRNA levels in the first days of development of heterozygous fish (**Fig. 2b**). In *dicer1* mutants, the initial build-up was the same, presumably by maternal Dicer, but production was arrested after ~1 week. The *mir-26a* precursor RNA was not observed in heterozygotes but began to accumulate in *dicer1* mutants at 4 d.p.f. The difference in levels of the

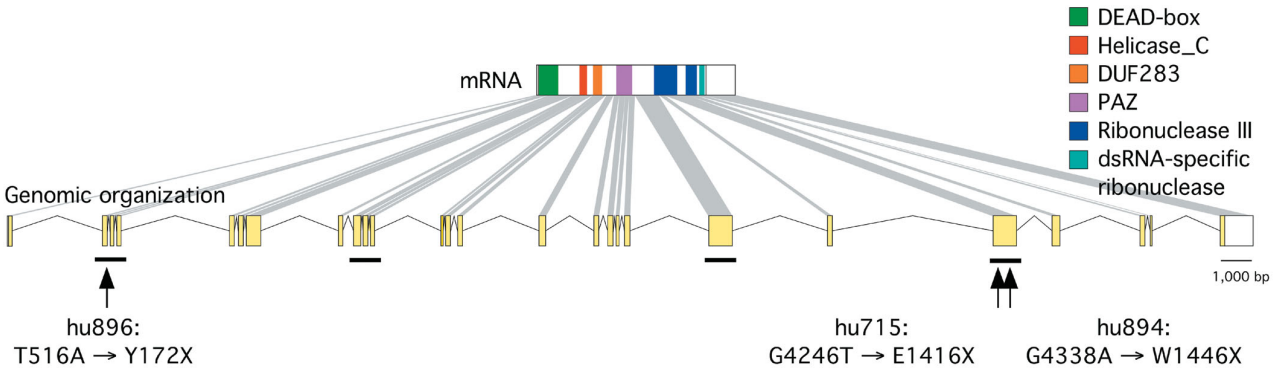


Figure 1 Target-selected inactivation of the zebrafish *dicer1* gene. Conserved domains are schematically indicated in color in the *dicer1* mRNA sequence. In the genomic organization, exons are represented as boxes and introns as lines (intron lengths are estimated). Fragments screened for mutations are underlined. Arrows indicate premature stop codons. Positions of the nucleotide and corresponding amino acid changes are given with regard to the predicted start codon of the mRNA.

Department of Functional Genomics, Hubrecht Laboratory, 3584 CT Utrecht, The Netherlands. Correspondence should be addressed to R.H.A.P (plasterk@niob.knaw.nl).

Published online 5 October 2003; doi:10.1038/ng1251

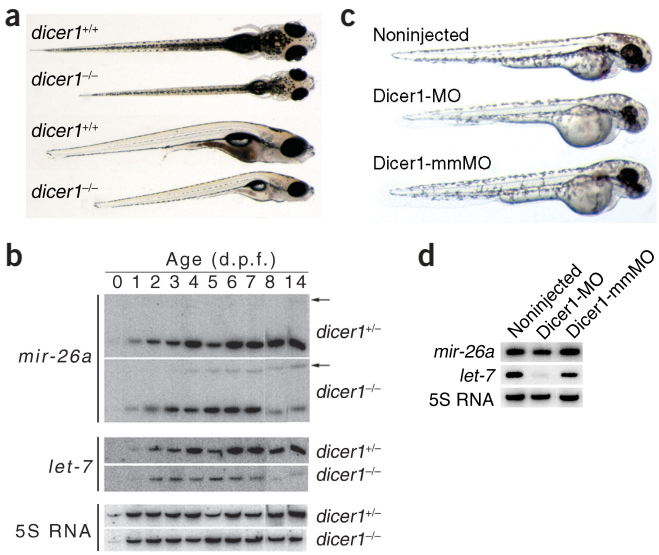


Figure 2 *Dicer1* knockout and knockdown zebrafish undergo developmental arrest and have reduced levels of miRNAs. (**a**) Dorsal and lateral views of wild-type and homozygous *dicer1* (*dcr*$^{hu715/hu715}$) knockout fish at 14 d.p.f. Developmental growth arrest is visible as reduced size and pigmentation. (**b**) Northern-blot analysis of miRNA expression during 1–8 d.p.f. and at 14 d.p.f. of heterozygous and homozygous *dicer1* knockout fish. Arrows indicate the precursor of *mir-26a*; the precursor of *let-7* was not observed. (**c**) Lateral views of embryos injected with morpholino (Dicer1-MO) and mismatched morpholino (Dicer1-mmMO) at 2 d.p.f. (**d**) Northern-blot analysis of miRNA expression of embryos injected with morpholino (Dicer1-MO) and mismatched morpholino (Dicer1-mmMO) at 2 d.p.f.

processed form of *let-7* was more marked. This might be because *let-7* is expressed later than *mir-26a* (data not shown) and, therefore, is not fully processed by maternal Dicer1. Tissue-specific expression may also account for differences between the miRNAs tested. We investigated whether morpholino-induced knockdown of maternal *dicer1* mRNA resulted in earlier effects. We found that *let-7* miRNA was almost absent at 2 d.p.f., in agreement with the earlier phenotypic growth delay (**Fig. 2d**). This effect was not seen for *mir-26a*. This may be because *mir-26a* is expressed earlier than *let-7* and is presumably still processed by maternal Dicer1 protein. Again, tissue-specific expression may also be causal.

This work shows that Dicer1 is essential in zebrafish development. Dicer1 mutants experience overall growth arrest. This is probably caused by the failure to produce miRNAs, but failure of additional functions of Dicer1, such as the production of small interfering RNAs[7], may also influence development. There are two, not mutually exclusive, explanations for the arrest caused by the failure to produce miRNAs: either some miRNAs are required for all cells to proliferate or each tissue or organ requires different miRNAs and all are depleted at approximately the same time. A future study should address the tissue-specific roles of different miRNAs in zebrafish development.

Material and methods are provided in **Supplementary Methods** online.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank M. Koster, J. Mudde, M. den Broeder and A. Keijser for help with finding of the Dicer mutants; the Sanger Centre for sequencing the BAC containing the *dicer1* gene; and D. Jongejan-Zivkovic, M. Tijsterman and R. Ketting for discussions. This work is supported by a grant from the Netherlands Organization for Scientific research (NWO).

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 6 August; accepted 16 September 2003
Published online at http://www.nature.com/naturegenetics/

1. Reinhart, B.J., Weinstein, E.G., Rhoades, M.W., Bartel, B. & Bartel, D.P. *Genes Dev.* **16**, 1616–1626 (2002).
2. Reinhart, B.J. *et al. Nature* **403**, 901–906 (2000).
3. Brennecke, J., Hipfner, D.R., Stark, A., Russell, R.B. & Cohen, S.M. *Cell* **113**, 25–36 (2003).
4. Grishok, A. *et al. Cell* **106**, 23–34 (2001).
5. Hutvagner, G. *et al. Science* **293**, 834–838 (2001).
6. Ketting, R.F. *et al. Genes Dev.* **15**, 2654–2659 (2001).
7. Bernstein, E., Caudy, A.A., Hammond, S.M. & Hannon, G.J. *Nature* **409**, 363–366 (2001).
8. Wienholds, E., Schulte-Merker, S., Walderich, B. & Plasterk, R.H. *Science* **297**, 99–102 (2002).
9. Lim, L.P., Glasner, M.E., Yekta, S., Burge, C.B. & Bartel, D.P. *Science* **299**, 1540 (2003).
10. Lagos-Quintana, M., Rauhut, R., Lendeckel, W. & Tuschl, T. *Science* **294**, 853–858 (2001).

Supporting Online Material

Supplementary Figure 1 Knockdown of *dicer* results in developmental growth arrest. Lateral views of morpholino (dicer-MO) and mismatched morpholino (dicer-mmMO) injected embryos at 4 d.p.f. Non-injected embryos at 3 and 4 d.p.f. are shown for comparison.

Note: Morpholino knockdown experiments

Two different antisense morpholinos, designed against the 5' UTR and the predicted start site of the *dicer* mRNA, were injected into one-cell stage embryos. Morpholino injected embryos start to show a delay around 1 d.p.f. (in more than 95% of the cases) and form underdeveloped structures in different tissues in later stages (e.g. the eyes, see **Fig. 2*c***). At 4 d.p.f. the embryos have a developmental delay of approximately 1 day (**Supplementary Fig. 1**). Thereafter embryos start to look severely malformed and eventually die around 7 d.p.f. Mismatched morpholino injected embryos show an initial delay, but fully recover and develop normally (**Fig. 2c** and **Supplementary Fig. 1**). Morpholino injections in *dcr*$^{-/-}$ background did not result in a different or enhanced phenotype (data not shown). A possibility for this is that zygotic transcription of *dicer* mRNA and the presence of morpholinos are not fully overlapping.

Materials and Methods

Cloning of the zebrafish dicer gene

The zebrafish *dicer* gene sequence was determined by several approaches. First, Blast searches[1] using the human gene as query were performed to obtain EST clones containing the zebrafish *dicer* cDNA. One clone (fc39d11) containing the 3'end of the zebrafish *dicer* cDNA was fully sequenced. Additional upstream dicer cDNA sequence was obtained by RT-PCR with primers designed in *dicer* containing traces from the whole genome shotgun library. BAC filters were screened with the EST clone as probe. One of the *dicer* positive BACs (zK51K10) was sequenced to gain the 5' end of the mRNA sequence and the genomic organization of the *dicer* gene. Finally intron/exon boundaries were determined by Genotrace software[2]. Blast searches against the zebrafish genome (Ensembl genome build, July 2003) indicated the presence of only one copy of *dicer* in the zebrafish genome.

Target-selected inactivation of the dicer gene

Four fragments of the *dicer* gene were selected for target-selected mutagenesis[3]. These fragments were screened for mutations by TILLING[4] in a library of 4,608 ENU-mutagenized F1 fish (E.W., F.v.E., R.H.A.P. and E.C.; manuscript in preparation). Fish with mutant alleles (*dcr*$^{hu715/+}$, *dcr*$^{hu894/+}$ and *dcr*$^{hu896/+}$) were outcrossed against wild-type fish and subsequently incrossed or directly incrossed against each other to obtain heteroallelic offspring. Genotyping was done by resequencing.

Morphlino injections

Two different, partly overlapping, antisense morpholinos (Gene Tools) were designed against the 5' UTR of the dicer mRNA (Dicer-MO1, CTGTAGGCCAGCCATGCTTAGAGAC and Dicer-MO2, GCTTAGAGACTGATAAGCAGGAGAC respectively). As control two corresponding mismatched morpholinos were designed (Dicer-mmMO1, CTCTAGGCCTGCCATCCTTAGTGAC and Dicer-mmMO2, GCTTTGAGAGTGATTAGCAGCAGAC respectively). Morpholinos were injected into the 1-cell stage of wild-type embryos derived from the TL line. Dicer-MO1 and Dicer-MO2 were injected in a range from 1 to 25 ng. Injection of both morpholinos gave the same phenotype. At 5ng consistent results were obtained without obvious nonspecific effects. This concentration was used for phenotypic and miRNA analysis. Both mismatched morpholinos could be injected at 10 ng without giving nonspecific effects.

Northern blot analysis

Total RNA from embryos and larvae was isolated using TRIzol Reagent (Invitrogen). From the genetic mutants RNA of single embryos/larvae was used for analysis. From the morpholino-injected embryos pooled RNA of 10 embryos was used for analysis. RNA was separated on 15% polyacrylamide gels. Radiolabeled probes complementary to *let-7* (TTCCAAACTATACAACCTACTACCTCACCGGATCCG), *mir-26a* (AGCCTATCCTGGATTACTTGAA) and 5S RNA (ATCGGACGAGATCGGGCGTA) were used for hybridizations.

References

1. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *J Mol Biol* **215**, 403-10 (1990).
2. Berezikov, E., Plasterk, R.H. & Cuppen, E. GENOTRACE: cDNA-based local GENOme assembly from TRACE archives. *Bioinformatics* **18**, 1396-7 (2002).
3. Wienholds, E., Schulte-Merker, S., Walderich, B. & Plasterk, R.H. Target-selected inactivation of the zebrafish rag1 gene. *Science* **297**, 99-102 (2002).
4. Colbert, T. et al. High-throughput screening for induced point mutations. *Plant Physiol* **126**, 480-4 (2001).

Chapter 6

MicroRNA expression in zebrafish embryonic development

Reprinted from Wienholds *et al*. (2005), *Science 309:310-311*

transfer from vision to olfaction (Fig. 3E, bottom). Data for all memory transfer experiments are summarized in Fig. 3F. Thus, persistent memory for the conditioned cue is essential for crossmodal memory transfer.

The neural circuits and cellular mechanisms underlying the crossmodal enhancement and transfer of memory are unknown. Further understanding requires the elucidation of visual and olfactory circuits and their interconnection, as well as the locus for storage of visual and olfactory memory. It is possible that "multisensory integrative neuron" may also exist in the *Drosophila* brain and that crossmodal interaction between different sensory modalities may also be achieved through synchronized activity between modality-specific brain regions (*19*). Crossmodal interaction between sensory systems can enhance the detection and discrimination of external objects and can provide information about the environment that is unobtainable by a single modality in isolation. Our findings indicate that individual flies make use of crossmodal interactions between two sensory systems during operant conditioning, which further suggests that crossmodal interactions using multiple sensory systems may also facilitate learning in the natural environment. These results provide a basis for further studies of the circuit mechanisms underlying crossmodal interactions during memory acquisition.

References and Notes

1. G. A. Calvert, *Cereb. Cortex* **11**, 1110 (2001).
2. M. A. Frye, M. Tarsitano, M. H. Dickinson, *J. Exp. Biol.* **206**, 843 (2003).
3. M. V. Srinivasan, S. W. Zhang, H. Zhu, *Nature* **396**, 637 (1998).
4. M. Heisenberg, R. Wolf, B. Brembs, *Learn. Mem.* **8**, 1 (2001).
5. L. Liu, R. Wolf, R. Ernst, M. Heisenberg, *Nature* **400**, 753 (1999).
6. S. Tang, A. Guo, *Science* **294**, 1543 (2001).
7. S. Tang, R. Wolf, S. Xu, M. Heisenberg, *Science* **305**, 1020 (2004).
8. R. J. Greenspan, *Neuron* **15**, 747 (1995).
9. J. Dubnau, T. Tully, *Annu. Rev. Neurosci.* **21**, 407 (1998).
10. S. Waddell, W. G. Quinn, *Trends Genet.* **17**, 719 (2001).
11. R. L. Davis, *Neuron* **44**, 31 (2004).
12. A. Guo *et al.*, *Learn. Mem.* **3**, 49 (1996).
13. Materials and methods are available as supporting material on *Science* Online.
14. J. A. Gottfried, R. J. Dolan, *Neuron* **39**, 375 (2003).
15. B. E. Stein, M. A. Meredith, *Merging of the Senses* (MIT Press, Cambridge, MA, 1993).
16. M. A. Frye, M. H. Dickinson, *J. Exp. Biol.* **207**, 123 (2004).
17. B. Brembs, M. Heisenberg, *J. Exp. Biol.* **204**, 2849 (2001).
18. G. Hall, *Anim. Learn. Behav.* **24**, 233 (1996).
19. B. van Swinderen, R. J. Greenspan, *Nat. Neurosci.* **6**, 579 (2003).
20. Supported by the National Science Foundation of China (NSFC) (30270341), Multidisciplinary Program (Brain and Mind) of the Chinese Academy of Sciences (CAS), Major State Basic Research Program (G2000077800), and CAS (KJCX1-09-03). We thank M.-m. Poo for invaluable advice, Y.-q. Peng for programming, and C.-y. Yuan for assistance.

Supporting Online Material
www.sciencemag.org/cgi/content/full/309/5732/307/DC1
Materials and Methods

18 February 2005; accepted 11 May 2005
10.1126/science.1111280

MicroRNA Expression in Zebrafish Embryonic Development

Erno Wienholds,[1] Wigard P. Kloosterman,[1] Eric Miska,[2,3] Ezequiel Alvarez-Saavedra,[2] Eugene Berezikov,[1] Ewart de Bruijn,[1] H. Robert Horvitz,[2] Sakari Kauppinen,[4] Ronald H. A. Plasterk[1]*

MicroRNAs (miRNAs) are small noncoding RNAs, about 21 nucleotides in length, that can regulate gene expression by base-pairing to partially complementary mRNAs. Regulation by miRNAs can play essential roles in embryonic development. We determined the temporal and spatial expression patterns of 115 conserved vertebrate miRNAs in zebrafish embryos by microarrays and by in situ hybridizations, using locked-nucleic acid–modified oligonucleotide probes. Most miRNAs were expressed in a highly tissue-specific manner during segmentation and later stages, but not early in development, which suggests that their role is not in tissue fate establishment but in differentiation or maintenance of tissue identity.

Current estimates of miRNA gene numbers in vertebrates are as high as 500 (*1*), of which many are conserved, and miRNAs may regulate up to 30% of genes (*2*). The miRNA first discovered, *lin-4*, is involved in developmental timing in the nematode *Caenorhabditis elegans* (*3*). In mammals, miRNAs have been implicated in hematopoietic lineage differentiation (*4*) and homeobox gene regulation (*5*). Zebrafish that are defective in miRNA processing arrest in development (*6*). Recently, miRNAs were shown to be dispensable for cell fate determination, axis formation, and cell differentiation but are required for brain morphogenesis in zebrafish embryos (*7*). Together, these findings indicate that miRNAs can play essential roles in development. However, little is known about the individual roles of most miRNAs. To focus future miRNA studies, we determined the spatial and temporal expression patterns of 115 conserved vertebrate miRNAs (see online Material and Methods; table S1; table S2) in zebrafish embryos.

First, we determined the temporal expression of miRNAs during embryonic development by microarray analysis (Fig. 1A and fig. S1A). Up to segmentation [12 hours post fertilization (hpf)], most miRNAs could not be detected. Most miRNAs became visible 1 to 2 days after fertilization and showed strong expression when organogenesis is virtually completed (96 hpf). In adults, the majority of miRNAs remained expressed (Fig. 1A). In addition we determined the expression of miRNAs in dissected organs of adult fish. For some miRNAs, a high degree of tissue specificity was observed (figs. S1B and S2, and table S3).

In situ hybridization of miRNAs had thus far not been possible in animals. Recently LNA (locked-nucleic acid)–modified DNA oligonucleotide probes have been shown to increase the sensitivity for the detection of miRNAs by Northern blots (*8*). By Northern blots analysis and in situ hybridization, using LNA probes, we detected predominantly mature miRNAs, which were reduced in dicer knockout zebrafish (fig. S3). We used these LNA probes for the whole-mount in situ detection of the conserved vertebrate miRNAs in zebrafish embryos and made a catalog of miRNA expression patterns (fig. S4 and database S1).

Most miRNAs (68%) were expressed in a highly tissue-specific manner. For example, miR-140 was specifically expressed in the cartilage of the jaw, head, and fins, and its presence was entirely restricted to those regions (Fig. 1B and database S1). Representative examples are shown (Fig. 1C) of six miRNAs that were expressed in different organ systems: nervous system, digestive system, muscles, circulatory system, sensory organs, and excretory system. Even within organs, there is specificity, as exemplified in Fig. 1D, where miR-217 can be seen to be expressed in the exocrine pancreas, and miR-7 in the endocrine pancreas (Langerhans islets). More than half of the miRNAs (43) were expressed in (specific regions of) the central nervous system (fig. S4). Many miRNA genes are clustered in the genome and, therefore, are probably expressed as one primary transcript, and indeed, we observed that many such clustered genes showed identical or overlapping expression patterns (figs. S4 and S5). We compared the in situ data with microarray

[1]Hubrecht Laboratory, Centre for Biomedical Genetics, 3584 CT Utrecht, the Netherlands. [2]Howard Hughes Medical Institute, Department of Biology and McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [3]Wellcome Trust, Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge CB2 1QN, UK. [4]Wilhelm Johannsen Centre for Functional Genome Research, Institute of Medical Biochemistry and Genetics, University of Copenhagen, DK-2200 Copenhagen N, Denmark.

*To whom correspondence should be addressed. E-mail: plasterk@niob.knaw.nl

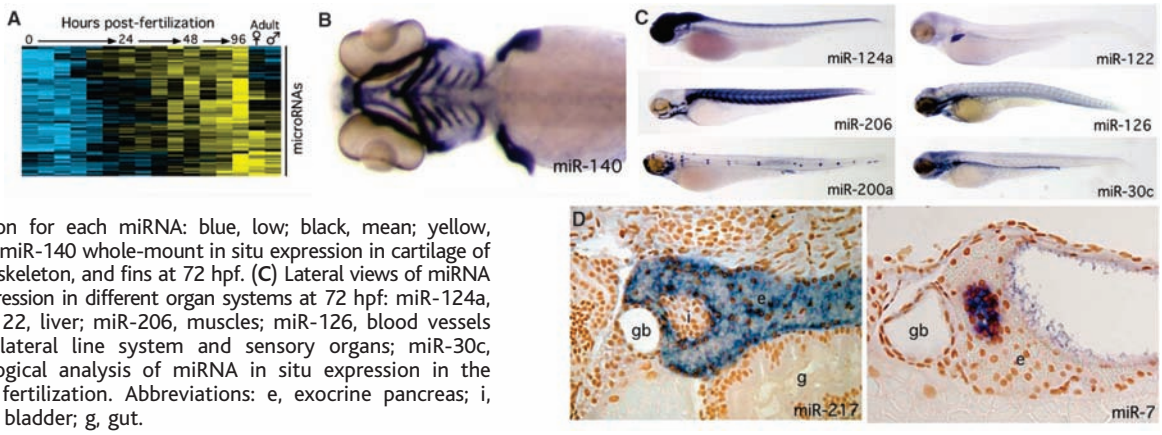


Fig. 1. miRNA expression in zebrafish embryonic development. (**A**) Microarray expression levels of 90 (of the 115) miRNAs during embryonic development. Colors indicate relative and mean-centered expression for each miRNA: blue, low; black, mean; yellow, high. (**B**) Ventral view of miR-140 whole-mount in situ expression in cartilage of pharyngeal arches, head skeleton, and fins at 72 hpf. (**C**) Lateral views of miRNA whole-mount in situ expression in different organ systems at 72 hpf: miR-124a, nervous systems; miR-122, liver; miR-206, muscles; miR-126, blood vessels and heart; miR-200a, lateral line system and sensory organs; miR-30c, pronephros. (**D**) Histological analysis of miRNA in situ expression in the pancreas 5 days after fertilization. Abbreviations: e, exocrine pancreas; i, pancreatic islet; gb, gall bladder; g, gut.

data for zebrafish and mammals (fig. S2 and table S3). Up to 77% of the in situ expression patterns were confirmed by at least one of the microarray data sets. In addition, miRNA in situ data showed patterns that cannot easily be detected by microarrays. For example, some miRNAs were expressed in hair cells of sensory epithelia (fig. S6).

In conclusion, we here describe the first comprehensive set of miRNA expression patterns in animal development. We found these patterns to be remarkably specific and diverse, which suggests highly specific and diverse roles for miRNAs. Most miRNAs are expressed in a tissue-specific manner during segmentation and later stages but were not detected during early development. Although we cannot exclude a role for undetectable early miRNAs, this observation indicates that most miRNAs may not be essential for tissue fate establishment but rather play crucial roles in differentiation or the maintenance of tissue identity.

References and Notes

1. E. Berezikov *et al.*, *Cell* **120**, 21 (2005).
2. B. P. Lewis, I. H. Shih, M. W. Jones-Rhoades, D. P. Bartel, C. B. Burge, *Cell* **115**, 787 (2003).
3. R. C. Lee, R. L. Feinbaum, V. Ambros, *Cell* **75**, 843 (1993).
4. C. Z. Chen, L. Li, H. F. Lodish, D. P. Bartel, *Science* **303**, 83 (2004).
5. S. Yekta, I. H. Shih, D. P. Bartel, *Science* **304**, 594 (2004).
6. E. Wienholds, M. J. Koudijs, F. J. van Eeden, E. Cuppen, R. H. Plasterk, *Nat. Genet.* **35**, 217 (2003).
7. A. J. Giraldez *et al.*, *Science* **308**, 833 (2005).
8. A. Valoczi *et al.*, *Nucleic Acids Res.* **32**, e175 (2004).

9. We thank D. Jongejan, F. van Eeden, S. Schulte-Merker, J. Bakkers, S. Chocron, N. Tolstrup, M. Bjørn, M. Bonde Mogensen, A. Hansen, W. Herzog, and ZFIN for technical assistance and/or help with annotation. We thank Exiqon for generously providing the LNA-modified probes. H.R.H. is an investigator of the Howard Hughes Medical Institute. This work was supported by a grant from the Netherlands Organization for Scientific Research (NWO) (E.W. and R.H.A.P.).

Supporting Online Material
www.sciencemag.org/cgi/content/full/1114519/DC1
Materials and Methods
Figs. S1 to S6
Tables S1 to S3
Database S1

6 May 2005; accepted 13 May 2005
Published online 26 May 2005;
10.1126/science.1114519
Include this information when citing this paper.

Ant Nestmate and Non-Nestmate Discrimination by a Chemosensory Sensillum

Mamiko Ozaki,[1]*† Ayako Wada-Katsumata,[1]* Kazuyo Fujikawa,[1] Masayuki Iwasaki,[2] Fumio Yokohari,[2] Yuji Satoji,[1] Tomoyosi Nisimura,[1] Ryohei Yamaoka[1]

In animal societies, chemical communication plays an important role in conflict and cooperation. For ants, cuticular hydrocarbon (CHC) blends produced by non-nestmates elicit overt aggression. We describe a sensory sensillum on the antennae of the carpenter ant *Camponotus japonicus* that functions in nestmate discrimination. This sensillum is multiporous and responds only to non-nestmate CHC blends. This suggests a role for a peripheral recognition mechanism in detecting colony-specific chemical signals.

The struggle to maintain order in societies has led social animals, including human beings, to evolve and develop various means of communication. Ants have developed a sophisticated chemical communication system that enables them to reject non-nestmate conspecifics and to accept nestmates (*1*, *2*). Many behavioral experiments have suggested that their aggressive behavior against non-nestmates is evoked by contact chemosensory detection of differences between colony-specific chemical signals (*3*–*8*). Despite this well-defined behavior, the sensory mechanism for nestmate and non-nestmate discrimination has been unclear. It is thought that a "neural template" of nestmate recognition cues is formed that represents a constantly changing, experience-derived memory (*9*, *10*). By comparing the chemosensory discriminators or "labels" of encountered individuals with the "template" previously acquired, ants decide between acceptance or aggression (*11*, *12*). For such a decision rule by "template-label matching," several models have been proposed (*13*–*15*). They are constructed on a threshold-response hypothesis (*2*) in which some neural mechanism in the brain sets a threshold of similarity between template and label, thus regulating aggression.

For the carpenter ant, *C. japonicus*, cuticular CHC blends consist of at least 18 compounds in colony-specific ratios (Fig. 1A). To investigate how these organisms discern nestmate from non-nestmate signals, we developed a bioassay whereby a glass bead was used as a surrogate ant. The aggressive behavior of worker ants toward encountered non-nestmates was mimicked by a glass bead inoculated with either cuticle extract or a CHC fraction derived from the non-nestmate body surface (Fig. 1B). No aggression was elicited in response to extract from the nestmate body surface. There was a significant difference (t test; $P < 0.001$) in ant aggression against nestmate and non-nestmate compounds. About 40% of the ants became ag-

[1]Department of Applied Biology, Faculty of Textile Science, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto 606-8585, Japan. [2]Department of Earth System Science, Faculty of Science, Fukuoka University, Fukuoka 814-0180, Japan.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: mamiko@kit.ac.jp

Supporting Online Material

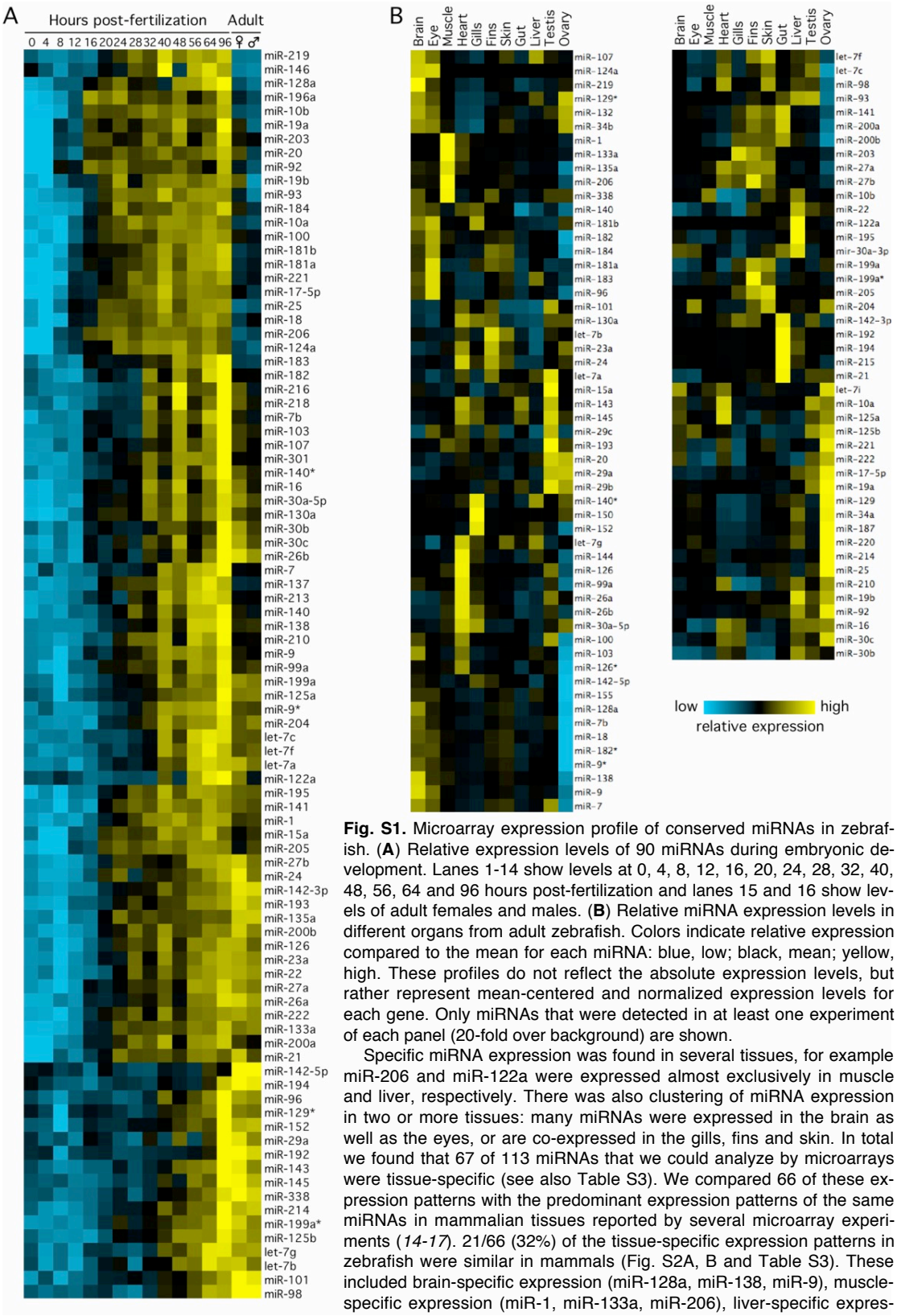


Fig. S1. Microarray expression profile of conserved miRNAs in zebrafish. (**A**) Relative expression levels of 90 miRNAs during embryonic development. Lanes 1-14 show levels at 0, 4, 8, 12, 16, 20, 24, 28, 32, 40, 48, 56, 64 and 96 hours post-fertilization and lanes 15 and 16 show levels of adult females and males. (**B**) Relative miRNA expression levels in different organs from adult zebrafish. Colors indicate relative expression compared to the mean for each miRNA: blue, low; black, mean; yellow, high. These profiles do not reflect the absolute expression levels, but rather represent mean-centered and normalized expression levels for each gene. Only miRNAs that were detected in at least one experiment of each panel (20-fold over background) are shown.

Specific miRNA expression was found in several tissues, for example miR-206 and miR-122a were expressed almost exclusively in muscle and liver, respectively. There was also clustering of miRNA expression in two or more tissues: many miRNAs were expressed in the brain as well as the eyes, or are co-expressed in the gills, fins and skin. In total we found that 67 of 113 miRNAs that we could analyze by microarrays were tissue-specific (see also Table S3). We compared 66 of these expression patterns with the predominant expression patterns of the same miRNAs in mammalian tissues reported by several microarray experiments (*14-17*). 21/66 (32%) of the tissue-specific expression patterns in zebrafish were similar in mammals (Fig. S2A, B and Table S3). These included brain-specific expression (miR-128a, miR-138, miR-9), muscle-specific expression (miR-1, miR-133a, miR-206), liver-specific expression (miR-122a) and gut-specific expression (miR-194). This overlap is an underestimate, because not all the same tissues have been analyzed by microarrays in zebrafish and mammals.

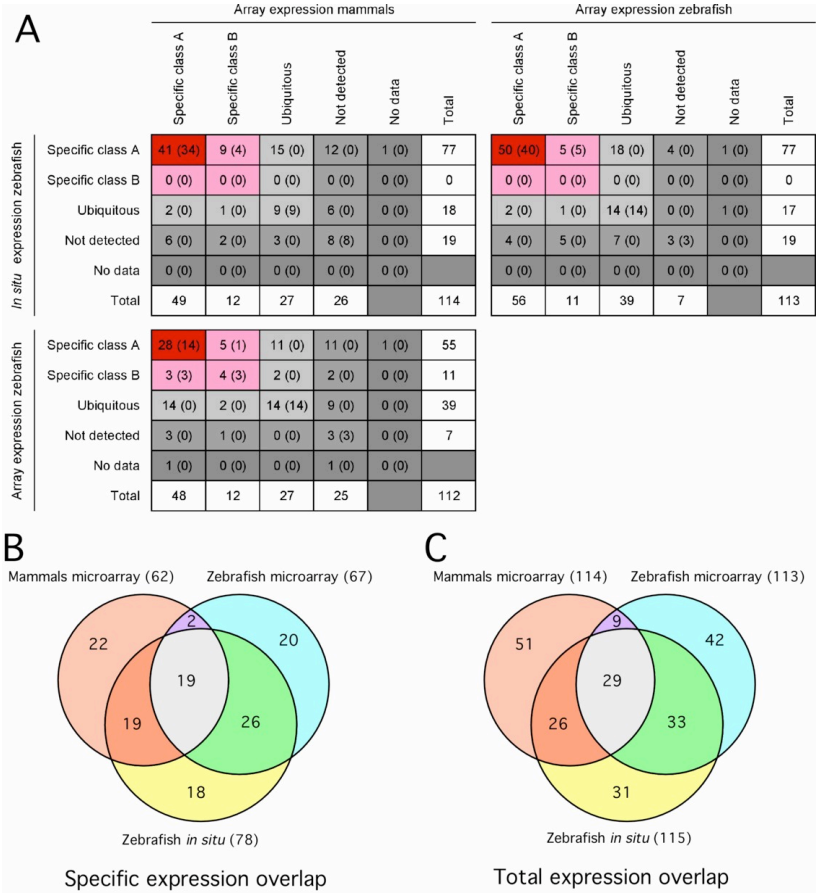


Fig. S2. Overlap between predominant miRNA expression patterns in mammals and zebrafish. (**A**) Pair-wise comparisons between miRNA expression data from mammals (microarray and literature data, see Table S3) and miRNA expression data from zebrafish (microarray and *in situ* hybridization data). To facilitate the comparisons, the miRNAs were first grouped according to their expression: specific class A (highly specific expression), red; specific class B (marginal specific- and/or low absolute expression), pink; ubiquitous expression, light gray; no detectable expression, gray; no data, dark gray. In the different data sets, a given miRNA can fall into other expression groups. Therefore we made square matrices, which reflect overlap between miRNAs from the different expression groups within each data set. The numbers of miRNAs belonging to each of these combinations of expression groups are shown. Next, the corresponding miRNA expression patterns were compared. The numbers of miRNAs with overlap in the expression patterns are denoted in brackets. miRNAs for which there was no data in either of the data sets were excluded from the comparison. miRNAs in specific class A and B were both classified as specific and therefore were both used for the calculations of the overlaps in specific expression patterns. (**B**) Three-way overlap between the specific (class A and B) miRNA expression of microarray data in mammals and zebrafish and *in situ* data in zebrafish. (**C**) Three-way overlap between the total miRNA expression patterns of microarray data in mammals and zebrafish and *in situ* data in zebrafish. Note that four miRNAs had overlap between *in situ* hybridizations and both microarrays in mammals and zebrafish but did not have overlap between the microarrays.

Forty-five miRNAs had overlap in specific expression patterns between the microarray data for zebrafish and the *in situ* hybridization data for zebrafish (A, B, see also Table S3). Thus, the *situ* hybridizations confirmed 67% of the 68 specific microarray expression patterns in zebrafish. The overlap between microarray data for mammals and our *in situ* data for zebrafish was similar. Of the 61 miRNAs with predominant microarray expressions in mammals, 38 (62%) had overlap with the *in situ* hybridization data (A, B, see also Table S3). Conversely, 45 (58%) and 38 (47%) of the 78 miRNAs with specific *in situ* patterns had overlap with the microarray expression patterns in zebrafish and mammals, respectively. Interestingly, 60 (77%) of the 78 miRNAs with specific *in situ* patterns had overlap with either the microarray data for zebrafish or for mammals (A, B, see also Table S3). In addition, overlap between non-specific patterns (ubiquitous and no detectable expression) was also observed (B, C, see also Table S3). Thus, in total a large fraction of the *in situ* patterns are consistent with microarray data for mammals or zebrafish.

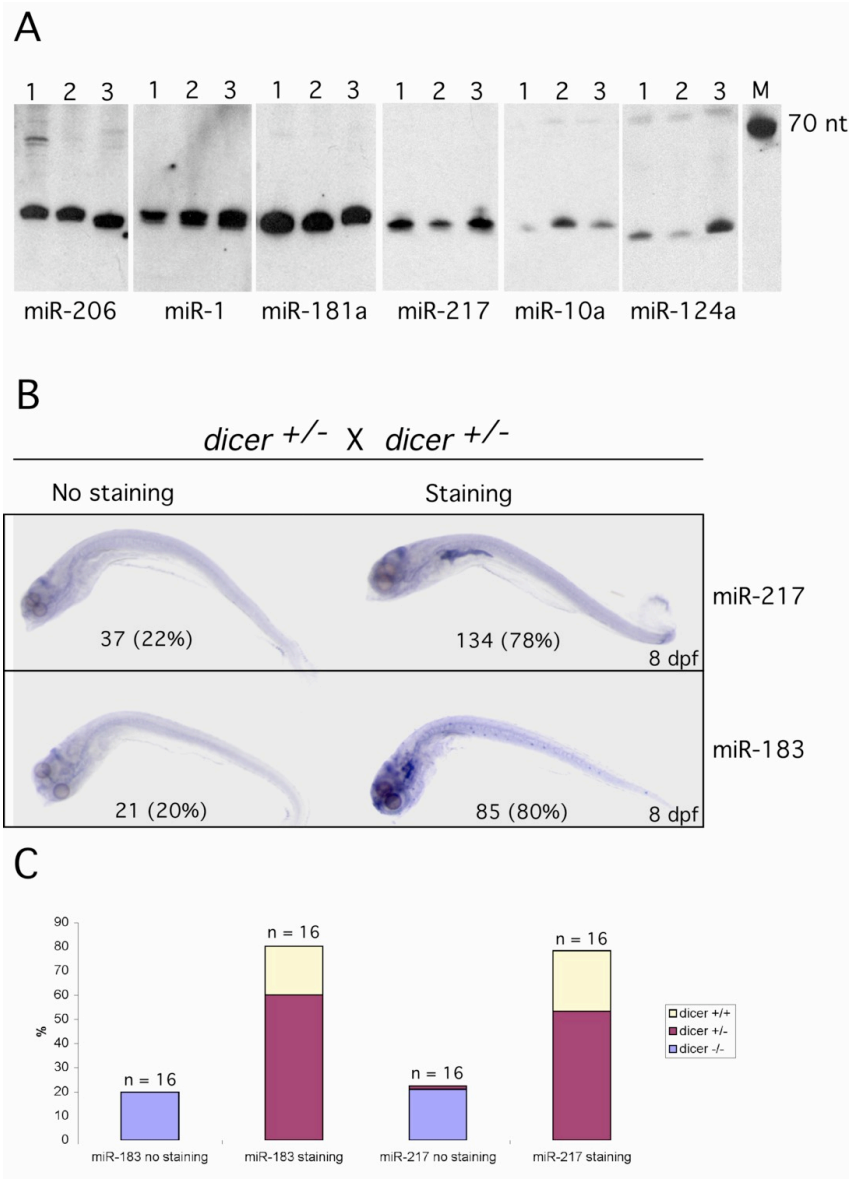
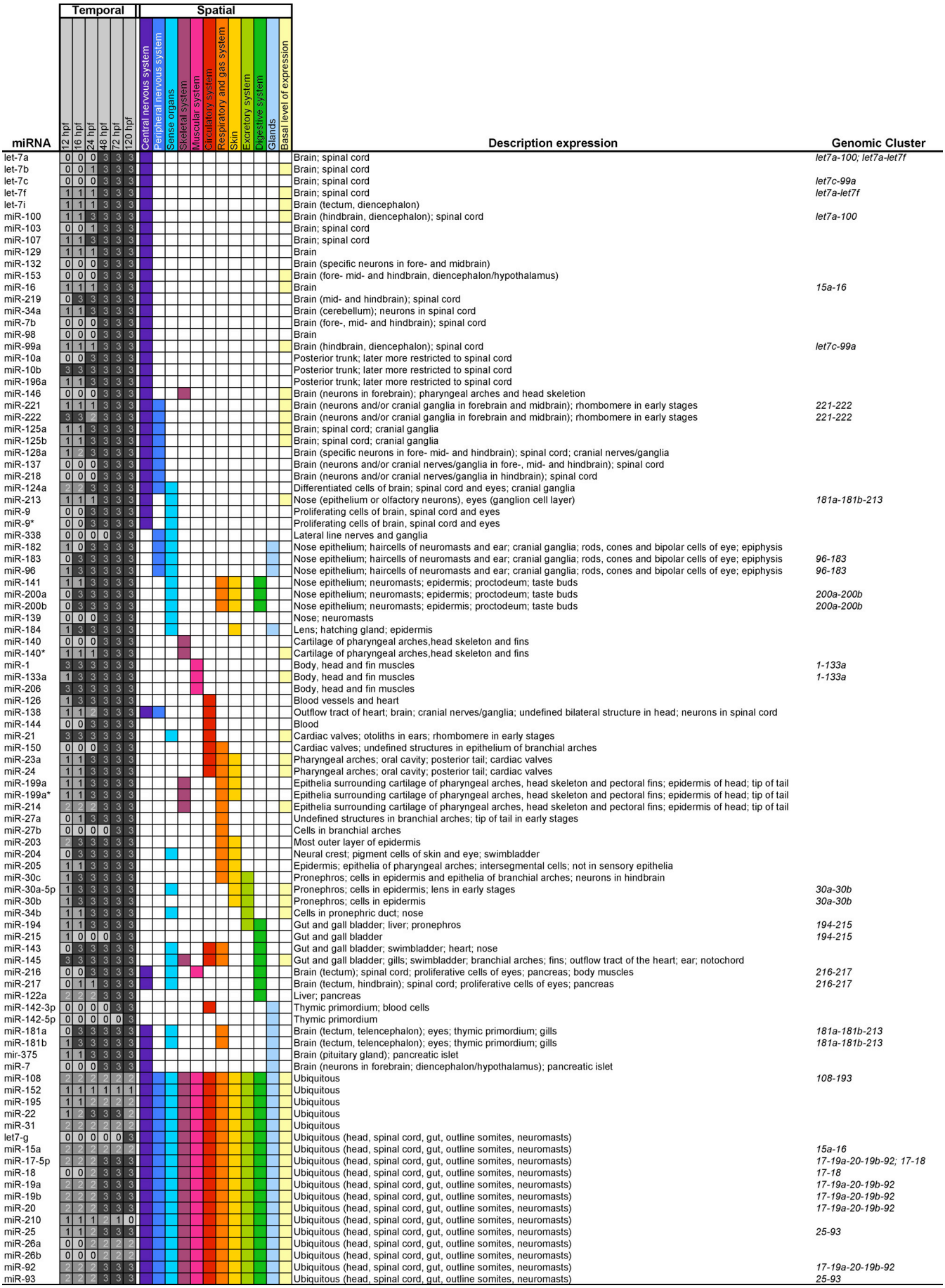


Fig. S3. miRNA detection in zebrafish using digoxigenin-labeled LNA-modified probes. (**A**). Northern blot analysis of six different miRNAs at 48 hpf (lane 1), 72 hpf (lane 2) and 5 days post-fertilization (lane 3). M, 70 nt marker. (**B**) Whole-mount *in situ* hybridizations on progeny of a *dicer*$^{hu715/+}$ incross at 8 days post-fertilization. Total number of embryos (5 and 8 days post-fertilization) that show staining or no staining for each miRNA is indicated. See Fig. S4 for description of the expression patterns. (**C**) Genotypes of embryos in (B). From each class 16 embryos were genotyped. In the 'no staining' classes 31 *dicer*$^{-/-}$ embryos and one *dicer*$^{+/-}$ embryo were detected. In the 'staining' classes only *dicer*$^{+/-}$ embryos or wild-type embryos were detected. This indicates that mature miRNAs cannot be detected in *dicer*$^{-/-}$ embryos and suggest that we specifically detect mature miRNAs in wild-type embryos.

Fig. S4. (Next page) Catalogue of miRNA expression patterns in zebrafish embryos. Expression patterns were determined by whole-mount *in situ* hybridizations using LNA-modified probes (*12*). To determine the miRNA expression patterns more precisely at cellular level, some embryos were sectioned. Expression was determined in embryos at 12, 16, 24, 48 and 72 h.p.f. and in five-day old larvae 120 h.p.f. Temporal expression, spatial expression in 11 different organ systems and basal level of expression in all organs is represented graphically. miRNAs that did not show expression were excluded from the figure. Temporal expression numbers indicate: 0, no expression; 1 background staining or weak ubiquitous expression; 2, ubiquitous expression; 3, specific expression in one or more organs systems. Some miRNAs in the latter class were expressed in many organs and therefore these miRNAs were finally annotated as ubiquitously expressed. Genomic miRNA clusters within 10,000 base-pairs are indicated. Note that some miRNAs are encoded by more than one gene, which can be in different clusters. Probe sequences and chromosomal positions of the miRNA genes are given in Tables S1 and S2. Pictures were mainly taken from embryos at 72 h.p.f. If additional structures were visible at earlier of later stages, pictures were also taken from these stages (Database S1).



miRNA	12 hpf	16 hpf	24 hpf	48 hpf	72 hpf	120 hpf	Description expression	Genomic Cluster
let-7a	0	0	0	3	3	3	Brain; spinal cord	*let7a-100; let7a-let7f*
let-7b	0	0	1	3	3	3	Brain; spinal cord	
let-7c	0	0	0	3	3	3	Brain; spinal cord	*let7c-99a*
let-7f	1	1	1	3	3	3	Brain; spinal cord	*let7a-let7f*
let-7i	1	1	1	3	3	3	Brain (tectum, diencephalon)	
miR-100	1	1	3	3	3	3	Brain (hindbrain, diencephalon); spinal cord	*let7a-100*
miR-103	0	0	1	3	3	3	Brain; spinal cord	
miR-107	1	1	3	3	3	3	Brain; spinal cord	
miR-129	1	1	1	3	3	3	Brain	
miR-132	0	0	0	3	3	3	Brain (specific neurons in fore- and midbrain)	
miR-153	0	0	0	3	3	3	Brain (fore- mid- and hindbrain, diencephalon/hypothalamus)	
miR-16	1	1	1	3	3	3	Brain	*15a-16*
miR-219	0	3	3	3	3	3	Brain (mid- and hindbrain); spinal cord	
miR-34a	1	1	3	3	3	3	Brain (cerebellum); neurons in spinal cord	
miR-7b	0	0	0	3	3	3	Brain (fore-, mid- and hindbrain); spinal cord	
miR-98	0	0	0	3	3	3	Brain	
miR-99a	1	1	1	3	3	3	Brain (hindbrain, diencephalon); spinal cord	*let7c-99a*
miR-10a	0	0	3	3	3	3	Posterior trunk; later more restricted to spinal cord	
miR-10b	3	3	3	3	3	3	Posterior trunk; later more restricted to spinal cord	
miR-196a	1	1	3	3	3	3	Posterior trunk; later more restricted to spinal cord	
miR-146	0	0	0	3	3	3	Brain (neurons in forebrain); pharyngeal arches and head skeletion	
miR-221	1	1	1	3	3	3	Brain (neurons and/or cranial ganglia in forebrain and midbrain); rhombomere in early stages	*221-222*
miR-222	3	3	2	3	3	3	Brain (neurons and/or cranial ganglia in forebrain and midbrain); rhombomere in early stages	*221-222*
miR-125a	1	1	3	3	3	3	Brain; spinal cord; cranial ganglia	
miR-125b	1	1	3	3	3	3	Brain; spinal cord; cranial ganglia	
miR-128a	1	2	3	3	3	3	Brain (specific neurons in fore- mid- and hindbrain); spinal cord; cranial nerves/ganglia	
miR-137	0	0	0	3	3	3	Brain (neurons and/or cranial nerves/ganglia in fore-, mid- and hindbrain); spinal cord	
miR-218	0	0	0	3	3	3	Brain (neurons and/or cranial nerves/ganglia in hindbrain); spinal cord	
miR-124a	2	2	3	3	3	3	Differentiated cells of brain; spinal cord and eyes; cranial ganglia	
miR-213	1	1	1	3	3	3	Nose (epithelium or olfactory neurons), eyes (ganglion cell layer)	*181a-181b-213*
miR-9	0	0	3	3	3	3	Proliferating cells of brain, spinal cord and eyes	
miR-9*	0	0	3	3	3	3	Proliferating cells of brain, spinal cord and eyes	
miR-338	0	0	0	0	3	3	Lateral line nerves and ganglia	
miR-182	1	0	3	3	3	3	Nose epithelium; haircells of neuromasts and ear; cranial ganglia; rods, cones and bipolar cells of eye; epiphysis	
miR-183	0	3	3	3	3	3	Nose epithelium; haircells of neuromasts and ear; cranial ganglia; rods, cones and bipolar cells of eye; epiphysis	*96-183*
miR-96	1	3	3	3	3	3	Nose epithelium; haircells of neuromasts and ear; cranial ganglia; rods, cones and bipolar cells of eye; epiphysis	*96-183*
miR-141	1	1	3	3	3	3	Nose epithelium; neuromasts; epidermis; proctodeum; taste buds	
miR-200a	0	3	3	3	3	3	Nose epithelium; neuromasts; epidermis; proctodeum; taste buds	*200a-200b*
miR-200b	0	3	3	3	3	3	Nose epithelium; neuromasts; epidermis; proctodeum; taste buds	*200a-200b*
miR-139	0	0	0	3	3	3	Nose; neuromasts	
miR-184	1	3	3	3	3	3	Lens; hatching gland; epidermis	
miR-140	0	0	0	3	3	3	Cartilage of pharyngeal arches,head skeleton and fins	
miR-140*	1	1	1	3	3	3	Cartilage of pharyngeal arches,head skeleton and fins	
miR-1	3	3	3	3	3	3	Body, head and fin muscles	*1-133a*
miR-133a	1	3	3	3	3	3	Body, head and fin muscles	*1-133a*
miR-206	3	3	3	3	3	3	Body, head and fin muscles	
miR-126	1	3	3	3	3	3	Blood vessels and heart	
miR-138	1	1	2	3	3	3	Outflow tract of heart; brain; cranial nerves/ganglia; undefined bilateral structure in head; neurons in spinal cord	
miR-144	0	0	3	3	3	3	Blood	
miR-21	3	3	3	3	3	3	Cardiac valves; otoliths in ears; rhombomere in early stages	
miR-150	0	0	0	3	3	3	Cardiac valves; undefined structures in epithelium of branchial arches	
miR-23a	1	1	3	3	3	3	Pharyngeal arches; oral cavity; posterior tail; cardiac valves	
miR-24	1	1	3	3	3	3	Pharyngeal arches; oral cavity; posterior tail; cardiac valves	
miR-199a	1	1	3	3	3	3	Epithelia surrounding cartilage of pharyngeal arches, head skeleton and pectoral fins; epidermis of head; tip of tail	
miR-199a*	1	1	3	3	3	3	Epithelia surrounding cartilage of pharyngeal arches, head skeleton and pectoral fins; epidermis of head; tip of tail	
miR-214	2	2	2	3	3	3	Epithelia surrounding cartilage of pharyngeal arches, head skeleton and pectoral fins; epidermis of head; tip of tail	
miR-27a	0	1	3	3	3	3	Undefined structures in branchial arches; tip of tail in early stages	
miR-27b	0	0	0	0	3	3	Cells in branchial arches	
miR-203	2	3	3	3	3	3	Most outer layer of epidermis	
miR-204	0	3	3	3	3	3	Neural crest; pigment cells of skin and eye; swimbladder	
miR-205	1	1	3	3	3	3	Epidermis; epithelia of pharyngeal arches; intersegmental cells; not in sensory epithelia	
miR-30c	1	3	3	3	3	3	Pronephros; cells in epidermis and epithelia of branchial arches; neurons in hindbrain	
miR-30a-5p	1	3	3	3	3	3	Pronephros; cells in epidermis; lens in early stages	*30a-30b*
miR-30b	1	3	3	3	3	3	Pronephros; cells in epidermis	*30a-30b*
miR-34b	1	1	3	3	3	3	Cells in pronephric duct; nose	
miR-194	1	1	3	3	3	3	Gut and gall bladder; liver; pronephros	*194-215*
miR-215	1	0	0	0	3	3	Gut and gall bladder	*194-215*
miR-143	0	3	3	3	3	3	Gut and gall bladder; swimbladder; heart; nose	
miR-145	3	3	3	3	3	3	Gut and gall bladder; gills; swimbladder; branchial arches; fins; outflow tract of the heart; ear; notochord	
miR-216	0	0	3	3	3	3	Brain (tectum); spinal cord; proliferative cells of eyes; pancreas; body muscles	*216-217*
miR-217	0	1	1	3	3	3	Brain (tectum, hindbrain); spinal cord; proliferative cells of eyes; pancreas	*216-217*
miR-122a	2	2	2	3	3	3	Liver; pancreas	
miR-142-3p	0	0	0	0	3	3	Thymic primordium; blood cells	
miR-142-5p	0	0	0	0	0	3	Thymic primordium	
miR-181a	0	3	3	3	3	3	Brain (tectum, telencephalon); eyes; thymic primordium; gills	*181a-181b-213*
miR-181b	1	3	3	3	3	3	Brain (tectum, telencephalon); eyes; thymic primordium; gills	*181a-181b-213*
mir-375	1	1	3	3	3	3	Brain (pituitary gland); pancreatic islet	
miR-7	0	0	0	3	3	3	Brain (neurons in forebrain; diencephalon/hypothalamus); pancreatic islet	
miR-108	2	2	2	2	2	2	Ubiquitous	*108-193*
miR-152	1	1	1	1	1	1	Ubiquitous	
miR-195	1	1	2	2	2	2	Ubiquitous	
miR-22	1	2	3	3	3	2	Ubiquitous	
miR-31	2	2	2	2	2	2	Ubiquitous	
let7-g	0	0	0	0	0	3	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)	
miR-15a	2	2	2	2	2	2	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)	*15a-16*
miR-17-5p	2	2	2	3	3	3	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)	*17-19a-20-19b-92; 17-18*
miR-18	0	0	2	3	3	3	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)	*17-18*
miR-19a	2	2	2	3	3	3	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)	*17-19a-20-19b-92*
miR-19b	2	2	2	3	3	3	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)	*17-19a-20-19b-92*
miR-20	2	2	2	3	3	3	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)	*17-19a-20-19b-92*
miR-210	1	1	1	2	1	0	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)	
miR-25	1	1	2	3	3	3	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)	*25-93*
miR-26a	0	0	2	2	2	2	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)	
miR-26b	0	0	0	2	2	2	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)	
miR-92	2	2	2	3	3	3	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)	*17-19a-20-19b-92*
miR-93	2	2	2	3	3	3	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)	*25-93*

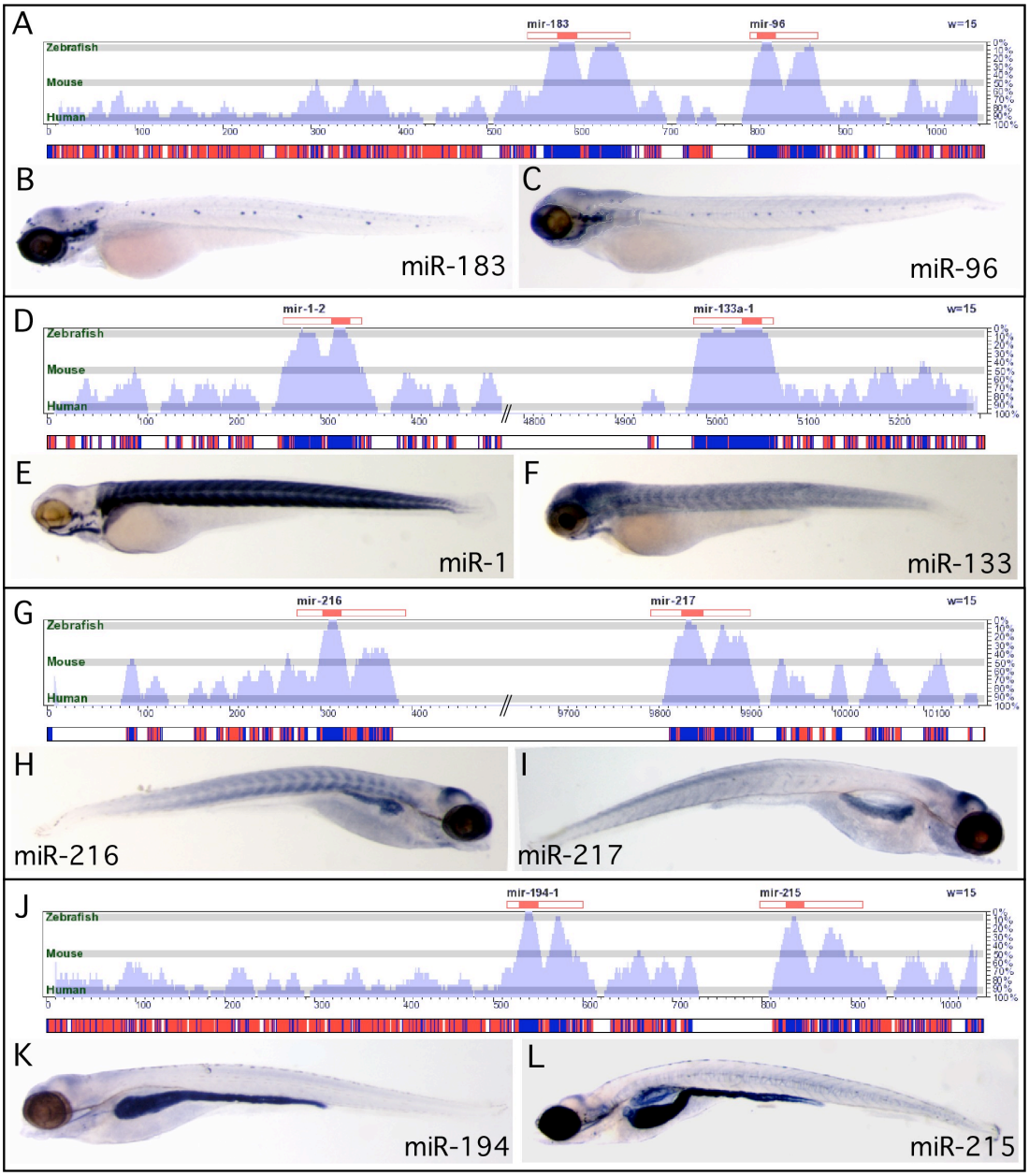


Fig. S5. Examples of miRNAs in genomic clusters. (**A, D, G, J**) VISTA-like plots of human/mouse/zebrafish comparisons of the genomic regions containing clustered miRNAs. Pre-miRNAs and mature miRNAs are indicated as open and filled boxes at the top of each figure, respectively. Divergence for each position is calculated in a 15-nucleotide window and plotted graphically. The bar below the plot indicates the nature of the variations observed in zebrafish regarding the consensus: blue, no variation; red, substitution; white, insertion or deletion. Distances between the miRNA pairs in (A, D, G, J) in zebrafish are 161, 2193, 235 and 174 nucleotides, respectively. (**B**, **C**, **E**, **F**) Lateral views of miRNA expression in embryos at 72 hpf. (**H**, **I**, **K**, **L**) Lateral views of miRNA expression in 5-day old larvae. Figures belonging to a cluster are boxed. See Fig. S3 for descriptions of the expression patterns. Many of these clusters consist of miRNAs from the same family, and in some cases we cannot exclude the possibility that the probes cross-hybridized. However, some clusters consist of unrelated miRNAs that nonetheless had similar expression patterns. We also observed differences between clustered miRNAs. For example, miR-216 and miR-217 were both expressed in the pancreas, but miR-216 was also expressed in the muscles. This indicates that miR-216 expression might be under additional control or that there are additional, yet unidentified, copies of *mir-216* in the genome that are differentially expressed.

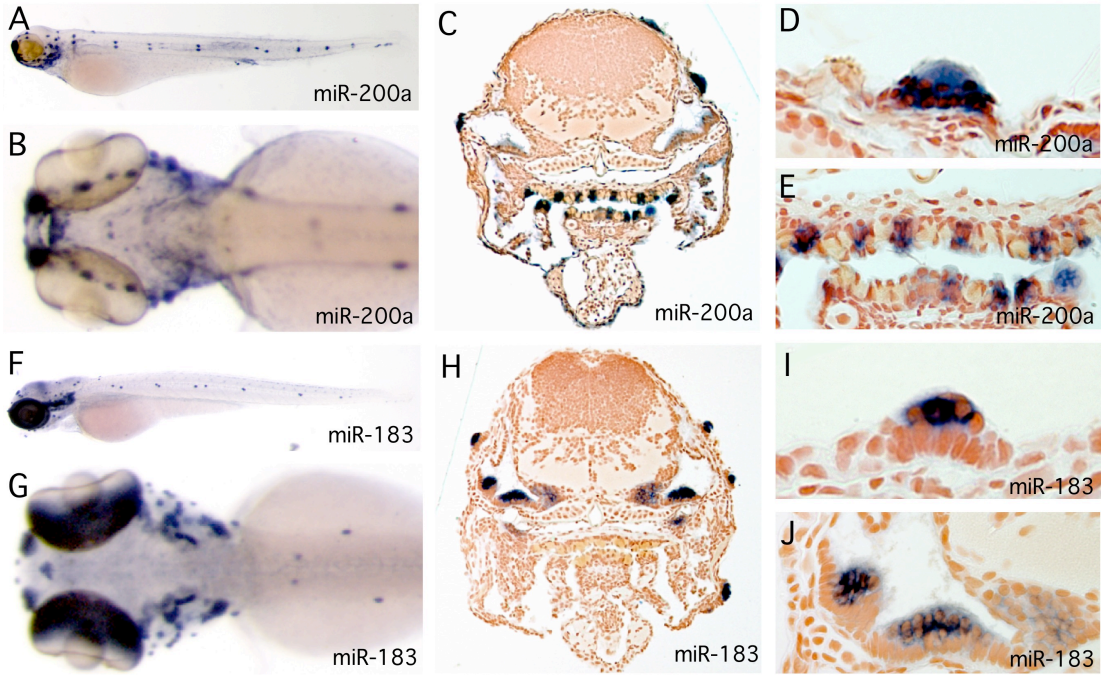


Fig. S6. miRNA expression in sensory epithelia. (**A**, **B**) In whole-mount embryos miR-200a was expressed in the nose, neuromasts of the lateral line, epithelium of the lips, mouth cavity and branchial arches. (**C**, **D**, **E**) On sections miR-200a was expressed in (C, D) hair- and supporting cells of the neuromasts and in (C, E) taste buds. (**F**, **G**) In whole-mount embryos miR-183 was expressed in the nose, neuromasts of the lateral line, ear, eye (rods, cones and bipolar cells), cranial ganglia and epiphysis. (**H**, **I**, **J**) On sections miR-200a was expressed in (H, I) hair cells of the neuromasts and in (H, J) sensory epithelia of the ear. (A, F) lateral views of embryos at 72 hpf. (B, G) Dorsal views of embryos shown in (A, F), respectively. (C, H) Cross sections through the head, at the position of the ear, of five-day old larvae. (D, E) Higher magnifications of image shown in (C). (I, J) Higher magnifications of image shown in (H). miR-200a apparently was expressed in sensory epithelial structures that can sense chemicals, and miR-183 was expressed in sensory epithelia that can sense light or vibrations. Thus, all of the sensory epithelial structures present in zebrafish had specific miRNAs expressed.

Materials and Methods

Animals

Zebrafish were kept under standard conditions (*1*). Embryos were staged according to (*2*). Wild-type zebrafish were used for northern blot analysis and microarray analysis. Homozygous *albino* embryos and larvae and progeny of a *dicer1*$^{hu715/+}$ (*3*) incross were used for the *in situ* hybridizations.

Identification of zebrafish miRNAs

Since the miRNA registry 5.0 (*4*) contained sequences for only 26 mature zebrafish miRNAs (corresponding to 30 precursor sequences), we designed a computational approach to identify additional zebrafish orthologs of known mammalian miRNAs. First, all human, mouse and rat precursor miRNA sequences from the miRNA registry 5.0 (*4*) were searched against the zebrafish genome assembly (Zv4) using the blastn program (*5*), and hits with lengths of at least 18 bp and identity of at least 16 bp were used as anchors for extracting relevant zebrafish genomic regions corresponding in length and position of the hit to the respective mammalian precursor miRNA sequence. Next, RNAfold software (*6*) was used to select only those regions that can form hairpin structures. As an additional filter, we used the Randfold program (*7*), which evaluates stability of the secondary structure relative to random sequences of the same nucleotide content. The regions with Randfold score of 0.005 or less were aligned with respective mammalian miRNA precursor sequences using CLUSTALW program, and average percentage identity was calculated for every alignment. In cases where several zebrafish regions were predicted for a given mammalian miRNA, the region with the highest alignment identity was considered as a true ortholog. Finally, redundancy in predictions (which occurred because of the use in the search of orthologous miRNAs from several species) was removed, and positions of mature miRNA sequences within predicted precursors were mapped by additional blast search. In this way we have identified 142 zebrafish miRNA regions corresponding to 126 unique mature miRNA sequences. 106 miRNAs that had no more than two mismatches were selected for further analyses. Of these, 23 had previously been identified in zebrafish (*8*), which validates this set. 74/106 (70%) matched perfectly and 100/106 (94%) had no more than one mismatch (Table S1 and Table S2). In total we investigated the expression of 115 miRNAs. These included the 106 miRNAs we identified from the zebrafish genome, three other previously identified zebrafish miRNAs, four other miRNAs that showed expression when assayed using microarrays, and two recently identified miRNAs (Table S1).

Microarray analysis

For the microarray analysis we used a microarray that was recently developed for the detection of mammalian miRNAs (*9*), containing probes for all the mammalian miRNAs currently known (miRNA registry 5.0). Since the microarray does not reliably discriminate between one or a few mismatches (*9*), we could use this array to study the expression of the zebrafish miRNAs. To perform the hybridizations, total RNA was isolated using trizol (Sigma) and size-selected for small RNAs in the range of mature miRNAs (18-26 nucleotides) by PAGE. Oligonucleotide microarrays were spotted on glass slides and hybridized as described previously (*9*). Briefly, oligonucleotides with complementary sequence to the miRNAs (miRNA registry 5.1(*4*)) were modified with a free amino group linked to the 5' termini through a 6-carbon spacer (IDT) and were printed onto amine-binding slides (CodeLink, Amersham Biosciences). Printing and hybridization were done according to the protocols from the manufacturer of the slides with the following modifications: the oligonucleotide concentration for printing was 20 μM in 150 mM sodium phosphate, pH 8.5. Printing was done on a MicroGrid TAS II arrayer (BioRobotics) at 50% humidity. For each microarray experiment 10 µg of total RNA was used. For the developmental time-course, RNA was labeled using a reverse transcription and amplification protocol as described previously (*9*). For all other samples the labeling method was modified to a direct labeling procedure as follows: after size-selection of the RNA, it was ligated to pCU-Cy3 using T4 RNA ligase. Ligation products were diluted 5 times in hybridization buffer and used directly for hybridization. Hybridization was done at 50°C for 10 hrs in 5X SSC, 0.1% SDS, 0.1 mg/ml salmon sperm DNA. The hybridized arrays were scanned using an arrayWoRx^{e} biochip reader (Applied Precision) and primary data were analyzed using the Digital Genome System suite (Molecularware). Data for the developmental time-course is log transformed, mean centered and normalized. Data for the different tissues are mean centered and normalized. Hierarchical clustering was performed using CLUSTER 3.0/TreeView software (*10*). Only miRNAs that were detected in at least one experiment (20-fold over background) are shown. Primary microarray data is deposited at the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under accession numbers GSE2625-GSE2628.

LNA-modified DNA oligonucleotides

LNA is a high-affinity RNA analogue with a bicyclic furanose unit locked in an RNA mimicking sugar conformation (*11*), which results in unprecedented hybridization affinity towards complementary single-stranded RNA molecules. This makes LNA-modified DNA probes ideally suited for RNA targeting. The sequences of the Locked Nucleic Acid (LNA)-modified DNA oligonucleotides (*12*) complementary to the mature miRNAs are listed in Table S1. The LNA probes were labeled with digoxigenin (DIG) using a DIG 3'-end labeling kit (Roche) and purified using Sephadex G25 MicroSpin columns (Amersham). For northern blot analysis and *in situ* hybridizations approximately 1-2 pmol of labeled probe was used.

Northern blot analysis

Total RNA was isolated from embryos at 48 hpf, 72 hpf and 5 day old embryos using trizol (Invitrogen). From each sample, 10 µg RNA was separated on 12.5% polyacrylamide gels and was blotted according to standard pro-

cedures. Blots were prehybridized in hybridization buffer (0.36M Na_2HPO_4, 0.14M NaH_2PO_4, 1mM EDTA, 7% SDS) for 30 min at 45°C and hybridized overnight in hybridization buffer containing 0.1 nM probe at 45°C. After stringent washes (2 times 30 min at 45°C in 2X SSC 0.1%SDS) blots were rinsed in washing buffer (0.1M maleic acid, 0.15M NaCl, 0.1% Tween, pH 7.5) and blocked in washing buffer containing 5% milk powder for 30 min at room temperature. Subsequently, blots were incubated with anti-DIG-AP antibody (Roche) in blocking buffer for 1 hour at room temperature, washed 3 times for 15 min in washing buffer and 2 times for 5 min with AP-buffer (0.1M Tris-HCl pH9.5, 50 mM $MgCl_2$, 0.1M NaCl, 0.1% Tween). Signal was detected using CDP-star chemiluminescent substrate (Roche).

Whole-mount in situ hybridizations

Whole-mount *in situ* hybridizations were performed essentially as described (*13*), with the following modifications: Hybridization, washing and incubation steps were done in 2.0 ml eppendorf tubes. All PBS and SSC solutions contained 0.1% Tween (PBST and SSCT). Embryos of 12, 16, 24, 48, 72 and 120 hpf were treated with proteinase K for 2, 5, 10, 30, 45 and 90 min, respectively. After proteinase K treatment and refixation with 4% paraformaldehyde, endogenous alkaline phosphatase activity was blocked by incubation of the embryos in 0.1 M ethanolamine and 2.5% acetic anhydride for 10 min, followed by extensive washing with PBST. Hybridizations were performed in 200 µl of hybridization mix. The temperature of hybridization and subsequent washing steps was adjusted to approximately 22°C below the predicted melting temperatures of the LNA-modified probes. Staining with NBT/BCIP was done overnight at 4°C. After staining, the embryos were fixed overnight in 4% paraformaldehyde. Next, embryos were dehydrated in an increasing methanol series and subsequently placed in a 2:1 mixture of benzyl benzoate and benzyl alcohol. Embryos were mounted on a hollow glass slide and covered with a coverslip.

Plastic sectioning

Embryos and larvae stained by whole-mount *in situ* hybridization were transferred from benzyl benzoate/benzyl alcohol to 100% methanol and incubated for 10 min. Specimens were washed twice with 100% ethanol for 10 min and incubated overnight in 100% Technovit 8100 infiltration solution (Kulzer) at 4ºC. Next, specimens were transferred to a mold and embedded overnight in Technovit 8100 embedding medium (Kulzer) deprived of air at 4ºC. Sections of 7 µm thickness were cut with a microtome (Reichert-Jung 2050), stretched on water and mounted on glass slides. Sections were dried overnight. Counterstaining was done by 0.05% neutral red for 12 sec, followed by extensive washing with water. Sections were preserved with Pertex and mounted under a coverslip.

Image acquisition

Embryos and larvae stained by whole-mount *in situ* hybridization were analyzed with Zeiss Axioplan and Leica MZFLIII microscopes and subsequently photographed with digital cameras. Sections were analyzed with a Nikon Eclipse E600 microscope and photographed with a digital camera (Nikon, DXM1200). A subset of images was adjusted for levels, brightness, contrast, hue and saturation with Adobe Photoshop 7.0 software to optimally visualize the expression patterns.

References

1. M. Westerfield, *The zebrafish book* (University of Oregon Press, 1993).
2. C. B. Kimmel, W. W. Ballard, S. R. Kimmel, B. Ullmann, T. F. Schilling, *Dev Dyn* **203**, 253-310 (1995).
3. E. Wienholds, M. J. Koudijs, F. J. van Eeden, E. Cuppen, R. H. Plasterk, *Nat Genet* **35**, 217-8 (2003).
4. S. Griffiths-Jones, *Nucleic Acids Res* **32 Database issue**, D109-11 (2004).
5. S. F. Altschul *et al.*, *Nucleic Acids Res* **25**, 3389-402 (1997).
6. I. L. Hofacker, *Nucleic Acids Res* **31**, 3429-31 (2003).
7. E. Bonnet, J. Wuyts, P. Rouze, Y. Van De Peer, *Bioinformatics* (2004).
8. L. P. Lim, M. E. Glasner, S. Yekta, C. B. Burge, D. P. Bartel, *Science* **299**, 1540 (2003).
9. E. A. Miska *et al.*, *Genome Biol* **5**, R68 (2004).
10. M. B. Eisen, P. T. Spellman, P. O. Brown, D. Botstein, *Proc Natl Acad Sci U S A* **95**, 14863-8 (1998).
11. A. A. Koshkin *et al.*, *Tetrahedron* **54**, 3607-3630 (1998).
12. A. Valoczi *et al.*, *Nucleic Acids Res* **32**, e175 (2004).
13. B. Thisse *et al.*, *Methods Cell Biol* **77**, 505-19 (2004).
14. S. Baskerville, D. P. Bartel, *Rna* **11**, 241-7 (2005).
15. O. Barad *et al.*, *Genome Res* **14**, 2486-94 (2004).
16. T. Babak, W. Zhang, Q. Morris, B. J. Blencowe, T. R. Hughes, *Rna* **10**, 1813-9 (2004).
17. J. M. Thomson, J. Parker, C. M. Perou, S. M. Hammond, *Nat Methods* **1**, 1-7 (2004).
18. M. N. Poy *et al.*, *Nature* **432**, 226-30 (2004).

Table S1. Zebrafish miRNAs and *in situ* probes

miRNA	miRNA sequence (5' to 3')	Conservation[ƒ]	Probe sequence (5' to 3')[†]	Target miRNAs[‡]
let-7a	UGAGGUAGUAGGUUGUAUAGUU	22/22/0	aactatacaacctactacctca	hsa-let-7a-2_zf10, rno-let-7a-2_zf506, hsa-let-7a-1_zf1
let-7b	UGAGGUAGUAGGUUGUGUGGUU	22/22/0	aaccacacaacctactacctca	mmu-let-7b_zf41
let-7c	UGAGGUAGUAGGUUGUAUGGUU	22/22/0	aaccatacaacctactacctca	rno-let-7c-1_zf519, hsa-let-7c_zf941
let-7f	UGAGGUAGUAGAUUGUAUAGUU	22/22/0	aactatacaatctactacctca	mmu-let-7f-1_zf76
let-7g	UGAGGUAGUUGUUUGUACAGU	21/21/1	actgtacaaacaactacctca	hsa-let-7g_zf1117
let-7i	UGAGGUAGUAGUUUGUGCU	19/19/0	agcacaaactactacctca	mmu-let-7i_zf98
miR-1	UGGAAUGUAAAGAAGUAUGUA	21/21/0	tacatacttctttacattcca	hsa-mir-1-2_zf1179, hsa-mir-1-1_zf1177
miR-100	AACCCGUAGAUCCGAACUUGUG	22/22/0	cacaagttcggatctacgggtt	mmu-mir-100_zf109
miR-101	UACAGUACUGUGAUAACUGAAG	22/22/0	cttcagttatcacagtactgta	rno-mir-101_zf566
miR-103	AGCAGCAUUGUACAGGGCUAUGA	23/23/0	tcatagccctgtacaatgctgct	hsa-mir-107_zf140
miR-107	AGCAGCAUUGUACAGGGCUAUCA	23/23/0	tgatagccctgtacaatgctgct	hsa-mir-103-2_zf101
miR-108	AUAAGGAUUUUUAGGGGCAUU	21/21/0	aatgcccctaaaaatccttat	hsa-mir-108_zf181
miR-10a	UACCCUGUAGAUCCGAAUUUGUG	23/23/0	cacaaattcggatctacagggta	dre-miR-10a
miR-10b	UACCCUGUAGAACCGAAUUUGU	22/22/0	acaaattcggttctacagggta	dre-miR-10b
miR-122a	UGGAGUGUGACAAUGGUGUUUGU	23/23/0	acaaacaccattgtcacactcca	hsa-mir-122a_zf223
miR-122a-1mm	-	-	acaaacaccatagtcacactcca	-
miR-122a-2mm	-	-	acaaacaccaaagtcacactcca	-
miR-124a	UUAAGGCACGCGGUGAAUGCCA	22/22/0	tggcattcaccgcgtgccttaa	hsa-mir-124a-2_zf296, rno-mir-124a-3_zf599, hsa-mir-124a-1_zf260, mmu-mir-124a-3_zf129
miR-124a-1mm	-	-	tggcattcaacgcgtgccttaa	-
miR-124a-2mm	-	-	tggcattcaaagcgtgccttaa	-
miR-125a	UCCCUGAGACCCUUAACCUGUG	23/23/1	cacaggttaagggtctcaggga	mmu-mir-125a_zf130, hsa-mir-125a_zf302
miR-125b	UCCCUGAGACCCUAACUUGUGA	22/22/0	tcacaagttagggtctcaggga	mmu-mir-125b-2_zf132, hsa-mir-125b-1_zf355
miR-126	UCGUACCGUGAGUAAUAAUGC	21/21/0	gcattattactcacggtacga	mmu-mir-126_zf135
miR-126*	CAUUAUUACUUUUGGUACGCG	21/21/0	cgcgtaccaaaagtaataatg	mmu-mir-126_zf135
miR-128a	UCACAGUGAACCGGUCUCUUUU	22/22/0	aaaagagaccggttcactgtga	mmu-mir-128a_zf137, hsa-mir-128a_zf412
miR-129	CUUUUUGCGGUCUGGGCUUGCU	21/21/0	agcaagcccagaccgcaaaaag	rno-mir-129-1_zf619, mmu-mir-129-2_zf144
miR-129*	AAGCCCUUACCCCAAAAAGCAU	22/22/0	atgctttttggggtaagggctt	mmu-mir-129-2_zf144
miR-130a	CAGUGCAAUGUUAAAAGGGC	20/20/0	gccctttaacattgcactg	hsa-mir-130a_zf492
miR-132	UAACAGUCUACAGCCAUGGUCG	22/22/0	cgaccatggctgtagactgtta	mmu-mir-132_zf150
miR-133a	UUGGUCCCCUUCAACCAGCUGU	22/22/0	acagctggttgaaggggaccaa	rno-mir-133a_zf636, mmu-mir-133a-2_zf152
miR-135a	UAUGGCUUUUUAUUCCUAUGUGA	23/23/0	tcacataggaataaaaagccata	hsa-mir-135b_zf711
miR-137	UAUUGCUUAAGAAUACGCGUAG	22/22/0	ctacgcgtattcttaagcaata	hsa-mir-137_zf754, mmu-mir-137_zf170
miR-138[§]	AGCUGGUGUUGUGAAUC	-	gattcacaacaccagct	-

Table S1. Continued

miRNA	miRNA sequence (5' to 3')	Conservation	Probe sequence (5' to 3')	Target miRNAs
miR-139	UCUACAGUGCAUGUGUCU	18/18/1	agacacatgcactgtaga	rno-mir-139_zf654
miR-140	AGUGGUUUUACCCUAUGGUAG	21/21/0	ctaccatagggtaaaaccact	mmu-mir-140_zf177
miR-140*	UACCACAGGGUAGAACCACGGAC	23/24/0	gtccgtggttctaccctgtggta	mmu-mir-140_zf177
miR-141	AACACUGUCUGGUAACGAUGC	-	gcatcgttaccagacagtgtt	mmu-mir-141_zf178
miR-142-3p	UGUAGUGUUUCCUACUUUAUGGA	23/23/0	tccataaagtaggaaacactaca	hsa-miR-142-3p
miR-142-5p	CAUAAAGUAGAAAGCACUAC	20/20/0	gtagtgctttctactttatg	hsa-miR-142-5p
miR-143§	UGAGAUGAAGCACUGUAGCUCG	-	cgagctacagtgcttcatctca	-
miR-144	UACAGUAUAGAUGAUGUACUA	21/22/0	tagtacatcatctatactgta	mmu-mir-144_zf180
miR-145	GUCCAGUUUUCCCAGGAAUCCCUU	24/24/0	aagggattcctgggaaaactggac	mmu-mir-145_zf186
miR-146	UGAGAACUGAAUUCCAAGGGU	21/22/1	acccttggaattcagttctca	hsa-mir-146_zf946
miR-148a	UCAGUGCAUUACAGAACUUUGU	22/22/1	acaaagttctgtaatgcactga	mmu-mir-148a_zf187
miR-150	UCUCCCAAUCCUUGUACCAGUG	22/22/1	cactggtacaaggattgggaga	rno-mir-150_zf663
miR-152	UCAGUGCAUGACAGAACUU	19/21/0	aagttctgtcatgcactga	rno-mir-152_zf664
miR-153	UUGCAUAGUCACAAAAGUGA	20/20/0	tcacttttgtgactatgcaa	hsa-mir-153-2_zf956
miR-155	UUAAUGCUAAUCGUGAUAGGGG	22/22/0	cccctatcacgattagcattaa	mmu-mir-155_zf197
miR-15a	UAGCAGCACAGAAUGGUUUGUG	22/22/1	cacaaaccattctgtgctgcta	mmu-mir-15a_zf200
miR-16	UAGCAGCACGUAAAUAUUGG	20/22/0	ccaatatttacgtgctgcta	hsa-mir-16-1_zf965
miR-17-5p	CAAAGUGCUUACAGUGCAGGUAGU	24/24/0	actacctgcactgtaagcactttg	hsa-mir-17_zf968
miR-18	UAAGGUGCAUUUAGUGCAGAUA	22/22/1	tatctgcactaaatgcacctta	hsa-mir-18_zf972
miR-181a	AACAUUCAACGCUGUCGGUGAGU	23/23/0	actcaccgacagcgttgaatgtt	mmu-mir-181a_zf210, mmu-mir-213_zf289, mmu-mir-181c_zf233
miR-181b	AACAUUCAUUGCUGUCGGUGGGUU	24/24/0	aacccaccgacagcaatgaatgtt	dre-miR-181b
miR-182	UUUGGCAAUGGUAGAACUCACA	-	tgtgagttctaccattgccaaa	dre-miR-182
miR-182*	UGGUUCUAGACUUGCCAACUA	-	tagttggcaagtctagaacca	dre-miR-182*
miR-183	UAUGGCACUGGUAGAAUUCACUG	23/23/0	cagtgaattctaccagtgccata	dre-miR-183
miR-184	UGGACGGAGAACUGAUAAGGG	21/22/0	cccttatcagttctccgtcca	hsa-mir-184_zf990
miR-187	UCGUGUCUUGUGUUGCAGCC	20/21/0	ggctgcaacacaagacacga	dre-miR-187 (one mismatch at end)
miR-189	GUGCCUACUGAGCUGAUAACAGU	23/23/1	actgttatcagctcagtaggcac	hsa-mir-24-2_zf1111
miR-190	UGAUAUGUUUGAUAUAUUAGGU	22/22/0	acctaatatatcaaacatatca	hsa-mir-190_zf992
miR-192	UGACCUAUGAAUUGACAGCC	20/21/0	ggctgtcaattcataggtca	dre-miR-192 (one mismatch at end), hsa-miR-192
miR-193	AACUGGCCUACAAAGUCCCAG	21/21/0	ctgggactttgtaggccagtt	rno-mir-193_zf697, hsa-mir-193_zf994
miR-194	UGUAACAGCAACUCCAUGUGGA	22/22/0	tccacatggagttgctgttaca	mmu-mir-194-1_zf238
miR-195§	UAGCAGCACAGAAAUAUUGGC	-	gccaatatttctgtgctgcta	-
miR-196a	UAGGUAGUUUCAUGUUGUUGGG	21/21/0	cccaacaacatgaaactaccta	dre-miR-196a

Table S1. Continued

miRNA	miRNA sequence (5' to 3')	Conservation	Probe sequence (5' to 3')	Target miRNAs
miR-199a	CCCAGUGUUCAGACUACCUGUUC	23/23/0	gaacaggtagtctgaacactggg	dre-miR-199a
miR-199a*	UACAGUAGUCUGCACAUUGGUU	22/22/0	aaccaatgtgcagactactgta	hsa-mir-199b_zf1017
miR-19a	UGUGCAAAUCUAUGCAAAACUGA	23/23/0	tcagttttgcatagatttgcaca	hsa-mir-19a_zf1019
miR-19b	UGUGCAAAUCCAUGCAAAACUGA	23/23/0	tcagttttgcatggatttgcaca	mmu-mir-19b-2_zf271
miR-20	UAAAGUGCUUAUAGUGCAGGUAG	22/22/0	ctacctgcactataagcacttta	hsa-mir-20_zf1037
miR-200a	UAACACUGUCUGGUAACGAUGU	22/22/0	acatcgttaccagacagtgtta	rno-mir-200a_zf731
miR-200b	UCUAAUACUGCCUGGUAAUGAUG	23/24/0	catcattaccaggcagtattaga	rno-mir-200b_zf736
miR-203	GUGAAAUGUUUAGGACCACU	20/22/0	agtggtcctaaacatttcac	dre-miR-203 (two mismatches at end), mmu-mir-203_zf280
miR-204	UUCCCUUUGUCAUCCUAUGCCUG	22/22/0	caggcataggatgacaaagggaa	dre-miR-204
miR-205	UCCUUCAUUCCACCGGAGUCUG	22/22/0	cagactccggtggaatgaagga	dre-miR-205
miR-206	UGGAAUGUAAGGAAGUGUGUGG	22/22/0	ccacacacttccttacattcca	hsa-mir-206_zf1052
miR-206-1mm	-	-	ccacacactaccttacattcca	-
miR-206-2mm	-	-	ccacacactatcttacattcca	-
miR-21	UAGCUUAUCAGACUGGUGUUG	21/22/1	caacaccagtctgataagcta	hsa-mir-21_zf1054
miR-210	CUGUGCGUGUGACAGCGGCU	20/21/0	agccgctgtcacacgcacag	dre-miR-210 (one mismatch at end), hsa-mir-210_zf1055
miR-213	ACCAUCGACCGUUGACUGUACC	22/22/1	ggtacagtcaacggtcgatggt	dre-miR-213 (two mismatches), mmu-mir-181c_zf233
miR-214	ACAGCAGGCACAGACAGGCAG	21/21/0	ctgcctgtctgtgcctgctgt	dre-miR-214
miR-215	AUGACCUAUGAAUUGACAG	19/21/0	ctgtcaattcataggtcat	mmu-mir-192_zf237
miR-216	UAAUCUCAGCUGGCAACUGUG	21/21/0	cacagttgccagctgagatta	dre-miR-216
miR-217	UACUGCAUCAGGAACUGAUUGGAU	24/24/0	atccaatcagttcctgatgcagta	dre-miR-217
miR-218	UUGUGCUUGAUCUAACCAUGU	21/21/0	acatggttagatcaagcacaa	hsa-mir-218-1_zf1073
miR-219	UGAUUGUCCAAACGCAAUUCU	21/21/0	agaattgcgtttggacaatca	dre-miR-219
miR-22	AAGCUGCCAGCUGAAGAACUGU	22/22/1	acagttcttcagctggcagctt	hsa-mir-22_zf1092
miR-220	CCACAACCGUAUCGGACACUU	-	aagtgtccgatacggttgtgg	dre-miR-220
miR-221	AGCUACAUUGUCUGCUGGGUUUC	23/23/0	gaaacccagcagacaatgtagct	dre-miR-221
miR-222	AGCUACAUCUGGCUACUGGGUCUC	24/24/0	gagacccagtagccagatgtagct	dre-miR-222
miR-223	UGUCAGUUUGUCAAAUACCCC	21/21/0	ggggtatttgacaaactgaca	dre-miR-223
miR-23a	AUCACAUUGCCAGGGAUUUCC	21/21/0	ggaaatccctggcaatgtgat	hsa-mir-23a_zf1099, rno-mir-23a_zf784
miR-24	UGGCUCAGUUCAGCAGGAACAG	22/22/0	ctgttcctgctgaactgagcca	hsa-mir-24-1_zf110, rno-mir-24-2_zf795
miR-25	CAUUGCACUUGUCUCGGUCUGA	22/22/0	tcagaccgagacaagtgcaatg	mmu-mir-25_zf346
miR-26a	UUCAAGUAAUCCAGGAUAGGCU	22/22/0	agcctatcctggattacttgaa	hsa-mir-26a-1_zf1122, hsa-mir-26b_zf1133, rno-mir-26b_zf807
miR-26b	UUCAAGUAAUCCAGGAUAGGUU	22/22/1	aacctatcctggattacttgaa	hsa-mir-26a-2_zf1128
miR-27a	UUCACAGUGGCUAAGUUCCGC	21/22/0	gcggaacttagccactgtgaa	mmu-mir-27a_zf370

Table S1. Continued

miRNA	miRNA sequence (5' to 3')	Conservation	Probe sequence (5' to 3')	Target miRNAs
miR-27b	UUCACAGUGGCUAAGUUCUG	20/20/0	cagaacttagccactgtgaa	mmu-mir-27b_zf374
miR-29a	CUAGCACCAUUUGAAAUCGGUU	22/22/1	aaccgatttcaaatggtgctag	hsa-miR-29a
miR-29b	UAGCACCAUUUGAAAUCAGUGU	20/20/0	acactgatttcaaatggtgcta	hsa-mir-29b-1_zf1145
miR-29c	UAGCACCAUUUGAAAUCGGUUA	22/22/0	taaccgatttcaaatggtgcta	mmu-mir-29a_zf379
miR-301	CAGUGCAAUAGUAUUGUCAAAGC	23/23/0	gctttgacaatactattgcactg	hsa-mir-301_zf1175
mir-30a-3p	CUUUCAGUUGGAUGUUUGCUGU	-	acagcaaacatccaactgaaag	mmu-mir-30a_zf414
miR-30a-5p	UGUAAACAUCCCCGACUGGAAGC	23/23/1	gcttccagtcggggatgtttaca	mmu-mir-30a_zf414
miR-30b	UGUAAACAUCCUACACUCAGC	21/21/0	gctgagtgtaggatgtttaca	rno-mir-30c-2_zf875
miR-30c	UGUAAACAUCCUACACUCUCAGC	23/23/0	gctgagagtgtaggatgtttaca	hsa-mir-30c-2_zf1194
miR-31	GGCAAGAUGUUGGCAUAGCUG	21/21/1	cagctatgccaacatcttgcc	hsa-mir-31_zf1206
miR-338	UCCAGCAUCAGUGAUUUUGUUG	22/23/0	caacaaaatcactgatgctgga	mmu-mir-338_zf433
miR-34a	UGGCAGUGUCUUAGCUGGUUGUU	22/22/0	aacaaccagctaagacactgcca	dre-miR-34a
miR-34b	UAGGCAGUGUUGUUAGCUGAUUG	22/22/1	caatcagctaacaacactgccta	hsa-mir-34b_zf1211
mir-375¶	UUUGUUCGUUCGGCUCGCGUUA	-	taacgcgagccgaacgaacaaa	-
miR-7	UGGAAGACUAGUGAUUUUGUU	21/21/0	aacaaaatcactagtcttcca	dre-miR-7, rno-mir-7b_zf906, hsa-mir-7-3_zf1219
miR-7b	UGGAAGACUUGUGAUUUUGUU	21/21/0	aacaaaatcacaagtcttcca	dre-miR-7b
miR-9	UCUUUGGUUAUCUAGCUGUAUGA	23/23/0	tcatacagctagataaccaaaga	mmu-mir-9-1_zf455, hsa-mir-9-3_zf1233, mmu-mir-9-2_zf461
miR-9*	UAAAGCUAGAUAACCGAAAGU	21/21/0	actttcggttatctagcttta	mmu-mir-9-1_zf455, hsa-mir-9-3_zf1233, mmu-mir-9-2_zf461
miR-92	UAUUGCACUUGUCCCGGCCUGU	22/22/0	acaggccgggacaagtgcaata	hsa-mir-92-1_zf1239
miR-93	AAAGUGCUGUUUGUGCAGGUAG	22/22/1	ctacctgcacaaacagcacttt	hsa-mir-93_zf1243
miR-96	UUUGGCACUAGCACAUUUUUGCU	22/22/0	agcaaaaatgtgctagtgccaaa	hsa-mir-96_zf1244
miR-98	UGAGGUAGUAAGUUGUGUUGUU	22/22/1	aacaacacaacttactacctca	rno-mir-98_zf931
miR-99a	AACCCGUAGAUCCGAUCUUGUG	22/22/0	cacaagatcggatctacgggtt	hsa-mir-99a_zf25

∫ length zebrafish/length mammals/internal mismatches. Mismatches at the ends of the zebrafish miRNAs (compared to mammalian miRNAs) are trimmed.

† The LNA-modified oligonucleotides complementary to the mature miRNA sequences are available as miRCURY detection probes at www.exiqon.com.

‡ Best matching hit in the zebrafish genome, see Table S2 for chromosomal positions.

§ Not detected by algorithm used in this study, but selected because it showed expression on microarrays.

¶ Not in miRNA registry 5.0, derived from (*18*).

Table S2. Chromosomal positions of zebrafish microRNAs[§]

miRNA	Target miRNA	Chromosome / Scaffold	Position start	Position end	Strand	Precursor sequence[¶]
miR-16	hsa-mir-16-1_zf965	1	48875524	48875613	-1	ggctgcctggctgTAGCAGCACGTAAATATTGGagtcaaagcacttgcgaatcctccagtattgaccgtgctgctggagttaggcgggcc
miR-15a	mmu-mir-15a_zf200	1	48875741	48875825	-1	cctgtcggtactgTAGCAGCACAGAATGGTTTGTGagttataacgggggtgcaggccgtactgtgctgcggcaacaacgacagga
miR-137	hsa-mir-137_zf754	2	27119586	27119688	-1	ttttctgaggctctcttcggtgacgggtattcttgggtggataatacggctctcgttgtTATTGCTTAAGAATACGCGTAGttgaggagagtcattgtcgact
miR-189	hsa-mir-24-2_zf1111	2	30198890	30198963	1	ttcaacctcctGTGCCTACTGAGCTGATAACAGTtgatgtatttagcactggctcagttcagcaggaaccggag
miR-152	rno-mir-152_zf664	3	14774990	14775075	-1	tgttcacctggctcaagttctgtgatacactcagactttgaatcagtggtagTCAGTGCATGACAGAACTTtggcccggacggacc
miR-150	rno-mir-150_zf663	3	23984075	23984160	-1	cagtccatctccgtcTCTCCCAATCCTTGTACCAGTGtctgatttacagatgacgctggacggggtttgggggggggctgagggagg
miR-199a	mmu-mir-199b_zf259	3	43600711	43600821	1	ccacctcttgcctccccctcgcctgCCCAGTGTTCAGACTACCTGTTCatcatgctgcagctgaacagtagtccgcacattggttaggctgggctgggacacacacacccg
miR-29b	hsa-mir-29b-1_zf1145	4	7154505	7154586	1	cctccagatgctggtttcacatggtggtttagatgtgttctaccaaagtcTAGCACCATTTGAAATCAGTGTtcttggggag
miR-29c	mmu-mir-29a_zf379	4	7155928	7156016	1	ccccaccaaacgatgactgatttccttggtgcttagagtcccatctgtcatcTAGCACCATTTGAAATCGGTTAtaatgactggggat
miR-96	hsa-mir-96_zf1244	4	14390900	14390978	-1	ttgcctgtTTTGGCACTAGCACATTTTTGCTtttttatatataccttgagcaattatgtgtagtgccaatatgggacaa
miR-183	dre-mir-183	4	14391043	14391153	-1	tgaagctgacagactcctgttctgtgTATGGCACTGGTAGAATTCACTgtgaaagcacactatcagtgaattaccaaagggccataaacagagcagagaaagaaccacgtg
miR-129	rno-mir-129-1_zf619	4	16487503	16487576	-1	caggtCTTTTTGCGGTCTGGGCTTGCTgttccttgaaccagtagccaggaagcccttaccccaaaaagtatctg
miR-181b	dre-mir-181b	4	21006433	21006522	-1	ctaatgactgcaataAACATTCATTGCTGTCGGTGGGTTtctaatagacacaactcactgatcaatgaatgcaaactgcggtgcaagaac
miR-181a	mmu-mir-181c_zf233	4	21006581	21006670	-1	ctggtcccttggtgAACATTCAACGCTGTCGGTGAGTtttgcgcttctgtaacaaACCATCGACCGTTGACTGTACCctgagggtggcca
miR-213	mmu-mir-181c_zf233	4	21006581	21006670	-1	ctggtcccttggtgAACATTCAACGCTGTCGGTGAGTtttgcgcttctgtaacaaACCATCGACCGTTGACTGTACCctgagggtggcca
miR-9	mmu-mir-9-2_zf461	5	21811624	21811696	1	gttgttaTCTTTGGTTATCTAGCTGTATGAgtattttgcacttcaTAAAGCTAGATAACCGAAAGTaaaaact
miR-9*	mmu-mir-9-2_zf461	5	21811624	21811696	1	gttgttaTCTTTGGTTATCTAGCTGTATGAgtattttgcacttcaTAAAGCTAGATAACCGAAAGTaaaaact
let-7a	rno-let-7a-2_zf506	5	24847663	24847759	-1	gagcgatgtctcgggaTGAGGTAGTAGGTTGTATAGTTtagagttacaacacgggagataactgtacagcctcctagctttcctcgagcagacgccc
miR-100	mmu-mir-100_zf109	5	24847984	24848064	-1	ccagctgccacaAACCCGTAGATCCGAACTTGTGgtgtctctgtgcacaagctcgtgtctatagtatgtgtctgcgaggt
miR-193	rno-mir-193_zf697	5	26889112	26889198	-1	atgtgttagaggttgggtctttgcgggcaaggtgagtagttaaatttactctcAACTGGCCTACAAAGTCCCAGtttctggctcatg
miR-144	mmu-mir-144_zf180	5	47178990	47179057	1	agacaggatatcatcgtatactgtaagttcattattgagacacTACAGTATAGATGATGTACTAtcca
miR-34b	hsa-mir-34b_zf1211	5	60620008	60620092	-1	ggggttggtctgTAGGCAGTGTTGTTAGCTGATTGtttcatatgaactataatcactaaccatactgccaacacaacaacctaca
miR-141	mmu-mir-141_zf178	6	24110434	24110506	-1	tctagggtacatcttacctgacagtgcttggctgttcactgatgttctAACACTGTCTGGTAACGATGCactc
let-7g	hsa-let-7g_zf1117	6	24976547	24976632	1	gggtTGAGGTAGTTGTTTGTACAGTttttagggtctgttattctgccctgttaaggagctaactgtacagactactgccttgccca
miR-204	dre-mir-204	7	22355610	22355717	-1	agtgaccagtttgtgacctcctgggtTTCCCTTTGTCATCCTATGCCTGcagttcctgatgaggctgggacagcaaagggaggttcagatgtcgacctgtactacagt
miR-206	hsa-mir-206_zf1052	7	33261492	33261578	-1	tgcttattgagacgacatacttccttatatccccatattcagaattaatctaTGGAATGTAAGGAAGTGTGTGGcttcagtgagatc
miR-153	hsa-mir-153-2_zf956	7	36134976	36135063	-1	agcggttgccagtgtcatttttgtgatgttgcagctagttatatgagcccagTTGCATAGTCACAAAAGTGAtcattggaaactgtat
miR-146	hsa-mir-146_zf946	7	49635442	49635541	1	ttacactgagctcttggcttTGAGAACTGAATTCCAAGGGTgtctgctttatattcagcccacggagttcagttcttaagtttggatgctcacggccgtc
miR-93	hsa-mir-93_zf1243	7	59085414	59085494	1	tgtgtgtgttaAAAGTGCTGTTTGTGCAGGTAGtgtgtttcctctactgtaggagcagcacttcacaacacacacactgct
miR-25	mmu-mir-25_zf346	7	59085679	59085763	1	agccggcgctgagaggcggagacttgggcagctgccgtcattcccagaaggCATTGCACTTGTCTCGGTCTGAcagtggcggcac
miR-126	mmu-mir-126_zf135	8	5287947	5288021	-1	tcacagtcCATTATTACTTTTGGTACGCGctaggccagactcaaacTCGTACCGTGAGTAATAATGCactgtggc
miR-126*	mmu-mir-126_zf135	8	5287947	5288021	-1	tcacagtcCATTATTACTTTTGGTACGCGctaggccagactcaaacTCGTACCGTGAGTAATAATGCactgtggc
miR-27b	mmu-mir-27b_zf374	8	24138295	24138365	1	aggtgcagagcttagctgattggtgaacagtgattgaactctttgTTCACAGTGGCTAAGTTCTGcatctg
miR-10b	dre-mir-10b	9	546919	547028	1	gtgaatatatgccgtcgtctatataTACCCTGTAGAACCGAATTTGTgtgaaaaaataacattcacagattcgattctaggggagtatatggtcgatgcaataacttcga
miR-221	dre-mir-221	9	29778587	29778682	-1	gtcttttgtgctgtcgtgaacctggcatacaatgtagatttctgtgtggtactatctacAGCTACATTGTCTGCTGGGTTTCaggccagcagaata
miR-222	dre-mir-222	9	29779119	29779196	-1	tcatgagatgctcagtagtcagtgtagatcctgtgtcacaatcagcAGCTACATCTGGCTACTGGGTCTCtgatggca
miR-26a	rno-mir-26b_zf807	9	40025165	40025250	-1	aggctttggcctggTTCAAGTAATCCAGGATAGGCTtgtgatgtccggaaagcctattcgggatgacttggttcaggaatgagacc
miR-194	mmu-mir-194-1_zf238	10	22254907	22254974	1	actgctTGTAACAGCAACTCCATGTGGAaggtttgtgtcttccagtggagctgctgttgcgtgcagat

Table S2. Continued.

miRNA	Target miRNA	Chromosome / Scaffold	Position start	Position end	Strand	Precursor sequence
miR-215	mmu-mir-192_zf237	10	22255081	22255170	1	aggacacagggtgATGACCTATGAATTGACAGccagtgtttgcagtccagctgcctgtcagttctgtaggccactgccctgtttatccta
miR-130a	hsa-mir-130a_zf492	10	26782106	26782195	1	ctgtctgtccagtgccccttttatattgtactactgataacccagttattaaagCAGTGCAATGTTAAAAGGGCattggccagggatttt
miR-301	hsa-mir-301_zf1175	10	26782235	26782321	1	gctgttaacaggtgctctgacttcattgcactactgtattggacagctagCAGTGCAATAGTATTGTCAAAGCgtctgagagcagct
miR-122a	hsa-mir-122a_zf223	10	30424298	30424384	-1	gtcctccagagctgTGGAGTGTGACAATGGTGTTTGTatcatctgtcgtcaaacgccattatcacactaaatagccacggtgtgacc
miR-129	mmu-mir-129-2_zf144	10	35266100	35266189	1	agtctttcacgaatCTTTTTGCGGTCTGGGCTTGCTgttctcaactatcaatgggAAGCCCTTACCCCAAAAAGCATttgcggagggcgc
miR-129*	mmu-mir-129-2_zf144	10	35266100	35266189	1	agtctttcacgaatCTTTTTGCGGTCTGGGCTTGCTgttctcaactatcaatgggAAGCCCTTACCCCAAAAAGCATttgcggagggcgc
let-7f	mmu-let-7f-1_zf76	11	21854433	21854522	-1	tgcagtgTGAGGTAGTAGATTGTATAGTTgtagggtagtgattttatcctgtgtagaagataactatacaatctattgccttccctgagg
let-7a	hsa-let-7a-1_zf1	11	21854638	21854717	-1	tgggaTGAGGTAGTAGGTTGTATAGTTtagggtcacacccacagtgggagataactatacaacctactgtctttctcaa
miR-184	hsa-mir-184_zf990	12	3428982	3429066	1	cgaacacgtctccttatcacttttccagcccagctatccatttagtattcgtTGGACGGAGAACTGATAAGGGcatgtgcccgat
miR-107	hsa-mir-103-2_zf101	12	10935083	10935161	-1	gtgtgctctgagcttctttacagtgttgtcttgtggcatggagatcaAGCAGCATTGTACAGGGCTATCAcagcacact
miR-24	rno-mir-24-2_zf795	12	14796737	14796845	-1	cgactgagggctgttcccacttgtgcctgctaaactggtatcagtatgttgatttagtgcTGGCTCAGTTCAGCAGGAACAGgtgtgaagtcctctactctcctccaat
miR-31	hsa-mir-31_zf1206	12	15270963	15271034	1	gaagagatGGCAAGATGTTGGCATAGCTGttaatgtttatgggcctgctatgcctccatattgccatttctg
miR-10a	hsa-mir-10a_zf184	12	24838903	24839009	1	atgatgtctgtcatctatataTACCCTGTAGATCCGAATTTGTGtgaatatacagtcgcaaattcgtgtcttggggaatatgtagttgacataaacacaacgcaata
miR-200b	rno-mir-200b_zf736	12	32776761	32776856	1	tgattttggtagtcgtctccatcttacgaggcagcattggatttcattacttttTCTAATACTGCCTGGTAATGATGatgattgctgcctacaagc
miR-200a	rno-mir-200a_zf731	12	32776897	32776986	1	ctggcacttagcagccatcttaccggacagtgctggactgtataactgttttcTAACACTGTCTGGTAACGATGTttgttgggtgacctc
miR-135a	hsa-mir-135b_zf711	14	36260647	36260744	-1	acagctgtcgtgtctTATGGCTTTTTATTCCTATGTGAaggtgaacaaggctcatgtagggatacaagccactaaacacgcagtcagaaatcagcttt
miR-145	mmu-mir-145_zf186	14	50667840	50667911	1	cccgggGTCCAGTTTTCCCAGGAATCCCTTgggcaatcgaaaggggattcctggaaatactgttcttgggg
miR-21	hsa-mir-21_zf1054	15	8190598	8190670	-1	tgtcatgTAGCTTATCAGACTGGTGTTGgctttgagtttttggcaacagcagtctaataggctgtctgacatt
miR-132	mmu-mir-132_zf150	15	13343593	13343659	-1	accgtggcattagattgttactgtaggaacagaatttttggTAACAGTCTACAGCCATGGTCGctag
miR-140	mmu-mir-140_zf177	15	14343316	14343386	-1	cccgtcAGTGGTTTTACCCTATGGTAGgttacgtcatgctgttcTACCACAGGGTAGAACCACGGACggga
miR-140*	mmu-mir-140_zf177	15	14343316	14343386	-1	cccgtcAGTGGTTTTACCCTATGGTAGgttacgtcatgctgttcTACCACAGGGTAGAACCACGGACggga
miR-125b	hsa-mir-125b-1_zf355	15	16707758	16707846	-1	tgtgcctctcacaaTCCCTGAGACCCTAACTTGTGAcgttttcctgttatgtgcacgggttaggttcttgggagctgagaggggtgctc
let-7a	hsa-let-7a-2_zf10	15	16716574	16716646	-1	aggcTGAGGTAGTAGGTTGTATAGTTtagaataacatcactggagataactgtacaacctcctagctttccct
miR-22	hsa-mir-22_zf1092	15	20966725	20966813	1	ggctgacctgcagcagttcttcactggcaagctttatgtccttgtgtaccagctaAAGCTGCCAGCTGAAGAACTGTtgtggttggctc
let-7c	hsa-let-7c_zf941	15	27968976	27969060	-1	gcatccaggcTGAGGTAGTAGGTTGTATGGTTtagaattttgccctgggagttaactgtacaaccttctagctttccttggagct
miR-99a	hsa-mir-99a_zf25	15	27969230	27969311	-1	ccacttgtcattAACCCGTAGATCCGATCTTGTGataagtttgatggcacaagctcgattctatgggtctctgtctctgtgg
miR-30b	rno-mir-30c-2_zf875	16	21215359	21215443	-1	ccagtgtagtcgcTGTAAACATCCTACACTCAGCtgtgagctgcagacgaggctgggcggagggtgtttgctgtgactgtctgga
miR-30a-5p	mmu-mir-30a_zf414	16	21215653	21215724	-1	atgccTGTAAACATCCCCGACTGGAAGCtgtgctacgcggaaaacgagCTTTCAGTTGGATGTTTGCTGTca
miR-30a-3p	mmu-mir-30a_zf414	16	21215653	21215724	-1	atgccTGTAAACATCCCCGACTGGAAGCtgtgctacgcggaaaacgagCTTTCAGTTGGATGTTTGCTGTca
miR-125a	mmu-mir-125a_zf130	16	21873828	21873895	-1	ctttgTCCCTGAGACCCTTAACCTGTGaggtcaaactaggtcacaggtgaggtcctcaggaacagggc
miR-9	mmu-mir-9-1_zf455	16	26459236	26459325	1	aggggttggctgttaTCTTTGGTTATCTAGCTGTATGAgtgtgattcattcttcaTAAAGCTAGATAACCGAAAGTaacaagaatcccat
miR-9*	mmu-mir-9-1_zf455	16	26459236	26459325	1	aggggttggctgttaTCTTTGGTTATCTAGCTGTATGAgtgtgattcattcttcaTAAAGCTAGATAACCGAAAGTaacaagaatcccat
miR-7	dre-mir-7	16	41008014	41008124	1	agtttttttggggtgtagtccactgcggtgTGGAAGACTAGTGATTTTGTTgtttttttatgatggaaggagacaacaagtcatagtctgcctcacggctctgagatgaac
miR-29b	rno-mir-29b-2_zf846	18	30297489	30297570	1	ctgctcctggaagctgaattcagatggtgccatagagtattttatggcatcTAGCACCATTTGAAATCAGTGTtcctgggcc
miR-125a	hsa-mir-125a_zf302	19	5067210	5067295	-1	gatgtatgtctgtgTCCCTGAGACCCTTAACCTGTGatgtcttccaaggtcacaggtgaggtccttgggaacacggctgtatatga
miR-219	dre-mir-219	19	8032840	8032950	-1	tgatactgacggtcaggtgcTGATTGTCCAAACGCAATTCTtgtgcttgtgtgaaacaggagttgtggatggacatcacgctcctgtctggtcctcacatctacaggatgg
miR-187	dre-mir-187	19	39362923	39363027	1	gaaggtgacctgtggctgggccaggggctgcaacacaggacatgggagctgtctctcactcccgcTCGTGTCTTGTGTTGCAGCCagtggaacggctacactgtc
miR-214	dre-mir-214	20	15017075	15017184	-1	gctgccattgactgagagcgttgtctgtctgcctgtctacacttgctgtgcagaacttcctgcacctgtACAGCAGGCACAGACAGGCAGacagatggcagcccgcctga
miR-124a	hsa-mir-124a-1_zf260	20	48892450	48892535	-1	gcctgcttttcttcgtgttcacagcggaccttgatttaaatgtccatacaaTTAAGGCACGCGGTGAATGCCAagagagatggcaa

Table S2. Continued.

miRNA	Target miRNA	Chromosome / Scaffold	Position start	Position end	Strand	Precursor sequence
miR-128a	hsa-mir-128a_zf412	21	2177418	2177500	-1	gtgctgggagacggggccgtggcactgtatgagattcatgtaggctttcTCACAGTGAACCGGTCTCTTTTTccagccctcac
miR-23a	rno-mir-23a_zf784	23	9240278	9240353	1	gccaggggaattcctggcagagtgatttttaaacctaatgactgaATCACATTGCCAGGGATTTCCaatggctcgt
miR-34a	dre-mir-34a	23	26362501	26362602	-1	gctgctgtgagtggttctcTGGCAGTGTCTTAGCTGGTTGTTgtgtggagtgagaacgaagcaatcagcaagtatactgccgcagaaactcgtcacctttaa
miR-98	rno-mir-98_zf931	23	28646250	28646330	1	gTGAGGTAGTAAGTTGTGTTGTTgttggggatcagtatagtatggcccttgaaggagataactatacaatttactgccttc
miR-26b	hsa-mir-26a-2_zf1128	23	33374078	33374163	-1	ggcctttgcctggTTCAAGTAATCCAGGATAGGTTagttcccactagtacggcctattcttggttacttgtttcaggaggaggcta
miR-205	dre-mir-205	23	37357630	37357698	-1	attctaTCCTTCATTCCACCGGAGTCTGtgtagttgttcaatcagatttcagtggtgtgaagtgtagga
miR-29c	rno-mir-29c_zf852	23	38110649	38110739	1	tcctctacccaaggttgaccgatttcttttggtgttcagagtcttttggggtttcTAGCACCATTTGAAATCGGTTAcaatgaaggaccaa
miR-196a	dre-mir-196a	23	38564367	38564437	-1	gtggttTAGGTAGTTTCATGTTGTTGGGattggcttcctggctcgacaacaagaaactgccttgattacgt
miR-124a	mmu-mir-124a-3_zf129	24	19760014	19760081	1	ctttccgtgttcacagcggaccttgatttaatgtcttacaaTTAAGGCACGCGGTGAATGCCAagagc
miR-137	mmu-mir-137_zf170	24	23233663	23233736	1	cttcggtgacgggtattcttgggtggataatacggctctcgttgtTATTGCTTAAGAATACGCGTAGtcgagga
let-7i	mmu-let-7i_zf98	25	731320	731405	-1	ctggcTGAGGTAGTAGTTTGTGCTgttggttgggatgtgacattgcccgttatggagatgactgcgcaagctactcccttgccagt
let-7b	mmu-let-7b_zf41	25	923357	923443	1	acagggTGAGGTAGTAGGTTGTGTGGTTtcagggtagtgattttgccccatcaggagttaactatacaacctactgccttcccctgaa
miR-9	hsa-mir-9-3_zf1233	25	5760436	5760524	-1	gggggttggtttctcTCTTTGGTTATCTAGCTGTATGAgttataacactgtcaTAAAGCTAGATAACCGAAAGTagaaataattcccaa
miR-9*	hsa-mir-9-3_zf1233	25	5760436	5760524	-1	gggggttggtttctcTCTTTGGTTATCTAGCTGTATGAgttataacactgtcaTAAAGCTAGATAACCGAAAGTagaaataattcccaa
miR-210	dre-mir-210	25	20084871	20084981	-1	aacttgtgatttctgaaagcaggtaagccactgactaacgcacattgcgcctattctccactccaCTGTGCGTGTGACAGCGGCTaaccagcttttggatctgcacctgac
miR-133a	mmu-mir-133a-2_zf152	25	26140224	26140328	-1	ccccctaaaatgctttgctaaagctggtaaaatggaaccaaatcaactgttttatggatTTGGTCCCCTTCAACCAGCTGTagctgtgcattgatcaccccgtgg
miR-139	rno-mir-139_zf654	Zv4_NA7794	92	160	-1	ctgtatTCTACAGTGCATGTGTCTccagtgtttctatggcgactggggaggcagcgctgttggaataac
miR-181a	mmu-mir-213_zf289	Zv4_NA6392	163	250	-1	gtttgcctcagtgAACATTCAACGCTGTCGGTGAGTttgagctaaatggaaaaaaaccattgaccgttgattgtaccctgcggccgag
miR-1	hsa-mir-1-2_zf1179	Zv4_NA16429	1246	1332	1	gcctacttggtgtacatacttctttatgtgcccatatgaacatataaaagctaTGGAATGTAAAGAAGTATGTATtcttggtcaggt
miR-1	hsa-mir-1-1_zf1177	Zv4_NA2028	1374	1445	-1	tggtagacatacttctttatatgcccatatgaacaagagcagctaTGGAATGTAAAGAAGTATGTAtcccag
miR-128a	mmu-mir-128a_zf137	Zv4_NA10772	1758	1828	-1	ggacaggggccgtttctactgtcagagatgctgccactcgtcTCACAGTGAACCGGTCTCTTTTcctgct
miR-27a	mmu-mir-27a_zf370	Zv4_NA5039	2548	2635	-1	tttcgtgaggtgcaggacttagctcactctgtgaacagatctcggatatcctatgTTCACAGTGGCTAAGTTCCGCtcctctgaggcc
miR-223	dre-mir-223	Zv4_NA16498	3294	3404	1	tcctcctgatctagactcttctcttagagtatttgacagactgtggttgacactcgatctaaaggggTGTCAGTTTGTCAAATACCCCaagagaggggcacgagtcttca
miR-133a	rno-mir-133a_zf636	Zv4_NA16429	3440	3527	1	caatgctttgctaaagctggtaaaatggaaccaaatcacctcttcaatggatTTGGTCCCCTTCAACCAGCTGTagctatgctttgat
miR-101	rno-mir-101_zf566	Zv4_NA3505	3490	3565	-1	tgccctggttcagttatcacagtgctgatgctgtccatcttaaaggTACAGTACTGTGATAACTGAAGgatggctg
miR-7b	rno-mir-7-1_zf896	Zv4_NA9334	6008	6105	-1	cgctggcttgcttctgtgTGGAAGACTTGTGATTTTGTTgttgttagttagatgaagtgacaacaaatcacggtctgccctacagcacaggcccagca
miR-338	mmu-mir-338_zf433	Zv4_NA8796	6425	6523	1	ctgtctcaaggtttctccctgcaacaatctcctgatgctgcctgagtgtttttcttccacTCCAGCATCAGTGATTTTGTTGccggaggtcaccacaca
miR-181b	dre-mir-181b	Zv4_NA11193	6850	6928	1	aggtcataatcAACATTCATTGCTGTCGGTGGGTTtagtcttgtaacagctctctgaacaatgaatgtaactgtggccc
miR-30c	hsa-mir-30c-2_zf1194	Zv4_NA12788	7771	7843	1	agggagTGTAAACATCCTACACTCTCAGCtggagcgcagccgaggccgggagtgggatgtttgcgctctctgg
miR-7	dre-mir-7b	Zv4_NA8362	8981	9091	-1	cagacaccagcactgacccgattcgctgTGGAAGACTAGTGATTTTGTTgttgtttcttctgctttctgacaacaagtcacagtctacctcagcgagcgggccctttggct
miR-148a	mmu-mir-148a_zf187	Zv4_NA11997	10039	10136	-1	agctctctggctttccaagtaaagttctgtgatacactccgactctgaatgtttgcagTCAGTGCATTACAGAACTTTGTtttgggagtttaaagctc
miR-92	hsa-mir-92-1_zf1239	Zv4_NA18101	12121	12199	-1	ctttctgcgcaggttgggattggtagcaatgctgtgtgttttgaaggTATTGCACTTGTCCCGGCCTGTaaaggattgt
miR-19b	mmu-mir-19b-2_zf271	Zv4_NA18101	12239	12323	-1	ctggacccccggtcagttttgctggtttgcattcagcttttaagactgtgcgcTGTGCAAATCCATGCAAAACTGAttgtggcag
miR-20	hsa-mir-20_zf1037	Zv4_NA18101	12394	12466	-1	gcagtgcTAAAGTGCTTATAGTGCAGGTAGtatttctgtcatctactgcagtgtgagcacttgaagtacttct
miR-155	mmu-mir-155_zf197	Zv4_NA14515	12482	12547	1	ggtTTAATGCTAATCGTGATAGGGGtttagtgctgatgaacacctatgctgttagcattaatcttg
miR-19a	hsa-mir-19a_zf1019	Zv4_NA18101	12604	12686	-1	gcagttctctgctagttttgcatagttgcactacaagaaaacgggagtTGTGCAAATCTATGCAAAACTGAtggtggcctgct
miR-17-5p	hsa-mir-17_zf968	Zv4_NA18101	12891	12976	-1	gtcaatgtattgtCAAAGTGCTTACAGTGCAGGTAGTattatggaatatctactgcagtggaggcacttctagcaatacacttgac
miR-219	hsa-mir-219-2_zf1086	Zv4_NA1719	14981	15078	1	aatttgagctcaagcagtTGATTGTCCAAACGCAATTCTtgtgaaatgtcgagcgatcagtcgagaattgtgcctggacatctgttgctggaggctcc
miR-23a	hsa-mir-23a_zf1099	Zv4_NA1574	15161	15234	1	gctggagggattcctggcagagtgatttgggattatatcataaaATCACATTGCCAGGGATTTCCaaccagctg

Table S2. Continued.

miRNA	Target miRNA	Chromosome / Scaffold	Position start	Position end	Strand	Precursor sequence
miR-125b	mmu-mir-125b-2_zf132	Zv4_NA10299	37184	37255	1	tcctggTCCCTGAGACCCTAACTTGTGAgctttgtgtgctaaaaatcacaggttaagctcttgggacctggg
miR-103	hsa-mir-107_zf140	Zv4_scaffold2426	65501	65582	-1	ctcttggctttgagcctctttacaatgctgccttgtagacccagaatcaAGCAGCATTGTACAGGGCTATGAaagcacagag
miR-26a	hsa-mir-26a-1_zf1122	Zv4_NA6547	74574	74651	1	gttcccttgTTCAAGTAATCCAGGATAGGCTgtctgtcctggaggcctattcatgattacttgcactaggtggcagcc
miR-216	mmu-mir-216_zf303	Zv4_NA14029	81597	81669	1	ttggcaTAATCTCAGCTGGCAACTGTGagtagtgttttcatccctctcacaggcgctgctggggttctgtcac
miR-217	rno-mir-217_zf756	Zv4_NA14029	81807	81912	1	gagctgcatgagaactttctgatgttggtgaTACTGCATCAGGAACTGATTGGATgatattcaggagccatcagttcctgatgcactcccatcagcatcgaaagag
miR-124a	hsa-mir-124a-2_zf296	Zv4_NA3687	90919	91029	-1	tcggaaacttgggtttttgctctttgtgttcacagtggaccttgatttaatttcaatacaaTTAAGGCACGCGGTGAATGCCAagagagaagccaacagcagcactcacct
miR-193	hsa-mir-193_zf994	Zv4_scaffold546	225530	225618	1	gagagtgtgtcagaggctgggtctttgcgggcaaggtgagttttcctttcattcAACTGGCCTACAAAGTCCCAGttttcggcccatgt
miR-108	hsa-mir-108_zf181	Zv4_scaffold546	227188	227281	-1	gcactgcaagagcaATAAGGATTTTTAGGGGCATTatgactgggtcagtataacacagctgcccctaaaagtccctcacttttcttgctgactt
no_probes	mmu-mir-135a-1_zf162	Zv4_scaffold771	235037	235127	-1	attgctctttgccccgtatggcttttattcctatctgagaattgctcaggactcatatagggatggaagccatgcagggctgggggactg
miR-181a	mmu-mir-181a_zf210	Zv4_scaffold1793	247451	247527	1	tctgggAACATTCAACGCTGTCGGTGAGTttggatgcataataaaccatcgaccgttggctgtgccctgagattacc
no_probes	hsa-mir-148a_zf948	Zv4_scaffold1523	541347	541415	1	cggggaaagttctgtggtccactctggctgtgtgagtgtgtggtcagtgcattacagaactttgcttta
miR-199a	dre-mir-199a	Zv4_scaffold537	1355450	1355560	-1	tttttgtggacgcccgtcccgcctgCCCAGTGTTCAGACTACCTGTTCaggaattagtgtttgTACAGTAGTCTGCACATTGGTTaggctggatggggatgctcgag
miR-199a*	hsa-mir-199b_zf1017	Zv4_scaffold537	1355450	1355560	-1	tttttgtggacgcccgtcccgcctgCCCAGTGTTCAGACTACCTGTTCaggaattagtgtttgTACAGTAGTCTGCACATTGGTTaggctggatggggatgctcgag
miR-199a	dre-mir-199a	Zv4_scaffold537	1355471	1355541	-1	cgcctgCCCAGTGTTCAGACTACCTGTTCaggaattagtgtttgTACAGTAGTCTGCACATTGGTTaggct
miR-199a*	dre-mir-199a*	Zv4_scaffold537	1355471	1355541	-1	cgcctgCCCAGTGTTCAGACTACCTGTTCaggaattagtgtttgTACAGTAGTCTGCACATTGGTTaggct
miR-218	hsa-mir-218-1_zf1073	Zv4_scaffold336	2162754	2162864	-1	gaggaaattcagcggggtttttcctTTGTGCTTGATCTAACCATGTggttgcagactcagactaatacatggttctgtcaagcaccatggaaggtcttgcagcatctacagc
miR-17-5p	mmu-mir-17_zf207	Zv4_scaffold2543	3302739	3302823	1	gtcactgtagtgtCAAAGTGCTTACAGTGCAGGTAGTtcaatataatctactgcagtggaggcacttcaagctttaccgtgacgc
miR-18	hsa-mir-18_zf972	Zv4_scaffold2543	3302855	3302926	1	tatgcTAAGGTGCATTTAGTGCAGATAgtgaaatagactagcacctactgccctaagtgccccttctggcat

[§] zebrafish genome assembly Zv4.

[¶] Mature part of *mir* sequences are in uppercase.

Table S3. Comparison of miRNA expression profiles in mammals and zebrafish determined by microarrays and *in situ* hybridization.

MicroRNA	Class*	Overlap†	1. Expression array mammals‡	2. Expression array zebrafish§	3. *In situ* expression pattern zebrafish
miR-1	A	1 - 2 - 3	Heart; muscle	Muscle, heart	Body, head and fin muscles
miR-122a	A	1 - 2 - 3	Liver	Liver, gut	Liver; pancreas
miR-124a	A	1 - 2 - 3	Brain	Brain, eye	Differentiated cells of brain; spinal cord and eyes; cranial ganglia
miR-128a	A	1 - 2 - 3	Brain	Brain	Brain (specific neurons in fore- mid- and hindbrain); spinal cord; cranial nerves/ganglia
miR-133a	A	1 - 2 - 3	Heart; muscle	Muscle, heart	Body, head and fin muscles
miR-138	A	1 - 2 - 3	Brain; bone marrow	Brain	Outflow tract of the heart; brain; cranial nerves/ganglia; undefined bilateral structure in head; neurons in spinal cord
miR-144	A	1 - 2 - 3	Bone marrow; heart	Heart	Blood
miR-194	A	1 - 2 - 3	Gut (small and large intestine); liver; stomach	Gut, liver	Gut and gall bladder; liver; pronephros
miR-206	A	1 - 2 - 3	Heart; muscle	Muscle	Body, head and fin muscles
miR-219	A	1 - 2 - 3	Brain	Brain	Brain (mid- and hindbrain); spinal cord
miR-338	A	1 - 2 - 3	Brain	Brain, muscle	Lateral line; cranial ganglia
miR-9	A	1 - 2 - 3	Brain	Brain	Proliferating cells of brain, spinal cord and eyes
miR-9*	A	1 - 2 - 3	Brain	Brain	Proliferating cells of brain, spinal cord and eyes
miR-200a	A	1 - 2 - 3	Pancreas; colon; kidney	Fins, skin, gut	Nose epithelium; lateral line organs; epidermis; gut (proctodeum); taste buds
miR-132	A	(1) - 2 - 3	(Brain)	Brain	Brain (specific neurons in fore- and midbrain)
miR-142-5p	A	1 - (2) - 3	Bone marrow; heart; muscle	(Gills)	Thymic primordium
miR-7	A	1 - (2) - 3	Brain; (adrenal gland)	(Brain)	Neurons in forebrain; diencephalon/hypothalamus; pancreatic islet
miR-143	A	1 - 3, 2 - 3	Bladder; ovary	Heart	Gut and gall bladder; swimbladder; heart; nose
miR-145	A	1 - 3, 2 - 3	Bladder	Heart	Gut and gall bladder; gills; swimbladder; branchial arches; fins; outflow tract of the heart; ear
miR-181a	A	1 - 3, 2 - 3	Brain, Thymus	Eye	Brain (tectum, telencephalon); eyes; thymic primordium; gills
miR-181b	A	1 - 3, 2 - 3	Brain, Thymus	Eye	Brain (tectum, telencephalon); eyes; thymic primordium; gills
miR-215	A	1 - (2) - 3	Gut (small intestine)	(Gut, liver)	Gut and gall bladder
let-7a	A	1 - 3	Brain; ubiquitous	Ubiquitous	Brain; spinal cord
let-7b	A	1 - 3	Brain; ubiquitous	Fins	Brain; spinal cord
miR-125a	A	1 - 3	Brain	Ubiquitous	Brain; spinal cord; cranial ganglia

Table S3. Continued.

MicroRNA	Class	Overlap	1. Expression array mammals	2. Expression array zebrafish	3. *In situ* expression pattern zebrafish
miR-125b	A	1 - 3	Brain	Ubiquitous	Brain; spinal cord; cranial ganglia
miR-142-3p	A	1 - 3	Thymus; bone marrow; heart; muscle; lung; spleen; Lymphocytes	Ubiquitous	Thymic primordium; blood cells
miR-200b	A	1 - 3	Pancreas; colon; kidney	Ubiquitous	Nose epithelium; lateral line organs; epidermis; gut (proctodeum); taste buds
miR-218	A	1 - 3	Brain	ND	Brain (neurons and/or cranial nerves/ganglia in hindbrain); spinal cord
miR-222	A	1 - 3	Brain	Ubiquitous	Neurons and/or cranial ganglia in forebrain and midbrain; rhombomere in early stages
miR-23a	A	1 - 3	Lung	Fins	Pharyngeal arches; oral cavity; posterior tail; cardiac valves
miR-27a	A	1 - 3	Lung	Fins, skin	Undefined structures in branchial arches; tip of tail in early stages
miR-34a	A	1 - 3	Brain	Ubiquitous	Brain (cerebellum); neurons in spinal cord
miR-375	A	1 - 3	Insulin-secreting cells of pancreas	no data	Pituitary gland; pancreatic islet
miR-99a	A	1 - 3	Brain	Heart	Brain (hindbrain, diencephalon); spinal cord
let-7i	A	2 - 3	Ubiquitous	Brain	Brain (tectum, diencephalon)
miR-100	A	2 - 3	Ovary	Brain	Brain (hindbrain, diencephalon); spinal cord
miR-103	A	2 - 3	ND	Brain	Brain; spinal cord
miR-107	A	2 - 3	ND	Brain	Brain; spinal cord
miR-126	A	2 - 3	Lung	Heart	Blood vessels and heart
miR-137	A	2 - 3	Ubiquitous	Brain	Brain (neurons and/or cranial nerves/ganglia in fore-, mid- and hindbrain); spinal cord
miR-140	A	2 - 3	ND	Gills	Cartilage of pharyngeal arches,head skeleton and fins
miR-140*	A	2 - 3	No data	Gills	Cartilage of pharyngeal arches,head skeleton and fins
miR-141	A	2 - 3	Ubiquitous	Fins, skin, gut	Nose epithelium; lateral line organs; epidermis; gut (proctodeum); taste buds
miR-150	A	2 - 3	Thymus; bone marrow;lymphocytes	Gills	Cardiac valves; undefined structures in epithelium of branchial arches
miR-182	A	2 - 3	(Thymus)	Eye	Nose epithelium; haircells of lateral line organs and ear; cranial ganglia; rods, cones and bipolar cells of eye; epiphysis
miR-183	A	2 - 3	ND	Eye	Nose epithelium; haircells of lateral line organs and ear; cranial ganglia; rods, cones and bipolar cells of eye; epiphysis
miR-184	A	2 - 3	ND	Eye, fins	Lens; hatching gland in early stages
miR-199a	A	2 - 3	Ubiquitous	Fins, skin	Epithelia surrounding cartilage of pharyngeal arches, oral cavity and pectoral fins; epidermis of head; tailbud

Table S3. Continued.

MicroRNA	Class	Overlap	1. Expression array mammals	2. Expression array zebrafish	3. *In situ* expression pattern zebrafish
miR-199a*	A	2 - 3	Ovary	Fins, skin	Epithelia surrounding cartilage of pharyngeal arches, oral cavity and pectoral fins; epidermis of head; tailbud
miR-203	A	2 - 3	Ubiquitous	Gills, fins, skin	Most outer layer of epidermis
miR-204	A	2 - 3	ND	Skin, eye	Neural crest; pigment cells of skin and eye; swimbladder
miR-205	A	2 - 3	Thymus	Skin, fins	Epidermis; epithelia of branchial arches; intersegmental cells; not in sensory epithelia
miR-221	A	2 - 3	Ubiquitous	Brain	Brain (Neurons and/or cranial ganglia in forebrain and midbrain; rhombomere in early stages)
miR-7b	A	2 - 3	(Adrenal gland)	Brain	Brain (fore-, mid- and hindbrain); spinal cord
miR-96	A	2 - 3	ND	Eye	Nose epithelium; haircells of lateral line organs and ear; cranial ganglia; rods, cones and bi-polar cells of eye; epiphysis
miR-217	B	(1) - (2) - 3	(Pancreas)	(Gut)	Brain (tectum, hindbrain); spinal cord; proliferative cells of eyes; pancreas
miR-126*	B	(1) - (2)	(Lung)	(Heart)	ND
miR-31	B	(1) - (2)	(Gut, large intestine)	(Gut)	Ubiquitous
miR-216	B	(1) - 3	(Pancreas)	ND	Brain (tectum); spinal cord; proliferative cells of eyes; pancreas; body muscles
miR-30a-5p	B	(1) - 3	(Kidney)	Heart	Pronephros; cells in epidermis; lens in early stages
miR-153	B	(2) - 3	ND	(Brain)	Brain (fore- mid- and hindbrain, diencephalon/hypothalamus)
miR-15a	C	1 - 2 - 3	Ubiquitous	Ubiquitous	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)
miR-17-5p	C	1 - 2 - 3	Ubiquitous	Ubiquitous	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)
miR-18	C	1 - 2 - 3	Ubiquitous	Ubiquitous	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)
miR-195	C	1 - 2 - 3	Ubiquitous	Ubiquitous	Ubiquitous
miR-19b	C	1 - 2 - 3	Ubiquitous	Ubiquitous	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)
miR-20	C	1 - 2 - 3	Ubiquitous	Ubiquitous	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)
miR-26a	C	1 - 2 - 3	Ubiquitous	Ubiquitous	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)
miR-92	C	1 - 2 - 3	Ubiquitious	Ubiquitous	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)
let-7c	C	1 - 2	Ubiquitous	Ubiquitous	Brain; spinal cord
miR-101	C	1 - 2	Ubiquitous, (brain)	Ubiquitous	ND
miR-16	C	1 - 2	Ubiquitous	Ubiquitous	Brain
miR-21	C	1 - 2	Ubiquitous	Ubiquitous	Cardiac valves; otoliths in ear; rhombomere in early stages
miR-30b	C	1 - 2	Ubiquitous	Ubiquitous	Pronephros; cells in epidermis

Table S3. Continued.

MicroRNA	Class	Overlap	1. Expression array mammals	2. Expression array zebrafish	3. *In situ* expression pattern zebrafish
miR-30c	C	1 - 2	Ubiquitous	Ubiquitous	Pronephros; cells in epidermis and epithelia of branchial arches; neurons in hindbrain
miR-26b	C	1 - 3	Ubiquitous	Heart	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)
let-7g	C	2 - 3	Brain	Ubiquitous	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)
miR-19a	C	2 - 3	ND	Ubiquitous	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)
miR-210	C	2 - 3	ND	Ubiquitous	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)
miR-22	C	2 - 3	Heart	Ubiquitous	Ubiquitous
miR-25	C	2 - 3	ND	Ubiquitous	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)
miR-93	C	2 - 3	ND	Ubiquitous	Ubiquitous (head, spinal cord, gut, outline somites, neuromasts)
miR-189	D	1 - 2 - 3	ND	ND	ND
miR-30a-3p	D	1 - 2 - 3	ND	ND	ND
miR-34b	D	1 - 2	ND	ND	Cells in pronephric duct; nose
miR-129*	D	1 - 3	ND	Brain	ND
miR-135a	D	1 - 3	ND	Muscle	ND
miR-182*	D	1 - 3	ND	(Eye)	ND
miR-187	D	1 - 3	ND	Ubiquitous	ND
miR-220	D	1 - 3	ND	Ubiquitous	ND
miR-301	D	1 - 3	ND	Brain	ND
miR-223	D	2 - 3	Bone marrow; lung	ND	ND
let-7f	-	-	Ubiquitous	Skin	Brain; spinal cord
miR-108	-	-	ND	no data	Ubiquitous
miR-10a	-	-	(Kidney)	Ubiquitous	Posterior trunk; later restricted to spinal cord
miR-10b	-	-	(Testes)	Muscle	Posterior trunk; later restricted to spinal cord
miR-129	-	-	ND	Ubiquitous	Brain
miR-130a	-	-	Ovary	Ubiquitous	ND
miR-139	-	-	Spleen; brain	ND	Nose; neuromasts
miR-146	-	-	Thymus	Ubiquitous	Neurons in forebrain; branchial arches and head skeletion
miR-148a	-	-	Ubiquitous	(Fins)	ND

Table S3. Continued.

MicroRNA	Class	Overlap	1. Expression array mammals	2. Expression array zebrafish	3. *In situ* expression pattern zebrafish
miR-152	-	-	ND	Gills	Ubiquitous
miR-155	-	-	(Ovary)	(Eye, liver)	ND
miR-190	-	-	Ubiquitous	(Brain)	ND
miR-193	-	-	Lung	Muscle	ND
miR-196a	-	-	(Bone marrow)	Ubiquitous	Posterior trunk; later restricted to spinal cord
miR-213	-	-	ND	Ubiquitous	Nose (epithelium or olfactory neurons), eyes (ganglion cell layer)
miR-214	-	-	ND	Ubiquitous	Epithelia surrounding cartilage of pharyngeal arches, oral cavity and pectoral fins; epidermis of head; tailbud
miR-24	-	-	Ubiquitous	Fins	Pharyngeal arches; oral cavity; posterior tail; cardiac valves
miR-27b	-	-	Ubiquitous	Fins	Cells in branchial arches
miR-29a	-	-	Brain	Ubiquitous	ND
miR-29b	-	-	Brain	Ubiquitous	ND
miR-29c	-	-	Brain	Ubiquitous	ND
miR-98	-	-	Ubiquitous	Skin	Brain

* Main class in which expression patterns were compared: A, specific expression; B, marginal specific expression or very low absolute expression; C, ubiquitous expression. D, no detectable expression.

† Overlap between the different datasets: 1, expression in mammals detected by microarray; 2, expression in zebrafish detected by microarray; 3, expression in zebrafish detected by *in situ* hybridization. Overlaps with marginal specific expression are in brackets.

‡ Predominant miRNA expression in mammals. This is based on microarray experiments by (*14-17*) and recent literature data (*18*). Marginal expression is indicated in brackets.

§ Predominant miRNA expression in adult zebrafish organs/parts. Expression is only indicated if it is 6-fold higher than the relative median expression for liver and 3-fold higher for the other tissues. Expression for testis and ovary is not given because it cannot be compared with the *in situ* data. Very low absolute expression is indicated in brackets. Due to difficulty in organ isolation, the heart and gills samples contain RNA of blood and the gut and liver samples might contain contaminating RNA of pancreas, gut and liver.

Chapter 7

Discussion

Reverse genetics in zebrafish

To date, the only functioning reverse genetic method to make gene knockouts in zebrafish is target-selected mutagenesis, as described in this thesis. In zebrafish, target-selected mutagenesis is accomplished by random *N*-Ethyl-*N*-nitrosourea (ENU) mutagenesis followed by targeted screening of the genes of interests. Several laboratories now routinely use this method to produce knockout- and missense alleles of their genes of interest. Since no alternative methods are available yet, target-selected mutagenesis will probably remain the method of choice in zebrafish to produce gene knockouts in the near future.

The success of finding gene knockouts by target-selected mutagenesis depends on a few factors: the size and quality of the mutant library and the efficiency of mutation detection. In principle, a well-mutagenized library consisting of 5,000 to 10,000 individuals contains at least one knockout mutation for an average gene [1]. Many laboratories, however, do not have the ability to raise and keep such high number of animals. To circumvent this, these laboratories could share the DNA samples of several small libraries or form a consortium to screen sufficient mutagenized animals.

The mutation frequency is another factor that determines the success rate of finding knockout mutations. The higher this frequency, the more chance there is of finding a knockout mutation and the smaller a library needs to be. Therefore it is extremely important to obtain the utmost mutation frequency. In zebrafish, the ENU mutagenesis has been investigated thoroughly to obtain the optimal mutagenesis regime [2-4]. Increasing the ENU dose or the number of consecutive treatments might result in higher mutation frequencies, but also in decreased survival, reduced fertility and increased chance of clonal mutants (although the last has never been shown). Since different zebrafish lines might have a different susceptibility to ENU it remains worthwhile to keep investigating these parameters for the optimal mutation frequency. Many of the ENU-induced lesions are restored by the DNA repair systems. Suppression of these DNA repair systems, such as the mismatch repair system [5], might also result in increased mutation frequencies.

Another major factor contributing to the success rate of target-selected mutagenesis in zebrafish is the mutation detection efficiency. The first gene knockout was found by direct resequencing of the target gene [6]. DNA sequencing is regarded as the golden standard in mutation detection. However, for the discovery of subsequent mutations in several other genes we implemented an enzyme-mediated mismatch recognition procedure (TILLING) to prescreen for mutations. This method reduced both time and costs for the mutation detection compared to DNA sequencing, while being equally sensitive as DNA sequencing [7]. Recently, we switched back to DNA sequencing again, because technological advances in capillary sequencing techniques and automation, considerably increased the throughput, while reducing the cost to a similar level as TILLING. Future advances in DNA sequencing technologies, such as two recently developed methodologies [8, 9] will probably further speed up and reduce the costs for the detection of mutations.

The detection of ENU-induced mutations is, in principle, not different from the detection of single nucleotide polymorphisms (SNPs). For the genome-wide discovery and genotyping of SNPs several microarray-based methods have been developed. These include high-density microarrays, which are used to genotype 10,000 to 100,000 SNPs on a single slide (see [10] for review). Instead of genotyping SNPs, these slides might also be used to resequence every position in a gene that can be mutated by ENU. Alternatively, if only the positions that result in nonsense codons and/or splice sites alterations when mutated are genotyped, thousands of genes might be screened simultaneously for potential knockout mutations in a single hybridization experiment. A tilling path of these slides is then needed to screen for all the potential knockout mutations present in a single mutagenized animal. Screening

a couple of thousand mutagenized animals may then result in knockouts in all the genes of a particular species, the mutome.

MicroRNAs in zebrafish development

miRNAs form a large class of post-transcriptional gene regulators that are widely present in multicellular organisms, ranging from plants to humans. Animal genomes encode several hundred miRNAs and recent estimates indicate that vertebrate genomes may encode up to a thousand different miRNAs [11, 12], of which the individual functions are largely unknown. However, the current set of known miRNAs is predicted to regulate several thousands target mRNAs, which may go up to 30% of all protein-coding genes [13-15]. Together with the strong evolutionary conservation of many miRNAs and their targets, this magnitude suggests that miRNAs have diverse and important roles in vertebrate development. To address this role we studied miRNAs in zebrafish embryonic development. This thesis describes the consequences of removing all miRNAs in zebrafish [16] and it describes the miRNA expression patterns during zebrafish development [17].

miRNAs are essential for vertebrate development

All miRNAs were removed from zebrafish embryos by disruption of the miRNA-producing enzyme Dicer [16]. Initially, *dicer* mutants develop normally, but they run out of mature miRNAs and arrest ~8 days after fertilization. At this point all the major organs have been formed. Knockdown of *dicer* mRNA using antisense morpholino oligonucleotides results in an earlier arrest, indicating the presence of maternal *dicer* mRNA. However, during the first 24 hours post-fertilization these knockdown embryos still develop without obvious phenotypes. Together, this suggests that Dicer and presumably miRNAs in general are not essential for the earliest developmental processes, but are needed for development and growth of tissues at later stages [16].

Since a role for maternal Dicer protein in early development could not be excluded, Giraldez and coworkers generated maternal-zygotic *dicer* mutant zebrafish from the *dicer* mutant described in this thesis [18]. Like *dicer* knockdown embryos, these maternal-zygotic mutants also develop quite normally during early embryogenesis and only display some mild defects. They have intact axis formation and cell regionalization and differentiate into multiple cell types and tissues, indicating normal cell-fate determination and patterning. However, they show morphogenesis defects during gastrulation, brain formation, somitogenesis and heart formation [18]. These results support our initial hypothesis that Dicer and miRNAs are not essential for the early zebrafish development, but are necessary for subsequent later steps in development.

Dicer is also essential for mouse development. Dicer knockout embryos die before axis formation and are depleted of ES cells [19]. Dicer-deficient mouse ES cells proliferate poorly, display defects in differentiation and fail to produce miRNAs [20, 21]. Furthermore, conditional knockout alleles and hypomorphic alleles indicate that Dicer and miRNAs are also important for late developmental processes in the mouse, such as angiogenesis [22], T cell lineage choice and differentiation [23, 24] and limb morphogenesis [25].

Although the absence of miRNAs in loss-of-Dicer zebrafish and mouse indicates that miRNAs are essential for vertebrate development, a role for other RNAi-related processes that involve Dicer, such as the formation of heterochromatin structures and centromeric silencing [20, 21, 26] cannot be excluded. However, rescue of the brain phenotype of Dicer mutant zebrafish embryos by injection of miR-430 duplexes [18] indicates that miRNAs are directly involved in vertebrate development. In addi-

tion, overexpression of individual miRNAs in mice and cell culture studies show that miRNAs are involved in developmental processes such as HOX gene regulation [27], cardiomyocyte differentiation [28], hematopoietic lineage differentiation [29], adipocyte differentiation [30] and cancer [31].

miRNAs are regulators of differentiation

To initiate the study on individual miRNAs during zebrafish embryonic development we determined the temporal and *in situ* expression patterns of all conserved miRNAs by microarray analysis and *in situ* hybridizations [17]. None of these miRNAs show detectable expression during the early developmental stages up to segmentation. A recent study describing the cloning frequencies of miRNAs during zebrafish development also shows this lack of miRNA expression in the early stages [32]. An exception to this is the miR-430 family [18, 32], which is involved in brain morphogenesis [18]. However, the overall absence indicates that miRNAs, in general, are not essential for early zebrafish development. In addition, it is in full agreement with the near absence of early phenotypes in the *dicer* mutants [16, 18], suggesting that that miRNAs are not required for cell-fate specification and patterning during early development.

At later stages of zebrafish embryonic development ~80% of the conserved miRNAs is expressed in a highly tissue-specific manner. Furthermore, most of the tissues that are present in zebrafish embryos have specific expression of one or several miRNAs [17]. This widespread tissue-specificity suggests that miRNAs have important roles in differentiation or more likely the maintenance of cell and tissue identity in multicellular animals, like vertebrates. As reviewed in the introduction of this thesis, several recent studies support this hypothesis. First, microarray analysis shows that the transfection of tissue-specific miRNAs into HeLa cells shifts the mRNA expression profile of the HeLa cells to that of the tissue in which these miRNAs normally are expressed. In other words, these miRNAs dampen the expression of many genes that are normally not expressed in the corresponding tissues [33]. Second, there is a general downregulation of miRNA expression in human tumors compared with normal tissues from which these tumors have arisen [34]. Third, zebrafish fully depleted of miRNAs differentiate multiple tissues types but arrest thereafter [18].

Together, with the high degree of conservation of many miRNAs (some are conserved from nematode to human) and the combinatorial coding capacity of miRNAs [33, 35], these results indicate that miRNAs are important regulators of differentiation that act by dampening the expression of large numbers of genes within a tissue or cell type. The identification of the full sets of miRNAs within animals and their interactions with target genes will be the key to find the function of all individual miRNAs. Towards this end the overexpression and/or depletion of individual and multiple miRNAs by either knockdown or knockout strategies, such as recently been done in *Drosophila* and *C. elegans* [36, 37], will undoubtedly reveal many functions of miRNAs during development and disease.

References

1. Wienholds, E., and Plasterk, R.H. (2004). Target-selected gene inactivation in zebrafish. Methods Cell Biol *77*, 69-90.
2. van Eeden, F.J., Granato, M., Odenthal, J., and Haffter, P. (1999). Developmental mutant screens in the zebrafish. Methods Cell Biol *60*, 21-41.
3. Solnica-Krezel, L., Schier, A.F., and Driever, W. (1994). Efficient recovery of ENU-induced mutations from the zebrafish germline. Genetics *136*, 1401-1420.
4. Mullins, M.C., Hammerschmidt, M., Haffter, P., and Nusslein-Volhard, C. (1994). Large-scale mutagenesis in the zebrafish: in search of genes controlling development in a vertebrate. Curr Biol *4*, 189-202.

5. Claij, N., van der Wal, A., Dekker, M., Jansen, L., and te Riele, H. (2003). DNA mismatch repair deficiency stimulates N-ethyl-N-nitrosourea-induced mutagenesis and lymphomagenesis. Cancer Res *63*, 2062-2066.
6. Wienholds, E., Schulte-Merker, S., Walderich, B., and Plasterk, R.H. (2002). Target-selected inactivation of the zebrafish rag1 gene. Science *297*, 99-102.
7. Wienholds, E., van Eeden, F., Kosters, M., Mudde, J., Plasterk, R.H., and Cuppen, E. (2003). Efficient target-selected mutagenesis in zebrafish. Genome Res *13*, 2700-2707.
8. Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., and Rothberg, J.M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. Nature.
9. Shendure, J., Porreca, G.J., Reppas, N.B., Lin, X., McCutcheon, J.P., Rosenbaum, A.M., Wang, M.D., Zhang, K., Mitra, R.D., and Church, G.M. (2005). Accurate multiplex polony sequencing of an evolved bacterial genome. Science *309*, 1728-1732.
10. Syvanen, A.C. (2005). Toward genome-wide SNP genotyping. Nat Genet *37 Suppl*, S5-10.
11. Berezikov, E., Guryev, V., van de Belt, J., Wienholds, E., Plasterk, R.H., and Cuppen, E. (2005). Phylogenetic shadowing and computational identification of human microRNA genes. Cell *120*, 21-24.
12. Bentwich, I., Avniel, A., Karov, Y., Aharonov, R., Gilad, S., Barad, O., Barzilai, A., Einat, P., Einav, U., Meiri, E., Sharon, E., Spector, Y., and Bentwich, Z. (2005). Identification of hundreds of conserved and nonconserved human microRNAs. Nat Genet *37*, 766-770.
13. John, B., Enright, A.J., Aravin, A., Tuschl, T., Sander, C., and Marks, D.S. (2004). Human MicroRNA targets. PLoS Biol *2*, e363.
14. Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell *120*, 15-20.
15. Xie, X., Lu, J., Kulbokas, E.J., Golub, T.R., Mootha, V., Lindblad-Toh, K., Lander, E.S., and Kellis, M. (2005). Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. Nature *434*, 338-345.
16. Wienholds, E., Koudijs, M.J., van Eeden, F.J., Cuppen, E., and Plasterk, R.H. (2003). The microRNA-producing enzyme Dicer1 is essential for zebrafish development. Nat Genet *35*, 217-218.
17. Wienholds, E., Kloosterman, W.P., Miska, E., Alvarez-Saavedra, E., Berezikov, E., de Bruijn, E., Horvitz, H.R., Kauppinen, S., and Plasterk, R.H. (2005). MicroRNA expression in zebrafish embryonic development. Science *309*, 310-311.
18. Giraldez, A.J., Cinalli, R.M., Glasner, M.E., Enright, A.J., Thomson, J.M., Baskerville, S., Hammond, S.M., Bartel, D.P., and Schier, A.F. (2005). MicroRNAs regulate brain morphogenesis in zebrafish. Science *308*, 833-838.
19. Bernstein, E., Kim, S.Y., Carmell, M.A., Murchison, E.P., Alcorn, H., Li, M.Z., Mills, A.A., Elledge, S.J., Anderson, K.V., and Hannon, G.J. (2003). Dicer is essential for mouse development. Nat Genet *35*, 215-217.
20. Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev *19*, 489-501.
21. Murchison, E.P., Partridge, J.F., Tam, O.H., Cheloufi, S., and Hannon, G.J. (2005). Characterization of Dicer-deficient murine embryonic stem cells. Proc Natl Acad Sci U S A *102*, 12135-12140.
22. Yang, W.J., Yang, D.D., Na, S., Sandusky, G.E., Zhang, Q., and Zhao, G. (2005). Dicer Is Required for Embryonic Angiogenesis during Mouse Development. J Biol Chem *280*, 9330-9335.
23. Cobb, B.S., Nesterova, T.B., Thompson, E., Hertweck, A., O'Connor, E., Godwin, J., Wilson, C.B., Brockdorff, N., Fisher, A.G., Smale, S.T., and Merkenschlager, M. (2005). T cell lineage choice and differentiation in the absence of the RNase III enzyme Dicer. J Exp Med *201*, 1367-1373.
24. Muljo, S.A., Ansel, K.M., Kanellopoulou, C., Livingston, D.M., Rao, A., and Rajewsky, K. (2005). Aberrant T cell differentiation in the absence of Dicer. J Exp Med *202*, 261-269.
25. Harfe, B.D., McManus, M.T., Mansfield, J.H., Hornstein, E., and Tabin, C.J. (2005). The RNaseIII enzyme Dicer is required for morphogenesis but not patterning of the vertebrate limb. Proc Natl Acad Sci U S A *102*, 10898-10903.
26. Fukagawa, T., Nogami, M., Yoshikawa, M., Ikeno, M., Okazaki, T., Takami, Y., Nakayama, T., and Oshimura, M. (2004). Dicer is essential for formation of the heterochromatin structure in vertebrate cells. Nat Cell Biol *6*, 784-791.
27. Yekta, S., Shih, I.H., and Bartel, D.P. (2004). MicroRNA-directed cleavage of HOXB8 mRNA. Science *304*, 594-596.
28. Zhao, Y., Samal, E., and Srivastava, D. (2005). Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. Nature *436*, 214-220.
29. Chen, C.Z., Li, L., Lodish, H.F., and Bartel, D.P. (2004). MicroRNAs modulate hematopoietic lineage differentiation. Science *303*, 83-86.
30. Esau, C., Kang, X., Peralta, E., Hanson, E., Marcusson, E.G., Ravichandran, L.V., Sun, Y., Koo, S., Perera, R.J., Jain, R., Dean, N.M., Freier, S.M., Bennett, C.F., Lollo, B., and Griffey, R. (2004). MicroRNA-143 regulates adipocyte differentiation. J Biol Chem *279*, 52361-52365.
31. He, L., Thomson, J.M., Hemann, M.T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S.W., Hannon, G.J., and Hammond, S.M. (2005). A microRNA polycistron as a potential human oncogene. Nature *435*, 828-833.
32. Chen, P.Y., Manninga, H., Slanchev, K., Chien, M., Russo, J.J., Ju, J., Sheridan, R., John, B., Marks, D.S., Gaidatzis, D., Sander, C., Zavolan, M., and Tuschl, T. (2005). The developmental miRNA profiles of zebrafish as determined by small RNA cloning. Genes Dev *19*, 1288-1293.

33. Lim, L.P., Lau, N.C., Garrett-Engele, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S., and Johnson, J.M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature *433*, 769-773.
34. Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A., Downing, J.R., Jacks, T., Horvitz, H.R., and Golub, T.R. (2005). MicroRNA expression profiles classify human cancers. Nature *435*, 834-838.
35. Bartel, D.P., and Chen, C.Z. (2004). Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. Nat Rev Genet *5*, 396-400.
36. Leaman, D., Chen, P.Y., Fak, J., Yalcin, A., Pearce, M., Unnerstall, U., Marks, D.S., Sander, C., Tuschl, T., and Gaul, U. (2005). Antisense-mediated depletion reveals essential and specific functions of microRNAs in Drosophila development. Cell *121*, 1097-1108.
37. Abbott, A.L., Alvarez-Saavedra, E., Miska, E.A., Lau, N.C., Bartel, D.P., Horvitz, H.R., and Ambros, V. (2005). The let-7 MicroRNA Family Members mir-48, mir-84, and mir-241 Function Together to Regulate Developmental Timing in Caenorhabditis elegans. Dev Cell *9*, 403-414.

Samenvatting

De zebravis (*Danio rerio*) is een belangrijk modelorganisme voor verschillende onderzoeksgebieden in de moderne biologie. Dit modelorganisme wordt voornamelijk veel gebruikt in de ontwikkelingsbiologie omdat het uitermate geschikt is om de embryonale ontwikkeling, de periode van bevruchte eicel tot aan het uitkomen van het ei of de geboorte, van gewervelde dieren te bestuderen. In tegenstelling tot zoogdieren, zoals de mens, ontwikkelen zebravisembryo's zich buiten het lichaam van de moeder. Bovendien is de embryonale ontwikkeling veel sneller dan bij zoogdieren. Ongeveer 48 uur na de bevruchting hebben de embryo's al de basiskenmerken van een gewerveld dier en zijn belangrijke organen, zoals hersenen, hart, ogen en spieren, gevormd. Tijdens dit gehele proces zijn de embryo's optisch transparant, waardoor de ontwikkeling van een embryo vanaf de eerste celdeling tot aan de vorming van de diepste inwendige organen goed zichtbaar is onder een standaardmicroscoop.

De zebravis is ook uitermate geschikt voor genetica, de studie naar erfelijkheid en variatie. Vooral de klassieke genetica, volgens de overervingwetten van Mendel, wordt veel gebruikt om de erfelijke factoren te vinden die verantwoordelijk zijn voor bepaalde genetische afwijkingen. De belangrijkste erfelijke factoren die in het DNA van alle levende organismen gecodeerd worden zijn genen. Het DNA van de mens en de zebravis bevat ieder ongeveer 25.000 tot 30.000 genen. Het merendeel hiervan heeft dezelfde of vergelijkbare functies in mens en vis. Het doel van genetica is de relatie tussen deze genen en hun functie te vinden. Vooral de genen die specifiek betrokken zijn bij de vorming en het functioneren van gewervelde dieren zijn veelal gelijk in structuur en functie. Deze zijn dus het meest relevant voor ontwikkelingsbiologisch onderzoek.

Een manier om de functie van een gen te vinden is het gevolg van de uitschakeling van het gen te onderzoeken. Genen kunnen worden uitgeschakeld door het aanbrengen van DNA-veranderingen, zogenaamde mutaties. In de zebravis kunnen mutaties willekeurig worden aangebracht en worden genen dus ook willekeurig uitgeschakeld. Dit levert een grote variëteit aan vissen met uiterlijke afwijkingen op, zogenaamde mutanten. Met 'voorwaartse' genetica probeert men de gemuteerde genen te vinden die de afwijkingen in deze mutanten veroorzaken. Op deze manier zijn duizenden mutanten met een abnormale embryonale ontwikkeling gevonden en zijn al honderden genen die hierbij betrokken zijn geïdentificeerd. Tegenwoordig probeert men met behulp van voorwaartse genetica ook genen te vinden die betrokken zijn bij afwijkingen die erg lijken op menselijke ziektes, zoals kanker. 'Omgekeerde' genetica is tegenovergesteld aan voorwaartse genetica. Bij omgekeerde genetica wordt de functie van een gen onderzocht door dit gen specifiek uit te schakelen en vervolgens de gevolgen hiervan te bestuderen. Het maken van deze zogenaamde 'knock-outs' was tot voor kort niet mogelijk in zebravissen.

In dit proefschrift wordt de eerste werkende methode van omgekeerde genetica in de zebravis beschreven (Hoofdstukken 2, 3 en 4). Deze methode hebben we 'doelgerichte mutagenese' genoemd. Bij doelgerichte mutagenese worden genen, net als in voorwaartse genetica, willekeurig uitgeschakeld in enkele duizenden zebravissen. In deze zebravissen wordt vervolgens gericht naar mutaties in een gen gezocht die het gen uitschakelen. Gemiddeld is dit bij één op de duizend zebravissen het geval. Om te laten zien dat deze methode werkt hebben we een *rag1* knock-out zebravis gemaakt (Hoofdstuk 2). In de mens is het *rag1* gen betrokken bij de werking van het immuunsysteem. Een defect immuunsysteem zorgt ervoor dat patiënten geen afweer meer hebben en dus extreem vatbaar zijn voor allerlei infectieuze ziektes. Zoals verwacht hebben de *rag1* knock-out zebravissen ook geen immuunsysteem meer en vormen deze vissen dus een goed model om afweer te bestuderen. De knock-out mutatie in het *rag1* gen is gevonden door de DNA-volgorde van dit gen in ongeveer 2.500 zebravissen met willekeurige mutaties te bepalen. Dit is een tijdrovend en kostbaar proces. Daarom hebben we een alternatieve methode ontwikkeld om vooraf

te kijken of vissen mutaties hebben in een bepaald gen om vervolgens de DNA-volgorde alleen in deze positieve vissen te bepalen (Hoofdstuk 3). Beide methodes hebben hun voor- en nadelen (Hoofdstuk 4). Vanwege technologische vooruitgang in het bepalen van DNA-volgorden wordt de eerste methode nu weer het meest toegepast.

In het algemeen wordt het DNA van genen overgeschreven in boodschapper RNA's, die op hun buurt worden vertaald in eiwitten, de machines van de cel. Recentelijk is er echter een groep van genen ontdekt die coderen voor microRNA's (miRNA's). miRNA's zijn kleine RNA-moleculen die niet worden vertaald in een eiwit, maar op zichzelf een functie hebben. miRNA's reguleren de activiteit van andere genen. miRNA's zijn aanwezig in veel, zoniet alle, meercellige organismen, variërend van plant tot mens. Gewervelde dieren hebben ongeveer vijfhonderd tot duizend verschillende miRNA's. Van het merendeel van deze miRNA's is de functie nog onbekend. De hoeveelheid en de evolutionaire conservatie suggereren echter dat miRNA's belangrijk zijn en betrokken bij allerlei biologische processen, waaronder embryonale ontwikkeling (Hoofdstuk 1).

Om de rol van miRNA's in de embryonale ontwikkeling van gewervelde dieren te bestuderen hebben we door middel van omgekeerde genetica zebravissen gemaakt waarin het miRNA-producerende enzym Dicer is uitgeschakeld (Hoofdstuk 5). Hierdoor worden er geen miRNA's meer aangemaakt. Deze *Dicer* knock-out zebravissen ontwikkelen zich normaal tijdens de vroege embryonale stadia, maar stoppen ongeveer na acht dagen van ontwikkeling, precies op het punt dat er geen nieuwe miRNA's meer worden aangemaakt. Op dit tijdstip zijn de meeste belangrijke organen in de zebravis aangemaakt. Dit suggereert dat miRNA's niet nodig zijn voor de vroegste ontwikkelingsprocessen, maar wel essentieel voor latere stappen van de ontwikkeling en groei van gewervelde dieren.

Om naar de expressie van miRNA's in zebravisembryo's te kijken hebben we een methode ontwikkeld die miRNA's in het weefsel (*in situ*) zichtbaar maakt. Vervolgens hebben we deze methode toegepast om de expressiepatronen van alle miRNA's in verschillende embryonale stadia van de zebravis te bepalen (Hoofdstuk 6). Hieruit blijkt dat de meeste miRNA's niet vroeg in embryonale ontwikkeling tot expressie komen, wanneer het lot van cellen bepaald wordt. Echter, tijdens latere stadia komen miRNA's zeer weefselspecifiek tot expressie. Er zijn bijvoorbeeld miRNA's gevonden die zeer specifiek zijn voor de hersenen, spieren, lever of kraakbeen (zie de foto op de omslag van dit proefschrift). Dit suggereert dat miRNA's betrokken zijn bij de vorming (differentiatie), instandhouding en/of groei van verschillende weefsels. De functie van miRNA's zou dus kunnen zijn dat ze cellen er constant aan herinneren wat voort soort cellen ze zijn.

Dankwoord

Dit proefschrift is het resultaat van het harde werk van vele mensen. Ik bedank iedereen die hier in welke vorm dan ook een bijdrage aan heeft geleverd. Een aantal mensen wil ik in het bijzonder bedanken.

Ronald, bedankt dat ik in jouw groep promotieonderzoek heb mogen doen. Ook al heb je niet zoveel verstand van zebravissen, je hebt mij precies behandeld zoals nodig is om zebravissen snel te laten groeien: flink veel ruimte en voedsel. Oftewel, bedankt dat je me alle vrijheid en middelen hebt gegeven om mezelf in het lab te ontwikkelen.

Al mijn collega's in het Hubrecht lab en met name in de Plasterk groep, bedankt voor jullie hulp, kritische vragen, opmerkingen en suggesties in het lab en tijdens werkbesprekingen. Tevens bedankt voor het creëren van een zeer aangename werksfeer.

Marit, jij hebt je werk op een enorm efficiënte en nauwkeurige manier aangepakt. Hier kunnen veel wetenschappers (waaronder ikzelf) nog wat van leren. Ik heb vooral ontzettend veel lol met je gehad tijdens het maken van de eerste levende knockout bank ter wereld. Ik vond je echt een topper.

Ewart, jij denkt als beste te weten hoe je met de vrouwtjes om moet gaan. Maar met die grote handtasjes act draaf je toch een beetje door! Bedankt voor je enorme inzet, met name voor de vele extra uren die je gemaakt hebt tijdens de microRNA *in situ* screen.

Mijn studenten, Marco, Maarten en Michel, alle drie zijn jullie een nieuw en lastig project begonnen. We hebben heel wat tegenslagen moeten verwerken. Ik hoop dat jullie hier iets van geleerd hebben, ik zeker wel.

Romke, samen zijn we begonnen met de zebravissen in Ronald's groep en we hebben dus veel dingen samen gedaan. Bedankt voor al je hulp en vooral voor de gezelligheid.

Josine, met jou heb ik een haat-liefde verhouding en ben ik altijd aan het bekvechten. Omdat ik nu het laatste woord heb kan ik natuurlijk zeggen: 'gelukkig ben ik nu van je af', maar ik zeg liever dat ik het best wel leuk vond om met je samen te werken.

Zhenya en Victor, bedankt voor al jullie hulp met de bioinformatica. Zonder jullie kennis had ik veel van de experimenten niet zo makkelijk kunnen doen.

Freek en Dana, jullie hebben me bijna alles geleerd over zebravissen. Ik vond het erg leuk om samen urenlang achter de microscoop te zitten en alle structuren en processen in de embryo's te bekijken. Bedankt dat ik altijd even bij jullie kon aankloppen als ik iets over zebravisontwikkeling moest weten of gewoon even moest klagen over de gang van zaken in het aquarium.

Edwin, jij bent vooral in het begin van mijn OIO-tijd voor mij een beetje de lopende 'Maniatis' geweest. De samenwerking die daarna ontstond was erg leuk en vruchtbaar. Samen zijn we met de 'high-throughput' aanpak begonnen, maar we hebben ook nogal wat verschillende gedachten hierover gehad. Gelukkig kwamen we na een korte discussie altijd wel tot dezelfde conclusies: gewoon doen. Verder hebben we

tijdens het borrelen wel meerdere briljante ideeën gehad. Maar, hoeveel kan jij je hier nog van herinneren?

Wigard, stiekem vond ik het best wel leuk om je te horen zeggen dat je onder mijn begeleiding werkte, ook al was dit niet zo en ontkende ik het meteen. Gelukkig heb je nu al meer dan eens bewezen dat je kwaliteiten genoeg hebt om zelfstandig te werken. We gaan hier vast en zeker binnenkort de resultaten van zien! Ik ben wel bang dat je voorlopig nog niet zelfstandig genoeg bent om alleen de kroeg in te gaan...

De vrijdagmiddag borrelclub, met name de harde kern, bedankt voor de gezellige weekafsluiters en voor de vele heftige wetenschappelijke discussies.

Ira, Brandon en René, bedankt voor jullie suggesties en hulp met het afronden van dit boekje.

My foreign collaborators, Stefan, Brigitte, Eric, Ezequiel, Bob and Sakari, thanks for some very fruitful collaborations that contributed to this thesis. I have learned a lot from it.

My friends, thank you for supporting me in my work and for showing me that there is more in life to enjoy.

Tenslotte, mam, pap, hartstikke bedankt dat jullie me altijd en in alles gesteund hebben. Zonder jullie was ik nooit zover gekomen.

Curriculum Vitae

Erno Wienholds werd op 15 juni 1977 geboren in Amsterdam. In 1995 behaalde hij het V.W.O. diploma aan de R.S.G in Enkhuizen. Hierna begon hij aan de studie medische biologie aan de Universiteit van Amsterdam. Voor het doctoraal examen werd een stage gedaan bij de vakgroep Moleculaire Biologie van de Universiteit van Amsterdam, onder begeleiding van M. Siep en Prof. Dr. L.A. Grivell. Tijdens deze stage werd er onderzoek gedaan naar de activatie van mRNA translatie in mitochondriën. Een tweede stage werd gedaan op het Nederlands Kanker Instituut in Amsterdam onder begeleiding van S. Fischer en Prof. Dr. R.H.A. Plasterk. Tijdens deze stage heeft hij onderzoek gedaan naar het verplaatsen van transposons in menselijke cellen en de nematode *C. elegans*. In 2000 werd het doctoraal examen afgerond. Van 2000 tot 2005 heeft hij promotieonderzoek gedaan in het Hubrecht Laboratorium in Utrecht in het kader van de Onderzoekschool Ontwikkelingsbiologie, onder leiding van Prof. Dr. R.H.A. Plasterk. De resultaten van dit onderzoek staan beschreven in dit proefschrift.

List of publications

1. Fischer, S.E., Wienholds, E., and Plasterk, R.H. (2001). Regulated transposition of a fish transposon in the mouse germ line. *Proc Natl Acad Sci U S A* **98**, 6759-6764.

2. Wienholds, E., Schulte-Merker, S., Walderich, B., and Plasterk, R.H. (2002). Target-selected inactivation of the zebrafish rag1 gene. *Science* **297**, 99-102.

3. Fischer, S.E., Wienholds, E., and Plasterk, R.H. (2003). Continuous exchange of sequence information between dispersed Tc1 transposons in the Caenorhabditis elegans genome. *Genetics* **164**, 127-134.

4. Hurlstone, A.F., Haramis, A.P., Wienholds, E., Begthel, H., Korving, J., Van Eeden, F., Cuppen, E., Zivkovic, D., Plasterk, R.H., and Clevers, H. (2003). The Wnt/beta-catenin pathway regulates cardiac valve formation. *Nature* **425**, 633-637.

5. Wienholds, E., van Eeden, F., Kosters, M., Mudde, J., Plasterk, R.H., and Cuppen, E. (2003). Efficient target-selected mutagenesis in zebrafish. *Genome Res* **13**, 2700-2707.

6. Wienholds, E., Koudijs, M.J., van Eeden, F.J., Cuppen, E., and Plasterk, R.H. (2003). The microRNA-producing enzyme Dicer1 is essential for zebrafish development. *Nat Genet* **35**, 217-218.

7. Kloosterman, W.P., Wienholds, E., Ketting, R.F., and Plasterk, R.H. (2004). Substrate requirements for let-7 function in the developing zebrafish embryo. *Nucleic Acids Res* **32**, 6284-6291.

8. Wienholds, E., and Plasterk, R.H. (2004). Target-selected gene inactivation in zebrafish. *Methods Cell Biol* **77**, 69-90.

9. Berezikov, E., Guryev, V., van de Belt, J., Wienholds, E., Plasterk, R.H., and Cuppen, E. (2005). Phylogenetic shadowing and computational identification of human microRNA genes. *Cell* **120**, 21-24.

10. Berghmans, S., Murphey, R.D., Wienholds, E., Neuberg, D., Kutok, J.L., Fletcher, C.D., Morris, J.P., Liu, T.X., Schulte-Merker, S., Kanki, J.P., Plasterk, R., Zon, L.I., and Look, A.T. (2005). tp53 mutant zebrafish develop malignant peripheral nerve sheath tumors. *Proc Natl Acad Sci U S A* **102**, 407-412.

11. Koudijs, M.J., den Broeder, M.J., Keijser, A., Wienholds, E., Houwing, S., van Rooijen, E.M., Geisler, R., and van Eeden, F.J. (2005). The Zebrafish Mutants dre, uki, and lep Encode Negative Regulators of the Hedgehog Signaling Pathway. *PLoS Genet* **1**, e19.

12. Wienholds, E., and Plasterk, R.H.A. (2005). MicroRNA function in animal development. *FEBS Lett.* **579**, in press.

13. Wienholds, E., Kloosterman, W.P., Miska, E., Alvarez-Saavedra, E., Berezikov, E., de Bruijn, E., Horvitz, H.R., Kauppinen, S., and Plasterk, R.H. (2005). MicroRNA expression in zebrafish embryonic development. *Science* **309**, 310-311.