BIOLOGICAL REVIEWS
Cambridge Philosophical Society
Biol. Rev. (2018), **93**, pp. 505–528.
doi: 10.1111/brv.12355

An evolutionary perspective on the systems of adaptive immunity

Viktor Müller[1,2,3,*], Rob J. de Boer[4], Sebastian Bonhoeffer[5] and Eörs Szathmáry[1,2,3,*]

[1] *Parmenides Center for the Conceptual Foundations of Science, 82049 Pullach/Munich, Germany*
[2] *Department of Plant Systematics, Ecology and Theoretical Biology, Institute of Biology, Eötvös Loránd University, 1117 Budapest, Hungary*
[3] *Evolutionary Systems Research Group, MTA Centre for Ecological Research, 8237 Tihany, Hungary*
[4] *Theoretical Biology, Department of Biology, Utrecht University, 3584 CH Utrecht, The Netherlands*
[5] *Institute of Integrative Biology, Department of Environmental Systems Science, ETH Zurich, 8092 Zurich, Switzerland*

ABSTRACT

We propose an evolutionary perspective to classify and characterize the diverse systems of adaptive immunity that have been discovered across all major domains of life. We put forward a new function-based classification according to the way information is acquired by the immune systems: Darwinian immunity (currently known from, but not necessarily limited to, vertebrates) relies on the Darwinian process of clonal selection to 'learn' by cumulative trial-and-error feedback; Lamarckian immunity uses templated targeting (guided adaptation) to internalize heritable information on potential threats; finally, shotgun immunity operates through somatic mechanisms of variable targeting without feedback.

We argue that the origin of Darwinian (but not Lamarckian or shotgun) immunity represents a radical innovation in the evolution of individuality and complexity, and propose to add it to the list of major evolutionary transitions. While transitions to higher-level units entail the suppression of selection at lower levels, Darwinian immunity re-opens cell-level selection within the multicellular organism, under the control of mechanisms that direct, rather than suppress, cell-level evolution for the benefit of the individual. From a conceptual point of view, the origin of Darwinian immunity can be regarded as the most radical transition in the history of life, in which evolution by natural selection has literally re-invented itself. Furthermore, the combination of clonal selection and somatic receptor diversity enabled a transition from limited to practically unlimited capacity to store information about the antigenic environment. The origin of Darwinian immunity therefore comprises both a transition in individuality and the emergence of a new information system – the two hallmarks of major evolutionary transitions.

Finally, we present an evolutionary scenario for the origin of Darwinian immunity in vertebrates. We propose a revival of the concept of the 'Big Bang' of vertebrate immunity, arguing that its origin involved a 'difficult' (i.e. low-probability) evolutionary transition that might have occurred only once, in a common ancestor of all vertebrates. In contrast to the original concept, we argue that the limiting innovation was not the generation of somatic diversity, but the regulatory circuitry needed for the safe operation of amplifiable immune responses with somatically acquired targeting. Regulatory complexity increased abruptly by genomic duplications at the root of the vertebrate lineage, creating a rare opportunity to establish such circuitry. We discuss the selection forces that might have acted at the origin of the transition, and in the subsequent stepwise evolution leading to the modern immune systems of extant vertebrates.

Key words: major evolutionary transition, adaptive immunity, Darwinian immunity, Lamarckian immunity, shotgun immunity, evolutionary scenario.

* Address for correspondence (Tel: (+36) 1 381 2187; Fax (+36) 1 381 2188; E-mail: mueller.viktor@gmail.com; szathmary.eors@gmail.com)

CONTENTS

I. INTRODUCTION

While there is no fundamental law of evolution that would dictate ever-increasing complexity, some lineages of life have clearly become more complex than others, and much of this increase can be attributed to a small number of radical evolutionary innovations (Maynard Smith & Szathmáry, 1995; Szathmáry & Maynard Smith, 1995; Queller, 1997; Szathmáry, 2015). The 'major transitions in evolution' involved transitions in individuality (such as the origin of protocells from populations of freely replicating molecules, the symbiogenetic origin of eukaryotes, or the evolution of multicellularity) and major changes in the way information is stored and transmitted between the generations (e.g. the origin of the genetic code, the epigenetic machinery of cell differentiation, and the origin of language) (Maynard Smith & Szathmáry, 1995; Szathmáry & Maynard Smith, 1995; Szathmáry, 2015). Recurring themes associated with the transitions involve the emergence of new levels of selection and potential conflicts between the levels, novel inheritance systems, possible mechanisms to acquire increasing complexity, and increasing division of labour between the components of the system. Herein we argue that the origins and evolution of adaptive immunity share many of these recurring themes and can therefore be analysed productively in the unified framework of the major evolutionary transitions. In addition, an evolutionary perspective offers a simple conceptual framework to categorize the many systems of immunity that involve adaptive elements across the diversity of life.

II. THE MANY FLAVOURS OF ADAPTIVE IMMUNITY

It is no longer possible to discuss 'adaptive immunity' without adjectives and further specification. While the expression used to be equated with the lymphocyte-based immune system of vertebrates, recent research has revealed a dizzying array of adaptive mechanisms in all major domains of life (Ghosh *et al.*, 2011; Rimer, Cohen & Friedman, 2014), which enable improved defence against pathogens that the host has already been exposed to. With the lines between 'innate' and 'adaptive' immunity becoming increasingly blurred, we need to make a clear distinction between different shades of adaptive immunity. We propose a simple classification based on 'design principles', focusing on the way immune responses acquire new targets during the individual lifespan. As it happens, the more advanced mechanisms of somatic targeting bear resemblance to the different – Darwinian and Lamarckian – modalities of evolution.

(1) Darwinian immunity

The adaptive immune system of vertebrates operates by a *sensu stricto* Darwinian mechanism. It generates, during the lifespan of each individual, a vast repertoire of 'random' receptor specificities, which is then shaped by a series of selection processes. Intricate genetic mechanisms generate variability in the antigen receptor genes during the development of lymphocyte precursors, and mature lymphocytes typically express a single variant of the receptor gene, which is then transmitted to the daughter cells upon

division. Finally, the antigenic specificity of a cell (what molecular patterns it can recognize and react to) affects its chances of survival and proliferation: during maturation, cells are selected for a functional, but not strongly auto-reactive antigen receptor (von Boehmer & Melchers, 2010; Klein *et al.,* 2014), while mature lymphocytes are induced to proliferate if their receptors 'recognize' non-self antigens (under pro-inflammatory conditions). As recognized almost 60 years ago by Burnet, the father of clonal selection theory, these characteristics enable a Darwinian evolutionary process by natural selection, fulfilling the criteria of multiplication and heritable variability in traits that affect survival or reproduction (Burnet, 1957, 1964).

Remarkably, two implementations of clonal selection-based immunity have evolved in the two main extant groups of vertebrates (jawed vertebrates and jawless fish) (Pancer *et al.,* 2004). While the two systems are likely to share deep roots (Kasahara & Sutoh, 2014), they differ in the way they generate variable immune receptors. Jawed vertebrates (gnathostomes) rely on the somatic rearrangement of multiple Variable, Diversity and Joining (V/D/J) gene segments (Oettinger *et al.,* 1990) of the immunoglobulin superfamily [enhanced by additional mechanisms, e.g. non-templated nucleotide addition diversity (Kallenbach *et al.,* 1992) or gene conversion (Reynaud *et al.,* 1987)] to produce variable T-cell and B-cell receptors (TCRs and BCRs, respectively). By contrast, jawless fish (agnathans) generate variable lymphocyte receptors (VLRs) by copying diverse leucin-rich-repeat (LRR) modules using gene conversion (Nagawa *et al.,* 2007).

We propose the term 'Darwinian immunity' to encompass all systems of immunity that rely on Darwinian evolutionary processes within the host, enabled by the combination of somatic receptor diversity and clonal selection. This function-based definition applies to both implementations of vertebrate adaptive immunity, and leaves open the possibility to include potential further instances of adaptive immunity, should independently evolved analogous systems be discovered in the future (e.g. in long-lived large-bodied invertebrates). As a rule, immunity acquired in this Darwinian framework is not readily transmitted between generations.

(2) Lamarckian immunity

A radically different mechanism of adaptive immunity has been discovered in bacteria and archaea: the CRISPR/Cas system (clustered regularly interspaced short palindromic repeats and CRISPR-associated genes) is able to incorporate short nucleotide sequences ('spacers') derived from invading bacteriophages or extragenomic DNA, and then use these to generate complementary 'guides' (short RNA molecules) to target foreign DNA or RNA for degradation in a sequence-specific manner (Mojica *et al.,* 2005; Makarova *et al.,* 2006; Horvath & Barrangou, 2010) (Fig. 1A). These mechanisms resemble Darwinian immunity in the narrow specificity of targeting (recognizing species or even strains, rather than broad classes of pathogens), and in the ability to shape the targeting repertoire during the lifespan of an individual. However, instead of selecting from a pre-existing stochastic repertoire, this system generates targeting motifs in response to the invasion of foreign genetic material, using it directly as a template to synthesize and insert a new 'spacer' unit into the CRISPR locus of the host chromosome (Nunez *et al.,* 2015). The integrated spacers are then typically copied along with the rest of the host chromosome in subsequent cell divisions, enabling the transmission of acquired immunity into the next generations of daughter cells. As Koonin & Wolf (2009, p. 8) pointed out, this type of immunity conforms to the Lamarckian mode of evolution, being 'based on variation directly caused by an environmental cue and resulting in a specific response to that cue', and allowing for the inheritance of acquired characteristics. We propose the term 'Lamarckian immunity' as a function-(i.e. principle-)based nomenclature for all systems of adaptive immunity that operate by templated targeting (i.e. by 'guided learning', as opposed to the 'trial-and-error' learning mechanism of Darwinian immunity). In Frank's (1996) classification, Lamarckian immunity constitutes an 'instructional', while Darwinian immunity a 'selective' system of learning.

Of course, the molecular machinery that enables Lamarckian immunity evolves by Darwinian evolution by natural selection. We also note that while Darwinian immunity involves Darwinian processes within an individual (in populations of cells within a host), Lamarckian immunity involves Lamarckian evolution at the level of populations of hosts.

In a broad sense, the host defence systems based on the large family of Argonaute proteins (Swarts *et al.,* 2014*b*) (Fig. 1B, C) can also be classified as Lamarckian immunity. Argonaute proteins in bacteria and archaea protect against mobile genetic elements through DNA- or RNA-guided DNA interference, or DNA-guided RNA interference (Swarts *et al.,* 2014*b*). They use short small interfering DNA (siDNA) guides (Swarts *et al.,* 2014*a*) or siRNA guides (Olovnikov *et al.,* 2013) acquired from the invading elements to target and cleave complementary DNA (or, in some cases, RNA). This system is analogous, although not homologous, to the CRISPR/Cas machinery, with the exception that the acquired target sequences are not integrated into the host genome, and therefore immunity is not automatically transmitted between the generations; this prompted Koonin & Krupovic (2015, p. 186) to regard it as the 'embodiment of innate immunity'. However, innate immunity is usually meant to encompass responses that rely on germline-encoded targeting, and transgenerational immunity is not a universal characteristic of adaptive immunity. Moreover, in the Argonaute-based prokaryotic system, guide sequences are probably not lost immediately after the elimination of the invading element (Olovnikov *et al.,* 2013), and might well be passed on to the daughter cells upon cell division.

Most eukaryotes have retained Argonaute-based Lamarckian immunity in the form of the virus-derived small

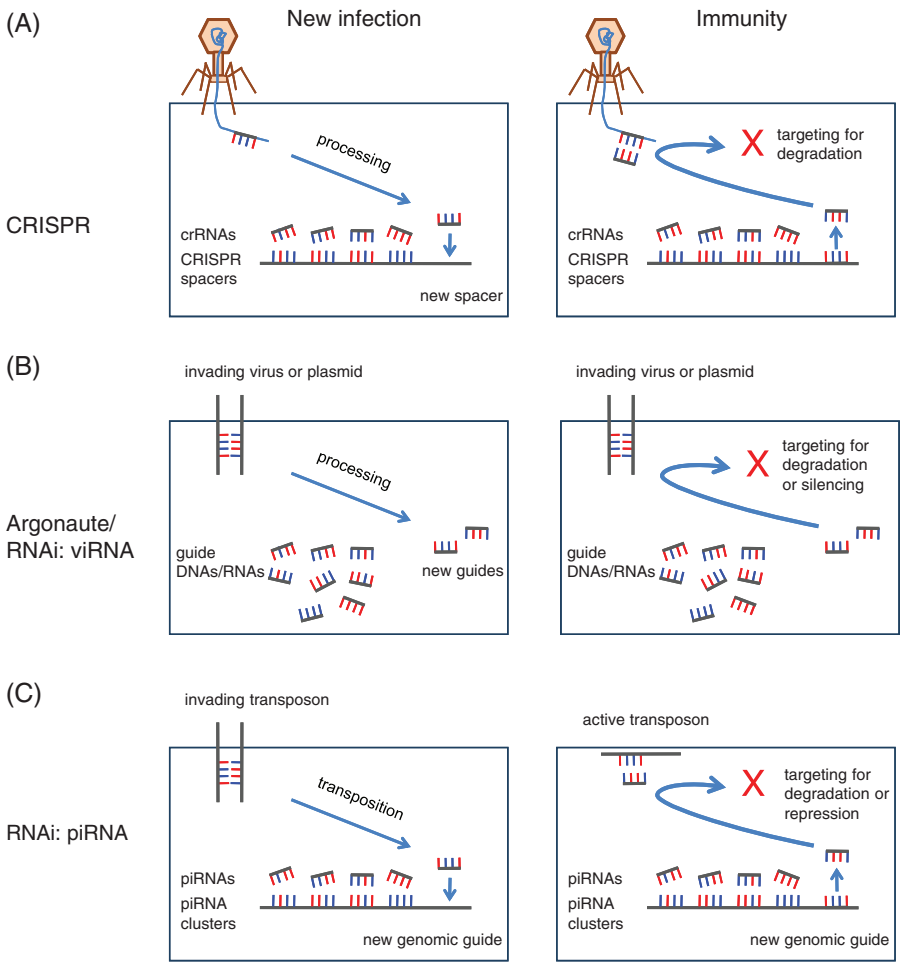

Fig. 1. Template-guided mechanisms of adaptive immune targeting. (A) In the CRISPR/Cas system of bacteria and archaea, the nucleic acid genome of the invading viruses or plasmids is cleaved to generate oligonucleotide 'spacers' that are inserted into the CRISPR locus of the bacterial genome. Upon repeated exposure to the invader, CRISPR-associated RNA (crRNA) transcribed from the spacer binds to the complementary region of the invader genome, targeting it for degradation by a molecular complex. (B) Argonaute-based systems of prokaryotic and eukaryotic DNA or RNA interference (RNAi) acquire specific immunity to genomic invaders by processing the invader genome to yield oligonucleotide guides (DNA in prokaryotes, RNA in eukaryotes) that target complementary DNA or RNA for degradation or silencing. The guides are not integrated into the genome but can, in some eukaryotic species, be amplified by host-encoded replicase enzymes. (C) In some eukaryotes, the Piwi-interacting RNA (piRNA) pathway of RNAi provides genomic defence against transposons: new transposons that transpose into the piRNA cluster of the host genome are transcribed to yield piRNA guides that target transposons to pathways of degradation or repression. Note that the molecular details, processing and effector mechanisms differ between the three systems: we omitted details to emphasize the fundamental similarity of the processes. CRISPR/Cas: clustered regularly interspaced short palindromic repeats and CRISPR-associated genes; viRNA: virus-derived short interfering RNA.

interfering RNA (viRNA) pathway of RNA interference (RNAi) (Ding, 2010). Remarkably, in many eukaryotic species RNA-dependent RNA polymerases (RdRPs) act to amplify siRNAs (Wassenegger & Krczal, 2006), which enables not only systemic spreading but in some cases also transgenerational transmission of specific immunity (Rechavi, Minevich & Hobert, 2011) which would amount to genuine Lamarckian inheritance of an acquired trait. We must note, however, that transgenerational RNAi has not yet been demonstrated in the context of natural exogenous viral infections, and might depend on germline-encoded templates (Ashe *et al.*, 2015).

The genomic defence afforded by the Piwi-interacting RNA (piRNA) guided pathway of RNAi can also be classified as adaptive (genomic) immunity of the Lamarckian kind (Koonin & Wolf, 2009). This pathway relies on germline-encoded piRNA templates to silence mobile genetic elements of complementary nucleic acid sequence in diverse groups of animals (Malone & Hannon, 2009), and the templates are thought to arise by the transposition of the mobile elements into the piRNA cluster (Fig. 1C). Genomic immunity is thus generated by an 'environmental cue' (transposition of the invading element) and this acquired characteristic then becomes heritable across generations.

Remarkably, even some giant viruses might have genomic defences (against virophages) that work on the principles of Lamarckian immunity (Levasseur *et al.*, 2016), although it remains to be seen how the targeting sequence can be dynamically replaced (or added) during the 'lifespan' of virions (or of the virus factories that produce them).

(3) Shotgun immunity

Finally, recent discoveries are beginning to reveal a variety of mechanisms that generate somatically diversified immune receptors or secreted effector molecules in several groups of invertebrates (Du Pasquier, 2006; Ghosh *et al.*, 2011). Similar to Darwinian immunity, somatic diversity in these systems is generated from germline-encoded templates, using various mechanisms of alternative splicing, RNA editing, post-translational modifications and somatic mutations (Du Pasquier, 2006; Ghosh *et al.*, 2011); however, there is no amplification by clonal selection. We propose the term 'shotgun immunity' because the swarm of immune receptors or effectors created by these mechanisms affect not a sharply defined target, but cover a whole region of the possible target range, not unlike pellets fired from a shotgun (Brites & Du Pasquier, 2015).

The best-studied examples involve fibrinogen-related proteins (FREPs) in molluscs (Zhang *et al.*, 2004; Adema, 2015) and the 185/333 family of immune receptor genes in echinoderms (Ghosh *et al.*, 2010). While initial reports implicated shotgun-type immunity also in the context of the somatically diversified Down syndrome cell adhesion molecules (Dscam) of arthropods (Watson *et al.*, 2005), the role of Dscam in immunity has recently been questioned (Brites & Du Pasquier, 2015) and requires further investigation. Plants might also have some form of shotgun immunity: exposure to pathogens increases the frequency of somatic recombination (Lucht *et al.*, 2002), which has been hypothesized to affect pathogen-resistance genes (Loker *et al.*, 2004), possibly generating somatic receptor diversity.

Somatic diversification creates unpredictability in immune targeting in both Darwinian and shotgun immunity, impeding the evasive adaptation of pathogens. The fundamental distinction between the two types is the lack of clonal selection in shotgun immunity. However, theoretical considerations have predicted that somatic diversity can be useful even in the absence of clonal selection (Loker *et al.*, 2004). It is likely that the germline-encoded starting repertoire is being shaped by evolution (between the generations) such that the 'general aim' of the somatically generated swarm would target molecular patterns that are reliably associated with pathogens, but are not found in the host species. Swarm targeting might be particularly useful against pathogenic structures that are themselves variable (Moné *et al.*, 2010).

We emphasize that we do not automatically equate shotgun immunity with invertebrate mechanisms of somatic targeting diversity. Both shotgun immunity and Darwinian immunity are defined on the basis of functional criteria (design principles): should a mechanism of invertebrate immunity turn out to involve clonal selection (i.e. somatically generated receptor diversity that is heritable across cell divisions, and clonal amplification based on receptor binding), then it will have to be re-assigned to the category of Darwinian immunity. The currently estimated distribution of the main types of adaptive immunity across the 'tree of life' is shown in Fig. 2; further occurrences of any type might be discovered in the future.

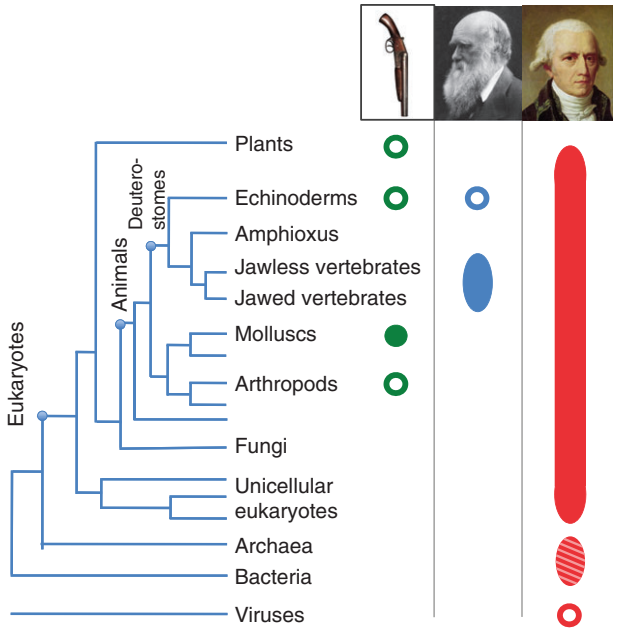

Fig. 2. The distribution of the three main types of adaptive immunity across the tree of life. Lamarckian immunity is ubiquitous in living organisms: both CRISPR/Cas and Argonaute-based systems occur in a diverse array of bacteria and archaea (showing a patchy phylogenetic distribution indicating repeated loss and/or acquisition), RNA interference seems to be an ancestral and common trait in eukaryotes, and even some viruses might have genomic defence based on this principle. Shotgun immunity has been characterized in molluscs, and might also be present in arthropods; further known families of highly diverse invertebrate immune receptors might turn out to belong to this category. The 185/333 family of sea urchin immune receptors either represents shotgun immunity, or possibly a limited form of proto-Darwinian immunity, and increased somatic recombination in plants in response to pathogens might also hint at a mechanism of shotgun immunity. Finally, Darwinian immunity is at present known from jawless and jawed vertebrates. Branch lengths are not drawn to scale; groups of different taxonomic ranks are shown with emphasis on groups that have been better characterized with respect to immunity and are discussed in the main text. Open circles indicate preliminary evidence; stripes indicate a patchy phylogenetic distribution.

Finally, we note that while this conceptual framework captures a fundamental organizing principle of adaptive immunity, aspects of the three main categories can appear in mixed or intermediate forms as well. There appears to be variability as to what degree CRISPR/Cas systems

conform to the Lamarckian mode of evolution [to what degree the acquisition of new spacers is selective towards invading parasites, as opposed to indiscriminate uptake of oligonucleotide templates (Weiss, 2015; Koonin & Wolf, 2016)]; Argonaute-based systems of immunity have varying capacity to transmit acquired information between generations of cells or individuals, while the transfer of maternal antibodies enables the transmission of some acquired information between generations in vertebrates (Lemke, Coutinho & Lange, 2004), and (see Section III.2) it is possible that sea urchins might have some limited form of 'proto-Darwinian' immunity. Such complications are inevitable considering that stepwise evolutionary trajectories exist between rudimentary and full-fledged forms of each type of immunity. The conceptual classification delineates the main underlying principles of the possible modes of acquiring targeting information for immunity, and provides a useful framework for the discussion of the evolutionary significance of each innovation.

III. THE MAKINGS OF A MAJOR EVOLUTIONARY TRANSITION

The paradigm of major evolutionary transitions (METs) posits that the evolution of complexity in the history of life depended on a small number of fundamental changes in the way information is stored and transmitted between generations (Maynard Smith & Szathmáry, 1995; Szathmáry & Maynard Smith, 1995; Szathmáry, 2015). Recurring themes associated with the transitions involve the emergence of new levels of selection and potential conflicts between the levels, novel informational (inheritance) systems, possible mechanisms to acquire increasing complexity, increasing division of labour between the components of the system, and, in some cases, contingent irreversibility. While not every transition possesses all of these features, each of the major transitions created either a new level of selection (transition in individuality) and/or a novel informational system capable of unlimited heredity (in which the number of possible types vastly exceeds the number of individual entities, and stored information is open-ended) (Queller, 1997; Szathmáry, 2015). Below we discuss within this framework why we believe the origin of Darwinian immunity constitutes a major evolutionary transition, while Lamarckian and shotgun immunity do not comply with the criteria.

(1) Darwinian immunity in the MET framework

The clonal selection of lymphocytes is a process of Darwinian evolution that occurs at the level of cells, and creates, as an emergent property, a new information system in the individual. The units of selection are lymphocyte clones bearing immune receptors that are variable among cells and heritable across cell divisions; the mechanisms of selection ensure that the process tends to increase the fitness of the individual. Clonal selection thus creates a new level (a 'controlled arena') of selection within, and for the benefit of, the individual. This is a radical evolutionary innovation. Other major evolutionary transitions (the origins of protocells, eukaryotes, multicellularity, and eusocial animal societies) involved the emergence of a higher level of selection, which favoured the 'de-Darwinization' of the lower-level units (Godfrey-Smith, 2009), i.e. the suppression of evolution at the lower level. By contrast, Darwinian immunity creates a 're-Darwinized' arena of cell-level evolution, both driven and controlled by the complex machinery of clonal selection. Control is a key feature: rather than being a disruptive reversal of multicellularity (as in the case of cancer), this controlled process of cell-level selection tends to confer a benefit at the (higher) level of the individual. The mechanisms of control that have evolved by selection acting at the higher, individual level, have 'domesticated' and harnessed the power of evolution acting at the lower (cellular) level.

The benefit for the individual arises from the complex adaptation of the immune repertoire to the antigenic environment, which also represents the other aspect of the transition: the generation of a new information system. The mechanisms of clonal selection have evolved to amplify immune responses that are useful against dangerous antigens: this ensures adaptation. In turn, the complexity of the adaptation depends on the open-ended nature of the receptor repertoire in vertebrate Darwinian immunity (Jerne, 1985). For example, the number of possible $\alpha\beta$ T-cell receptors has been estimated to be at least 10^{15} in humans (Davis & Bjorkman, 1988) and might be as high as 10^{20} (Zarnitsyna *et al.*, 2013); the alternative Darwinian immune system of jawless fish is estimated to be able to generate around 10^{14}–10^{17} distinct receptors (Pancer *et al.*, 2004). This immense potential diversity constitutes a system of 'unlimited heredity' within the immune system (Maynard Smith & Szathmáry, 1995; Szathmáry, 2015), which is further strengthened by the flexibility of targeting: at least for peptide antigens, all possible peptides of appropriate length are likely to be potential targets. In fact, the open-endedness of the vertebrate immune repertoire applies not only to potential, but also to realized targeting: already in sharks, the antibody repertoire appears to be broad enough to recognize the 'potential universe of antigens' (Adelman, Schluter & Marchalonis, 2004).

Such levels of diversity [e.g. individual humans harbour at least 10^8 distinct $\alpha\beta$ T-cell receptors (Qi *et al.*, 2014)] cannot be encoded in the genome. The number of distinct rearranged lymphocyte receptor genes in one individual can exceed the size of the entire germline genome by several orders of magnitude, while the highest observed diversity of germline-encoded immune receptors (found in invertebrates) does not exceed a few hundred genes (Buckley & Rast, 2015). The somatic generation of variable immune receptors is thus a strict prerequisite for the complexity of vertebrate (and, in general, Darwinian) adaptive immunity. This is an example of 'delegated complexity' (Szathmáry, Jordán & Pál, 2001) [also called 'predicted complexity' in Frank (1996)]:

phenotypic complexity arises not from information directly encoded in the genes but from the interplay of a genetically encoded generative mechanism and a complex environment.

The evolutionary process of Darwinian immunity continues to operate throughout the lifetime of the individual. Shifting exposure to diverse antigenic challenge creates dynamic selection pressure, while variability is replenished by the generation of new naïve cell clones and somatic mutations. In addition to baseline somatic mutations, mammals operate an enhanced mechanism of affinity maturation that allows for repeated rounds of efficient Darwinian selection in B cells (De Silva & Klein, 2015). However, we must emphasize that affinity maturation is not a criterion for Darwinian immunity: it merely enhances the evolutionary process that occurs in the populations of all conventional T and B cells. Furthermore, while Darwinian immunity has been best characterized (and might be most advanced) in mammals, all of the major components appear to be present in all lineages of vertebrates. Teleost fish have been demonstrated to mount both T-cell (Boudinot, Boubekeur & Benmansour, 2001) and B-cell (Castro *et al.*, 2013) responses involving the clonal expansion of antigen specific cells; and both cartilaginous fish (sharks) (Dooley & Flajnik, 2005) and jawless fish (lamprey) (Alder *et al.*, 2008) can be immunized to mount highly specific antibody responses. Functional comparisons have revealed astonishing similarities between the mammalian and the most distantly related (non-homologous) lamprey systems of antibody-based immunity (Altman *et al.*, 2015). Darwinian immunity thus indeed appears to be a general property of vertebrates.

While we have empirical data only on particular examples of vertebrate immunity, we can summarize and generalize the key features of Darwinian immunity at this point. Somatic generation of immune recognition motifs is required to generate a system of practically unlimited (open-ended) information capacity, and clonal selection is required to 'upload' that system with information acquired through continuous rounds of antigenic challenge. The interplay of the two mechanisms enables a radically new way continuously to acquire and store open-ended information about the antigenic environment. In addition, the 're-Darwinization' of lower-level selection in a controlled sub-system of the individual can be regarded as a transition in individuality – this innovation has recently been recognized and classified as a 'filial transition' in the framework of the major evolutionary transitions (Szathmáry, 2015). The origin of Darwinian immunity thus fulfils both criteria of a major evolutionary transition: it created both a new (in this case, embedded) level of selection and a new information system (Table 1); we therefore propose adding it to the list of the major transitions in evolution.

We note that the evolutionary trajectory towards Darwinian immunity encompassed the evolution of evolvability within the immune system: in a sense, evolution by natural selection has 're-invented' itself, taking advantage of its outstanding power to learn (Watson & Szathmáry, 2016). Of some philosophical importance, and in contrast to the evolution of organisms, this embedded system of evolution serves a clear 'purpose' (efficient immunity), conferred by selection acting at a higher level.

At the moment, the Darwinian immunity of vertebrates represents the only biological system with firm evidence for such embedded and controlled evolutionary processes. However, there is some indication that controlled evolutionary dynamics of neural activation patterns might underlie complex problem-solving in the brain (Szilágyi *et al.*, 2016), in which case the evolution of nervous systems capable of Darwinian neurodynamics would constitute another example of 'filial' evolutionary transitions (Szathmáry, 2015). An important implication of this similarity has been recognized by Frank (1996, p. 464) who proposed that 'selection may be the only way to build a complex and meaningful information system from simple rules'.

Finally, we note that in the presentation of the conceptual significance of Darwinian immunity, we have avoided the (sometimes heated) discussion on how much benefit Darwinian immunity provided to vertebrates in the long run (Hedrick, 2009; Pradeu, 2009). Due to the lack of space to address this question properly, we contend that the conceptual novelty of Darwinian immunity and its apparently ubiquitous presence in the highly successful lineage of vertebrates justify its discussion in the framework of the major evolutionary transitions, even without considering issues of improved efficiency.

(2) Other types of adaptive immunity in the framework

How do the other systems of adaptive immunity fit into the paradigm of major evolutionary transitions? Lamarckian immunity operates on an alternative principle of acquiring and storing information about parasites. It certainly constitutes an information system, in some aspects similar to Darwinian immunity. Targeting is open-ended in the sense that, in principle, any nucleotide sequence of appropriate length could be targeted, and the number of possible distinct targets can be astronomical. The typical length of targeting motifs is 32–38, 13–25 and 21–24 nucleotides for CRISPR/Cas (Barrangou & Marraffini, 2014), the prokaryotic Argonaute-based system (Swarts *et al.*, 2014*a*), and RNAi (Ding, 2010), respectively, implying 10^8–10^{23} possible targeting motifs. The number of realized targeting motifs (in any one individual) is typically much lower. In the CRISPR/Cas system, the capacity for information storage is limited to a few hundred (typically <50) targeting motifs (Horvath & Barrangou, 2010), probably due to the rapid erosion of spacers from the genome. Furthermore, most CRISPR targeting motifs are inherited, rather than acquired. By contrast, prokaryotic Argonaute proteins acquire targets from invading plasmids (Swarts *et al.*, 2014*a*), and the actual diversity of targeting motifs has been estimated to be of the order 10^6 in a cell culture (i.e. in a small population) of bacteria (Olovnikov *et al.*, 2013).

In the case of RNAi, the capacity of eukaryotic cells allows the accumulation of a large realized targeting repertoire.

Table 1. Defining traits of a major evolutionary transition in adaptive immunity

	Somatic diversity	Clonal selection (new level of selection)	New information system	Open-ended repertoire
Lamarckian immunity	+	−	−	+*
Shotgun immunity	+	−	−	−
Darwinian immunity	+	+	+	+
Proto-Darwinian	−	+	+/−**	−

*Lamarckian immunity possesses an open-ended repertoire only in the case of RNA interference (RNAi).
**A clonal selection-based information system has strongly limited information capacity in the absence of somatic diversity.

Indeed, viRNAs often provide continuous coverage of every genomic position of the invading viruses (Ding, 2010), presenting a further example of 'delegated complexity'. In addition, the amplification of siRNAs by RdRPs could, in principle, involve multiplication and heritable variation, which would create the potential for a Darwinian selection process. However, to be adaptive for the individual, selection should amplify those targeting motifs that are most effective against the pathogens, and avoid evolution towards autoimmunity or 'selfish' motifs that are simply good at being replicated. While clonal selection in Darwinian immunity relies on complex signalling networks that integrate positive and negative signals, such complex regulation is probably not possible at the level of molecules in RNAi. Instead, RNAi appears to rely on simple mechanisms of self-limitation that prohibit repeated cycles of replication (Bergstrom, McKittrick & Antia, 2003; Pak *et al.*, 2012), but also preclude sustained evolution. Thus, while RNAi involves open-ended delegated complexity, there is no new level of selection, and no new informational system: the acquisition of information relies on the adopted informational system of template (DNA/RNA) copying.

Shotgun immunity represents 'delegated complexity' in that it generates a large array of immune receptors (or effectors) not directly encoded in the genome. However, variability is not combined with heritability and multiplication – the conditions for evolution by natural selection. As a consequence, there is no new level of selection, and no feedback to acquire information about the antigenic environment. The lack of clonal selection also excludes antigen-specific tolerance, which very likely restricts the generative mechanisms of somatic diversity to targeting motifs reliably associated with pathogens and not occurring in the host. Further work is needed to elucidate the targeting specificity of these variable receptors, but the little that we know so far seems to be consistent with this prediction (Ghosh *et al.*, 2011). Targeting in shotgun immunity therefore cannot be open-ended.

Finally, we propose the possibility of a limited form of Darwinian immunity, which would rely on clonal selection, but from a fixed repertoire of germline-encoded receptors, without the generation of somatic diversity. In such a system, if receptors are expressed on cells in a monoallelic manner, and receptor engagement induces proliferation (as in vertebrate Darwinian immunity), then multiplication and heritable variation could enable Darwinian selection, and the repertoire would adapt to antigenic challenge. However, the fixed, limited set of receptor variants would constrain the immune system to a finite, relatively low number of possible states, analogous to the concept of 'limited heredity' (Szathmáry & Maynard Smith, 1995; Szathmáry, 2015). We propose to call such, at present hypothetical, systems of immunity 'proto-Darwinian': without an open-ended information system, they represent a limited form of the (filial) evolutionary transition. Proto-Darwinian systems might have represented earlier stages in the evolution of vertebrate Darwinian immunity, and extant examples might yet be found in some invertebrates as well. The nearly monoallelic expression of single variants from the highly variable 185/333 family of immune-response genes in sea urchin (Majeske *et al.*, 2014) makes this system a promising candidate [possibly even enhanced with some degree of somatic receptor diversity (Ghosh *et al.*, 2010)]. In addition, mammalian natural killer (NK) cells might also have retained or reconstituted this level of complexity. These cells express a stochastically sampled subset of germline-encoded antigen receptors and can establish some degree of antigen-specific immune memory (O'Sullivan, Sun & Lanier, 2015). If antigen-specific NK cell immunity is based on the clonal amplification of cells with a fixed (i.e. heritable) receptor expression profile, then NK cells might represent an extant subsystem of proto-Darwinian immunity within the vertebrate immune system.

To conclude, it is the combination of heritable somatic receptor diversity and clonal selection that creates a new open-ended information system and constitutes a major evolutionary transition. Shotgun and proto-Darwinian immunity lack clonal selection and/or heritable somatic diversity, and are therefore unable to store open-ended information. Lamarckian immunity lacks clonal selection (and a transition in individuality), and while it has the potential for open-ended information capacity (in the form of RNAi), it relies on the pre-existing information system of nucleotide template copying. We summarize these traits in Table 1, while the relative adaptability and the realized repertoire size of the different kinds of immune systems is compared in Fig. 3. Below we discuss how some further recurrent themes of the major evolutionary transitions feature in the systems of adaptive immunity.

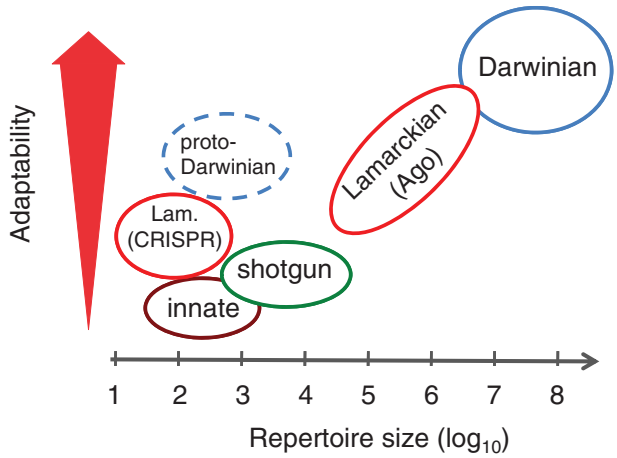

Fig. 3. Realized repertoire size and within-individual (lifespan) adaptability of the different kinds of immunity. Shotgun immunity can generate on the order of 10^4 receptor variants and, given the large number of molecules produced, realized diversity might well approach this magnitude; however, adaptability is restricted to the activation of whole classes of receptors. CRISPR-based systems of prokaryotic Lamarckian (Lam.) immunity operate with tens to hundreds of targeting motifs, some of which are acquired during the lifespan of an individual, which adds a qualitative dimension to adaptability. Argonaute-(Ago-)based Lamarckian immunity ranges from the Piwi-associated targeting of transposons that is rarely updated during the individual lifespan, to the siDNA- or viRNA-guided targeting of whole plasmid or viral genomes, which can utilize millions of overlapping guides that are acquired primarily during the lifespan of an individual. Darwinian immunity uses, in both jawless and jawed vertebrates, 10^7–10^8 receptor variants that are generated during development and are highly inducible. Proto-Darwinian immunity (a hypothetical intermediate stage of the evolutionary transition) involves clonal selection acting on effector cells bearing a limited set of germline-encoded receptors: differential amplification of narrowly targeted effector cell clones can make it highly adaptive; however, new targeting motifs are not generated during the individual lifespan. For comparison, we also show *sensu stricto* innate immune receptors that are coded in the germline: the different types number hundreds or at most a few thousands in one individual, and their adaptability is restricted to the differential activation of broad 'hard-wired' classes of immune responses.

(3) Recurring features of major evolutionary transitions in adaptive immunity

(*a*) *New level of selection: conflict and control*

Major transitions that create a new level of selection open up conflicts if selection favours different traits at the different levels of organisation (Buss, 1987). In Darwinian immunity, multiple mechanisms of control have evolved to modulate cell-level selection for the benefit of the individual. For example, positive and negative selection (von Boehmer & Melchers, 2010; Klein *et al.*, 2014) remove non-functional or potentially self-reactive clones, and the initiation of lymphocyte responses is under the control of innate components of immunity (Iwasaki & Medzhitov, 2015) that can efficiently restrict these responses to 'dangerous' antigens. However, the long-term evolution of lymphocytes can circumvent these 'rules', and give rise to lymphocytes with increased cross-reactivity or auto-reactivity (Deshpande, Parrish & Kuhns, 2015). Therefore, mechanisms of control have evolved also to constrain the scope of 'selfish' evolution: e.g. relying on stem-cell memory T cells (Gattinoni *et al.*, 2011) can limit the accumulation of mutations in memory T-cell clones (Shahriyari & Komarova, 2013; Derényi & Szöllősi, 2017). Finally, we note that, in the form of anti-tumour immunity, the Darwinian immune system has itself assumed a 'policing' function against selfish cellular evolution (malignant transformation) (Michod, 2003; Pradeu, 2013) occurring outside the immune system; an idea that was also first developed by Burnet (1970).

(*b*) *Increasing complexity*

The major evolutionary transitions typically involved one or more of three possible ways of increasing complexity: gene duplication and divergence, the integration of genetic information from independent lineages, and epigenesis. Genomic duplications and horizontal gene transfer laid the foundations for practically all major molecular mechanisms of adaptive immunity [from bacterial systems of Lamarckian immunity (Krupovic *et al.*, 2014; Swarts *et al.*, 2014*b*) to vertebrate Darwinian immunity (Kasahara, 1997; Agrawal, Eastman & Schatz, 1998)], and epigenetic mechanisms enable the astonishing diversity of lymphocyte classes and activation states in vertebrate immunity. The evolution of multiple cell types and effector mechanisms also involved increasing division of labour (specialization and interplay among components), which is a further recurring theme of the major evolutionary transitions. Of particular importance is the division of labour between innate and adaptive components of vertebrate immunity (Iwasaki & Medzhitov, 2015), involving also the transfer of heritable information (recognition of pathogen class) from innate to adaptive immunity (Borghans & De Boer, 2002).

(*c*) *Contingent irreversibility*

The traits emerging from major evolutionary transitions can subsequently be co-opted for new functions, which creates dependencies that act against the reversibility of the transition. Possible examples in adaptive immunity include the acquired role of RNAi in gene regulation and genome maintenance (Castel & Martienssen, 2013), the function of mammalian T cells in wound healing (Havran & Jameson, 2010) and tissue repair (Sadtler *et al.*, 2016), and the regulatory role of major histocompatibility complexes (MHCs) in brain development (Elmer & McAllister, 2012) and their possible co-option for mate choice (Kamiya *et al.*, 2014). These additional functions impose barriers against the loss of these immune components, independent of the original selection forces that created them.

IV. THE ORIGIN OF DARWINIAN IMMUNITY IN VERTEBRATES

Above, we argued that the origin of Darwinian immunity constitutes a major transition in evolution. We now speculate on how it might have happened in the lineage of vertebrates. We propose that the transition occurred only once, before the split between jawed and jawless vertebrates, and explain why we believe that the transition was limited by a difficult evolutionary innovation, rather than the presence or absence of selection pressure for Darwinian immunity. We offer a hypothesis on the nature of the limiting innovation, and outline possible routes of stepwise evolution once the bottleneck had been passed.

(1) A single origin

There are good reasons to believe that the Darwinian immune systems of jawless and jawed vertebrates can be traced back to a common root, and thus that the major transition occurred only once, in a common ancestor of the two lineages. Lampreys have three distinct classes of lymphocytes that provide cellular and humoral immunity, resembling both major lineages of T cells and B cells of jawed vertebrates, respectively (Guo *et al.,* 2009; Hirano *et al.,* 2013). The similarities between not only functions, but also gene expression profiles suggest that the three kinds of lymphocytes are homologous between the two groups and pre-date the divergence of jawed and jawless vertebrates (Flajnik, 2014; Kasahara & Sutoh, 2014). In addition, jawless fish have thymus-like lympho-epithelial structures ('thymoids') that are thought to serve as the sites of lymphocyte development (Bajoghli *et al.,* 2011), and express the lamprey orthologue of the gene encoding forkhead box N1 (Foxn1) transcription factor, a marker of the thymopoietic microenvironment in jawed vertebrates. Finally, receptor diversity in jawless fish is generated by the action of enzymes that are closely related to the gnathostome activation-induced cytosine deaminase (AID) (Rogozin *et al.,* 2007), which is active in the diversification of B-cell receptors. The apparent homology of multiple components of clonal selection-based immunity between jawed and jawless vertebrates strongly suggests that the roots of the system originated in the common ancestor of all vertebrates.

(2) Chance or necessity

The apparently unique origin of Darwinian immunity can be explained in two possible ways. Either, the transition involved a difficult (i.e. low-probability) event that occurred only once, 'by chance', in a common ancestor of all vertebrates [classifying this transition as 'variation-limited' (Számadó & Szathmáry, 2006)]; or, the selective forces that favour the emergence of Darwinian immunity appeared first (and only) in vertebrates and have then driven, 'by necessity', the stepwise evolution of the system [in the frame of a 'selection-limited' transition (Számadó & Szathmáry, 2006)].

Historically, the discovery of the intricate molecular mechanisms of V(D)J recombination (the only mechanism of somatic receptor diversity then known) led researchers to favour the first alternative, assuming a once-only low-probability event for the origin of this system. Marchalonis & Schluter (1990, p. 16) termed this event 'a "Big Bang" because sophisticated rearranging systems consisting of multiple elements appear in a fully functional form without foreshadowing in the antecedent species'. This notion was further strengthened by the recognition that the molecular machinery of V(D)J recombination likely arose by the integration of recombination-activating genes (RAGs) into the vertebrate genome by horizontal gene transfer (Bernstein *et al.,* 1996; Fugmann, 2010). However, subsequent discoveries have challenged the key role of V(D)J recombination in the origin of Darwinian immunity.

First, the discovery of RAG1/2 in sea urchins (Fugmann *et al.,* 2006) suggested that the original horizontal gene transfer event must have preceded the origin of vertebrates. Second, we now know that RAG-mediated V(D)J recombination [enhanced with non-templated nucleotide addition diversity (Kallenbach *et al.,* 1992)] is far from being the only mechanism that can generate somatic receptor diversity. Jawless fish generate receptor diversity by RAG-independent gene conversion (Nagawa *et al.,* 2007), subsets of immunoglobulin (Ig) genes in some jawed vertebrates (sharks, birds, rabbits, sheep) rely heavily on gene conversion and hypermutation to generate antibody diversity (Flajnik & Kasahara, 2010), and invertebrate systems of shotgun immunity generate somatic diversity by gene conversion, alternative splicing, RNA editing, post-translational modifications, and possibly even somatic recombination (Ghosh *et al.,* 2011). Mechanisms of somatic diversity have thus evolved multiple times independently, and are unlikely to be a limiting 'bottleneck' in the evolution of Darwinian immunity. Although it cannot be ruled out that the evolution of V(D)J recombination in particular might have been triggered by a second (intragenomic) transposition event that inserted the RAG transposon into a variable innate immune receptor gene (Koonin & Krupovic, 2015), this can no longer be regarded as 'the Big Bang' of adaptive immunity, but rather as one of several 'smaller bangs' (Bartl *et al.,* 2003; Flajnik, 2014).

Kasahara (1997, 1998) argued that the triggering event of the 'Big Bang' might have been the one or two rounds of whole-genome duplication (WGD) that occurred close to the origin of vertebrates (Smith *et al.,* 2013; Smith & Keinath, 2015). This event duplicated many genes related to immunity, and it 'might have provided unique opportunities to create many accessory and effector molecules of the adaptive immune system' (Kasahara *et al.,* 1997, p. 92). We will return to this idea in Section IV.4, proposing possible scenarios as to how the WGD event might have triggered the origin of Darwinian immunity.

In turn, several studies have argued against the key role of a single triggering event. Klein & Nikolaidis (2005, p. 174) [along the lines of an earlier argument by Bartl *et al.* (2003)]

favour gradual evolution that 'consisted initially of changes unrelated to immune response that were selected to serve other functions' and that, by chance, attained a combination that integrated the elements into a new function giving rise to adaptive immunity. Litman, Rast & Fugmann (2010) also emphasized co-option and redirection of pre-existing systems as the main source of innovation, at the same time perceiving 'no reason to assume that vertebrates require a complex immune system any more than do complex invertebrates' (Litman *et al.,* 2010, p. 552). However, if the origin of Darwinian immunity is not dependent on a 'difficult' (i.e. low-probability) transition, then vertebrates must have some specific traits that favour Darwinian immunity in this group, but are absent from others.

(3) Selective scenarios: not exclusive to vertebrates

Long lifespan (Klein, 1989; Lee, 2006) and slow reproduction (Flajnik, 1998; Lee, 2006; Flajnik & Kasahara, 2010), high metabolic intensity (Rolff, 2007; Sandmeier & Tracy, 2014), efficient closed circulation (van Niekerk, Davis & Engelbrecht, 2015), low population density (Klein, 1989) and large (Klein, 1989; Flajnik & Kasahara, 2010) or morphologically complex (Boehm, 2012) bodies have been invoked as factors favouring (Darwinian) adaptive immunity. However, these traits are not exclusive to vertebrates, and, in fact, the last common chordate ancestor (and therefore also the ancestral vertebrate) was probably a lancelet-(amphioxus-)like creature (Lowe *et al.,* 2015): small, not particularly long-lived, and rather inconspicuous. The most extensive phylogenetic analysis so far estimated that the lineages of jawed vertebrates and jawless fish diverged about 650 million years ago (Blair & Hedges, 2005). While molecular clock estimates might be sensitive to assumptions on the tempo and mode of evolution, fossil evidence of two distinct types of jawless fish dated to around 520 million years ago (Shu *et al.,* 1999) confirms that the split must have occurred before or shortly after the Cambrian Explosion: Darwinian immunity must therefore have provided a selective advantage already in the Precambrian or early Cambrian world of small body sizes and simple body plans.

Many extant invertebrates very likely surpass the last common ancestor of vertebrates in both size and life expectancy, and yet (to our current knowledge) lack Darwinian immunity. Cephalopods can have large bodies and long lifespan, but Darwinian immunity (clonal selection acting on heritable somatic receptor diversity) has not been found in the group (Castellanos-Martínez & Gestal, 2013). It must nonetheless be noted that the species investigated so far have been octopuses that have short lifespans; studies of immunity in *Nautilus* species that can live for several decades (Saunders, 1984) are much awaited.

Some further ancestral traits of vertebrates might also have facilitated or favoured the evolution of Darwinian immunity. A closed circulatory system, which seems to be an ancestral chordate character (Stach, 2008), may well be a prerequisite of effective immune surveillance by lymphocytes; however, cephalopods also have a closed circulation. Filter feeding seems to be an ancestral trait for deuterostomes (Gans & Northcutt, 1983; Yu & Holland, 2009; Lowe *et al.,* 2015), and is present in echinoderms (sea urchins, sea cucumbers), tunicates (sea squirts), and also cephalochordates (amphioxus), which are thought to most closely resemble the common ancestor of vertebrates (Gans & Northcutt, 1983; Yu & Holland, 2009). The evolution of this lifestyle probably generated selection pressure for improved immunity (to fight pathogens, and to avoid unnecessary or harmful responses to the myriad harmless microorganisms in the filtrate). Echinoderms (Hibino *et al.,* 2006), amphioxus (Huang *et al.,* 2008) and, independently, also mussels (Gerdol & Venier, 2015) and sponges (Degnan, 2015), the most ancient group of filter-feeding organisms, took the path of expanding their repertoire of innate pattern-recognition receptors. While expanded innate receptors indeed imply selection pressure for improved immunity, the defining traits of Darwinian immunity have not been found in any of these groups to date.

It has also been noted that vertebrates harbour more complex microbiomes than invertebrates, which tend to have either relatively simple microbial communities or rely on microbial partners that are shielded from immunity within the cells or in compartments enclosed in physical barriers (McFall-Ngai, 2007). Managing a complex microbiome has been invoked as a selection pressure that may have driven the evolution of Darwinian immunity specifically in vertebrates (Pancer & Cooper, 2006; Weaver & Hatton, 2009; Lee & Mazmanian, 2010; Boehm, 2012). However, this explanation only leads one step back, to another question: why would vertebrates be special in terms of needing a complex microbiome? We find it more plausible that Darwinian immunity evolved for another reason (a rare event that opened up a difficult evolutionary path), and could then enable the acquisition of a more complex microbiome – which then might have provided an evolutionary edge to vertebrates.

A further hypothesis has been proposed by Pancer & Cooper (2006, p. 512), who posited that novel selection pressure might have arisen at the origin of vertebrates because a large arsenal of innate receptors 'presented serious autoimmunity problems at a time of rapid developmental and morphologic innovation', and rapid changes in the endosymbiotic communities might also have occurred. As a consequence, the complexity of the innate immune system might have been reduced, creating increased selection pressure for the evolution of an alternative system. However, innate receptors, even those belonging to complex families, tend to target classes of molecules that are not present in the host, and the complexity of the vertebrate body plan increased not so much by expanding the set of molecular building blocks, but rather by regulatory and organizational complexity (Heimberg *et al.,* 2008; Lowe *et al.,* 2011). Such an evolutionary trajectory would not have raised the risk of autoimmunity by innate recognition. It is also unclear why the evolution of vertebrate characteristics would have generated a selection pressure for rapid shifts in the

microbiome, sufficiently strong to compensate for drastically reduced (innate) immune defence against pathogens.

The discoverer of clonal selection, Burnet himself entertained the idea that it might have been the increased developmental flexibility of vertebrates that created the selection pressure for adaptive immunity (Burnet, 1968). He argued that flexible development resulted in an increased risk of cancer, and the threat from the 'modified self' of tumours called for a mechanism that was itself variable and adaptable. However, Darwinian immunity requires reliable mechanisms of immune tolerance to be able to target patterns that are similar to those found in the host self. As we will explain in later sections, it is likely to have started targeting motifs that showed relatively small similarity to host motifs, and could expand to riskier targets only as gradual evolution improved the specificity of targeting and the capacity for antigen-specific tolerance. Distinguishing tumours from normal self is likely to be the most challenging task for Darwinian immunity that could only be added at advanced stages of its evolution – it cannot have been the initial trigger.

Finally, we note that extant vertebrates encompass huge diversity in terms of lifestyles, body size (from shrews to the blue whale) and lifespan (from weeks to >100 years), and while some species have lost or simplified elements of adaptive immunity, the presence of clonal selection-based Darwinian immunity seems ubiquitous across this dizzying diversity of size and form. (Moreover, the species with reduced adaptive immunity do not seem to follow any discernible pattern of size or lifestyle: these examples may simply reflect stochastic loss in some lineages). Considering that most components of vertebrate Darwinian immunity appear to be scalable in terms of diversity, and a higher diversity of innate immune recognition would probably be quite straightforward to re-evolve [indeed Atlantic cod (*Gadus morhua*) have lost MHC class II and have expanded their innate Toll-like receptor (TLR) repertoire (Star *et al.,* 2011)], the ubiquitous maintenance of Darwinian immunity in vertebrates suggests that this type of adaptive immune defence provides benefits across a very wide range of life-history parameters. It is hard to see how this wide range also would not cover the lifestyles of a large number of invertebrate species.

To conclude, while a number of life-history traits likely exerted selection pressure on the ancestral vertebrate to develop sophisticated immunity, and some features of the vertebrate body plan might have acted as necessary pre-adaptations, none of these selection pressures and physical traits seem to be exclusive to this group, and Darwinian immunity would likely be beneficial for many invertebrates as well. We therefore argue that a key piece of the puzzle is still missing: there must have been a difficult evolutionary innovation that emerged, as far as we know, only in vertebrates.

(4) Immunological Big Bang 2.0

What had to be invented for the transition from the invertebrate immunity of an amphioxus-like ancestor to Darwinian vertebrate immunity? The necessary components for the somatic generation of receptor diversity were all in place: amphioxus has RAG1 (Huang *et al.,* 2014; Zhang *et al.,* 2014) and proto-MHC (Abi-Rached *et al.,* 2002); sea urchins have RAG1/2 (Fugmann *et al.,* 2006); and the presence of orthologous ancestral genes in both jawed and jawless vertebrates indicates that the vertebrate ancestor had both BCR/TCR and VLR precursors (Flajnik & Kasahara, 2010). In addition, lymphocyte-like cells have been found in amphioxus (Huang *et al.,* 2007), along with homologues of several genes that are active in immune signalling in the Darwinian immunity of vertebrates (Yu *et al.,* 2005), and recently discovered innate lymphoid cells in mammals perform many functions associated with T cells without expressing T-cell receptors (Walker, Barlow & McKenzie, 2013). These cells can be induced by microbial products, and NK cells that bear germline-encoded antigen receptors (specific, e.g. for conserved structures of viruses) establish immune memory by the survival of an amplified cell population (O'Sullivan *et al.,* 2015), possibly constituting a system of proto-Darwinian immunity. Similar lymphocyte-like cells bearing germline-encoded receptors might have existed in the ancestral vertebrate [innate lymphoid cells might be present in jawless fish, as well (Eberl, Di Santo & Vivier, 2015)], and might already have possessed both the genetic circuitry required for pathogen-induced proliferation and antimicrobial effector mechanisms.

In addition to these pre-existing components, clonal selection-based Darwinian immunity requires two key properties (Du Pasquier, 2006). First, as recognized very early by Burnet (1970), monoallelic (or at most oligoallelic) expression of the somatically generated, clonally heritable antigen receptors is needed to allow for specific amplification (clonal selection) of an appropriate response. Stable expression and clonal heritability are required to maintain targeting specificity over time and across cell divisions; monoallelic expression is necessary to prevent the simultaneous presence of useful and useless or harmful receptors on the same cell, which would greatly abrogate the efficiency of clonal selection. Second, antigen-specific immune tolerance is needed to avoid autoimmunity when a somatically generated receptor responds to a molecular pattern of the host ('self').

We propose that the evolution of antigen-specific immune tolerance is a difficult (low-probability) transition that requires major innovations in gene regulation, and therefore imposes a critical bottleneck in the evolution of Darwinian immunity. We argue that in the evolution of vertebrates this transition was made possible by an abrupt increase in regulatory complexity [precipitated by a WGD event and a series of segmental genome duplications (Smith *et al.,* 2013; Smith & Keinath, 2015)] before the divergence of jawless and jawed vertebrates, and once this difficult transition had been achieved, pre-existing mechanisms of somatic receptor diversity could quickly be co-opted for clonal selection. We term this concept the 'Immunological Big Bang 2.0', and below provide further arguments in its support.

Of the two components of the transition, monoallelic expression of receptor genes does not seem to be particularly difficult to evolve. In addition to the antigen receptors of lymphocytes in jawed and jawless vertebrates (Pancer *et al.,* 2004), monoallelic expression occurs in many mammalian genes not associated with immunity (Nag *et al.,* 2013), while inhibitory receptors on mammalian NK cells (Cichocki, Miller & Anderson, 2011) are characterized by the stochastic expression of a subset of receptor genes from a larger germline-encoded repertoire, and the 185/333 immune-response genes expressed in sea urchin coelomocytes (a type of immune cell) display near-monoallelic expression from a set of germline-encoded alleles (Majeske *et al.,* 2014). However, amplifying lymphocytes with 'random' (i.e. somatically generated) receptors carries the risk of autoimmunity – which brings us to the necessity of antigen-specific tolerance for clonal selection-based Darwinian immunity.

Whereas an autoreactive response without amplification inflicts damage analogous to a fixed dose of a toxic substance, an amplifiable response is analogous to an infectious agent that can multiply and do great harm even at a very low initial dose. As soon as clonal amplification extends to immune recognition motifs that can potentially target self patterns, protective mechanisms are needed to neutralize effector cells based on their self-reactive targeting specificity. Two main mechanisms operate in jawed vertebrates: clonal deletion ('negative selection') removes autoreactive cells during the maturation of lymphocytes (Palmer, 2003) to enable 'recessive tolerance' (tolerance by the absence of autoreactivity); by contrast, regulatory T cells (Tregs) enable 'dominant tolerance' by actively downregulating autoreactive immune responses in the targeted tissues (Coutinho *et al.,* 2001; Sakaguchi, 2004). Both mechanisms are based on intricate gene regulation mechanisms that are likely to be difficult to evolve, and the (near) simultaneous appearance of both systems is highly unlikely. We propose that the 'Big Bang' of vertebrate immunity might have been triggered by the evolution of Treg-mediated dominant tolerance, facilitated by the greatly increased potential for regulatory complexity following the WGD event that gave rise to vertebrates. Below we explain why dominant, rather than recessive tolerance might have been the key innovation, and show that its main genomic components probably originated at or near the WGD event.

We argue that reliable immune tolerance can be achieved by Treg-mediated dominant tolerance, but not by negative selection alone. Both mechanisms are necessarily imperfect (and must have been even less efficient in the beginning), but there is an important difference in the way the two mechanisms can 'fail'. Imperfect negative selection is imperfect in terms of coverage: some auto-reactive clones escape selection; imperfect dominant tolerance is imperfect in terms of degree: all autoimmune reactions are affected, but the degree of control is limited. In the former case, a single escaped clone could wreak havoc without additional control by Tregs in the peripheries, because repeated rounds of clonal expansion would induce exponential growth of the autoimmune reaction. By contrast, imperfect dominant tolerance can afford mistakes, because a self-reactive clone activated by a stochastic glitch in tolerance could still be brought under control later: negative selection has one chance to act, dominant tolerance has many. To suffice alone, negative selection should be perfect; dominant tolerance just needs to be 'good enough' to have a statistically high chance of bringing self-reactive clones under control before they can do too much damage.

We therefore argue (in agreement with Janeway, 2001) that the evolution of regulatory T cells (dominant tolerance) was probably necessary for the emergence of Darwinian immunity. Once dominant tolerance jumpstarted the evolution of Darwinian immunity, the evolution of mechanisms for negative selection against major self-antigens could provide an economical advantage, removing highly autoreactive cells before they had their first chance to expand.

If dominant immune tolerance was a necessary innovation to achieve Darwinian immunity, it certainly cannot have been an easy one. Foxp3 acts as a central switch: it forms complexes with hundreds of genes (Rudra *et al.,* 2012), and affects the expression of more than 2000 genes in mouse T cells (Xie *et al.,* 2015). The task is indeed not trivial. Foxp3+ Tregs often have to respond in the opposite reaction compared with conventional (non-regulatory) effector T cells: TCR signalling (with co-stimulation) induces effector functions in conventional T cells, but repressor functions acting on neighbouring T cells in Tregs. In addition, regulatory activity must strike a delicate balance between too little regulation resulting in runaway autoimmunity, and too much, which could downregulate useful responses against pathogens (self antigens are also presented in the vicinity of pathogen invasion). To achieve this complex functionality, Foxp3 acts not only as a repressor of activation-associated genes, but also upregulates a large number of genes (Zheng *et al.,* 2007), and is likely to operate a bistable autoregulatory loop to maintain a stable identity of regulatory cell clones (Rubtsov *et al.,* 2010). The complexity and difficulty of the task supports the notion that Treg-mediated tolerance might indeed constitute the major bottleneck towards Darwinian immunity that, in vertebrates, could only be passed by a rare burst of regulatory complexity.

Phylogenetic evidence is compatible with the origin of Treg-mediated dominant tolerance in the vertebrate common ancestor. *Foxp3*, the key regulatory gene for the development of regulatory T cells (Hori, Nomura & Sakaguchi, 2003), belongs to the ancient eukaryotic family of Forkhead box (Fox) transcription factors. Remarkably, the Foxp class of the family has a single orthologue in invertebrates (including sea urchin), but four members in most vertebrates (Andersen, Nissen & Betz, 2012), which is consistent with the origin of the class at the WGD event [followed by segmental duplication involving Foxp loci (Smith & Keinath, 2015)]. The analysis of the sea lamprey (*Petromyzon marinus*) genome identified homologues of Foxp1, 2 and 4, but did not find Foxp3 (Smith *et al.,* 2013). However, Foxp3 is most closely related to Foxp4, and both were created

by the last gene duplication in the family (Santos *et al.,* 2011). The Foxp4 ortholog identified in lamprey might therefore be homologous to the common ancestor of Foxp3 and Foxp4 [a situation with known precedents among duplicated transcription factors (Kasahara & Sutoh, 2014)], and might perform the regulatory role of Foxp3 in jawless fish.

While Foxp3 is at the top of the regulatory cascade of dominant tolerance, the evolution of this complex regulatory function likely required the involvement of a whole suite of regulatory genes – which may have depended on the sudden availability of duplicated genes in the ancestral vertebrate. Of note, the transcription factors Helios and GATA-3, which are key interacting partners of Foxp3 in the orchestration of the regulatory phenotype (Rudra *et al.,* 2012; Kim *et al.,* 2015), both belong to gene families that were duplicated in the WGD event (Gillis *et al.,* 2009; John, Yoong & Ward, 2009). Another member of the Foxp class, Foxp1 is involved in the regulation of B- and T-cell development and homeostasis (Hu *et al.,* 2006; Feng *et al.,* 2010), and further classes of duplicated regulatory genes might also have contributed to the expanding genetic circuitry of immune cell fates (Rothenberg & Pant, 2004; John *et al.,* 2009).

In addition to duplicated transcription factors, the increased regulatory complexity of vertebrates arose partly from a massive increase in microRNAs (miRNAs) in the stem lineage of vertebrates (preceding the split between jawless and jawed vertebrates), both due to genome duplication and to the acquisition of new miRNA families (Heimberg *et al.,* 2008, 2010). miRNAs play multiple complex roles in the development and control of vertebrate adaptive immunity (Xiao & Rajewsky, 2009; Mehta & Baltimore, 2016), including mechanisms of both central and peripheral tolerance (reviewed in Simpson & Ansel, 2015). In particular, the selective disruption of miRNAs in Tregs results in autoimmune pathology closely resembling that caused by deficiency in Foxp3 (Zhou *et al.,* 2008), while the selective knockout of miRNAs in thymic epithelial cells compromises promiscuous gene expression (Ucar *et al.,* 2013) that is crucial for the thymic induction of tolerance against peripheral self-antigens. By contrast, V(D)J recombination does not seem to require miRNA control (Xiao & Rajewsky, 2009). Compatible with our scenario, the operation of specific immunological tolerance depends on regulatory complexity acquired at the origin of vertebrates, but the generation of receptor diversity does not.

Thus many components of the genetic circuitry (transcription factors, miRNAs) seem to have appeared in the series of genomic duplications that occurred at the root of the vertebrate lineage. Since duplicated genes tend to get inactivated then lost unless they acquire new functions, the integration of a large number of elements, duplicated within a short time frame, is consistent with a rapid, 'Big Bang' like episode of evolution. Conversely, the construction of the highly complex genetic circuitry of dominant tolerance might have depended on the simultaneous presence of a large number of recently duplicated elements. The analysis of the gene regulatory networks (Martinez-Sanchez *et al.,* 2015) might eventually elucidate how the duplicated regulatory elements might have triggered the evolution of a Treg cell phenotype.

We have thus argued that dominant immune tolerance might be a necessary condition for Darwinian immunity, that the regulatory circuitry required for this function might be very difficult to evolve, and that in vertebrates the origin of the involved genetic machinery apparently goes back to the rare burst of genomic innovation that gave rise to the lineage. Thus, while Burnet (1968) believed that the greater flexibility of development in vertebrates created the selection pressure for adaptive (Darwinian) immunity, we suggest that it created not the need, but the opportunity.

However, we note that the presence of Tregs in jawless fish still needs to be demonstrated, and while it is plausible to assume a crucial role of dominant tolerance in Darwinian immunity, the evidence is not unequivocal. We argued that both somatic receptor diversity and clonal selection might have had pre-existing components, and it was the linking of the two that required a difficult evolutionary innovation: specific (and probably dominant) immune tolerance. However, while the existence of multiple mechanisms of somatic diversity has clearly been demonstrated, clonal selection (amplification) has not been described in any invertebrate to date. It is possible that the machinery for clonal amplification by itself is difficult to evolve, and we cannot exclude that it was this step that imposed a bottleneck for the evolution of Darwinian immunity (L. Du Pasquier, personal communication) that could only be passed by the increased regulatory complexity of vertebrates.

We also note that the 'burst of regulatory complexity' at the root of the lineage is not quite straightforward to explain. WGDs have occurred rarely, but still multiple times in animals, and much more frequently in plants (Otto & Whitton, 2000). However, while some of these events have given rise to successful new clades and/or duplicated regulatory factors, the origin of vertebrates appears to be unique with respect to the number of regulatory elements retained, and the abrupt increase in regulatory complexity and developmental flexibility that accompanied it. It remains to be elucidated what additional factors (selection pressures, pre-adaptations, low-probability genomic events) might have contributed to the rare constellation of conditions that allowed for the rapid increase in regulatory complexity that very likely laid the foundations for the evolutionary success of vertebrates, and opened the trajectory towards Darwinian immunity.

To summarize, the lack of a selective scenario specific to vertebrates argues very strongly for a 'Big Bang'-type origin of Darwinian immunity, limited by a difficult evolutionary innovation; the abrupt increase in regulatory complexity at the origin of vertebrates was very likely a prerequisite (and possible trigger) to passing this bottleneck; and antigen-specific dominant immune tolerance is a plausible (but not the sole) candidate for the limiting evolutionary innovation.

(5) The chicken and egg problem of Darwinian immunity

Beyond the initial bottleneck for Darwinian immunity, an apparent chicken and egg problem arises. Clonal amplification of immune responses with stochastic (somatically diversified) targeting is unsafe without specific (dominant) tolerance; however, specific tolerance might not make much sense without stochastic immune targeting. We argued previously that the emergence of dominant tolerance might have been the key to the evolution of Darwinian immunity in vertebrates – but what drove it in the first place? If specific tolerance evolved against the backdrop of innate or shotgun immunity that did not allow for the clonal amplification of somatically diversified immune responses, what was then the initial selective advantage?

There are two ways to resolve this apparent paradox. First, some limited form of clonal amplification, involving immune responses with a limited scope of diversified targeting, might be beneficial even without specific tolerance, if the benefits of improved defence outweigh the costs (including some limited auto-immunity). Below we shall discuss possible incremental stages in the evolution of randomized immune targeting: it is not impossible that the very first steps could be taken without dominant immune tolerance. In this scenario, a slightly enhanced form of proto-Darwinian immunity (with restricted somatic diversification, and amplification limited in both space and time) might have preceded the emergence of dominant immune tolerance, and mitigating the low-level auto-immunity associated with the former might have provided an immediate evolutionary benefit. If this is true, this level of proto-Darwinian immunity should eventually be found in extant invertebrates [the near-monoclonally expressed 185/333 immune-response genes of sea urchins (Majeske *et al.,* 2014) might constitute a candidate system]. Note that while this scenario somewhat blurs the line between shotgun immunity and Darwinian immunity, a large gap still remains, and bridging that gap very likely required the evolution of specific immune tolerance.

Alternatively, Treg-mediated dominant tolerance might have evolved first to afford specific tolerance to beneficial symbiotic bacteria (Weaver & Hatton, 2009). Innate and shotgun immunity tend to target broad classes of conserved microbial patterns and cannot discriminate and selectively spare potentially beneficial species. By providing this function (downregulating innate mechanisms with narrow targeting), specific tolerance, even in early rudimentary forms, might have provided an immediate benefit even in the absence of somatically diversified immune effector targeting. Remarkably, the gut of extant vertebrates (mice), which holds the largest diversity and biomass of the microbiome, is enriched in Tregs that are reactive to commensal microbes and are essential for the maintenance of immune tolerance against these (Chai, Zhou & Hsieh, 2014; Sefik *et al.,* 2015). Improved microbiome management might afford a huge metabolic benefit (McFall-Ngai, 2007), and is thought to have been a major driver of immune evolution from the earliest animals (Bosch, 2014). In this scenario, Tregs might even have been the first cell type to evolve somatically diversified targeting, which could then be co-opted for effector targeting, as the broadening scope of specific tolerance allowed it. Because the generation of regulatory cells depends on an education period when they encounter antigens under non-inflammatory conditions, the scope of specific tolerance under this scenario could easily be extended to cover self-antigens.

(6) Stepwise evolution of Darwinian immunity after the 'Big Bang'

After the emergence of an early form of specific immune tolerance, the subsequent evolution of vertebrate Darwinian immunity could proceed in small incremental steps, increasing the potential of somatic receptor diversification and clonal amplification to match and drive further the improving capacity of specific tolerance. We consider in turn how the scope and potential of somatic receptor diversification, clonal amplification and specific immune memory, and specific tolerance might have evolved through a series of gradual improvements.

Receptor targeting might have evolved in terms of broadening epitope coverage, shifting from germline-encoded receptors to increasing somatic diversification, and towards higher specificity. Mechanisms of receptor diversification very likely existed even before the immunological 'Big Bang' (Loker *et al.,* 2004), either to generate shotgun immunity [e.g. by somatic hypermutation (Du Pasquier *et al.,* 1998; Lee *et al.,* 2002)] or expressed in the germline to generate variation rapidly across generations. LRR- and RAG/Ig-based systems of gene assembly might also have had their origin at this stage, e.g. the sea urchin homologues of Rag1/2 are expressed in coelomocytes that perform immune functions (Fugmann *et al.,* 2006), and RAG transposition still appears to play a role in generating germline-encoded receptor diversity across generations in sharks (Lee *et al.,* 2000; Hsu *et al.,* 2006). Such pre-existing mechanisms of receptor diversity could then be conveniently co-opted for clonally selected lymphocytes once specific tolerance had appeared. Initially, the germline-encoded receptors must have targeted safe molecular patterns that were reliably associated with potential pathogens but were absent from the host species (Ohno, 1990), and the scope of somatic diversification in the frame of shotgun immunity must have been optimized (limited) to keep the repertoire within these safe boundaries. Then, after the 'Big Bang', the gradual improvements of specific tolerance allowed these safe boundaries to expand, and the mechanisms of somatic receptor diversity had multiple ways to take advantage of this opportunity and expand accordingly (Fig. 4).

First, germline-encoded receptor genes (or their modular components) might have expanded by gene duplication and divergence to allow the targeting of novel domains in the 'epitope space' of possible targeting motifs. Second, the extent of somatic diversification (the possible distance from the germline-encoded target specificities) might also have increased gradually, e.g. by increasing the rate of

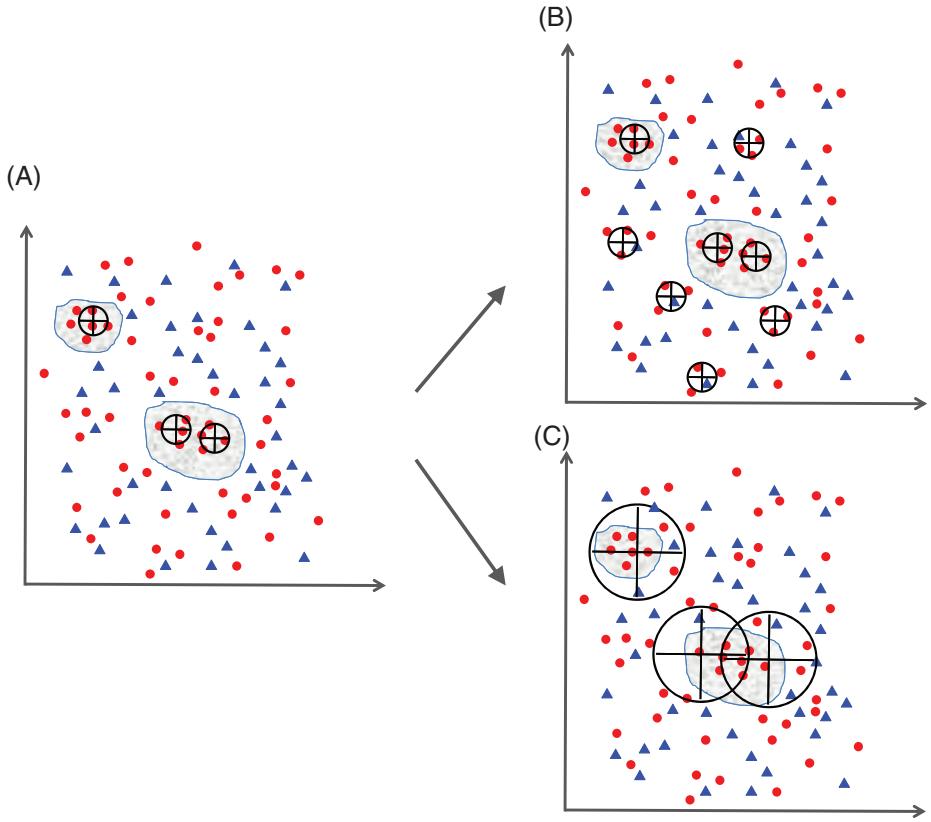

Fig. 4. Two alternative ways to expand epitope coverage. The plot shows a schematic of 'epitope space' (projected into two dimensions for clarity). Red circles indicate immune-targeting motifs associated with potential pathogens, blue triangles denote motifs of the host ('self'); crosshair symbols indicate domains targeted by immune responses. (A) Initially, germline-encoded receptors must have targeted domains of epitope space that are enriched in conserved pathogen-associated patterns, but devoid of host self signatures (textured areas); and the scope of somatic diversification (area under the crosshair) must have been limited to keep the repertoire within these safe boundaries. As gradual improvements in specific tolerance allowed these safe boundaries to expand, receptor genes (or gene segments) might have expanded by duplication and divergence to allow the targeting of novel domains in 'epitope space' (B), and/or the extent of somatic diversification (illustrated by the diameter of crosshairs) could also increase gradually, e.g. by increasing the rate of hypermutation (C). The first mechanism could create new foci of epitope targeting, while the second could increase the action radius of existing foci in the epitope space.

hypermutation or by expanding the genomic regions affected. The first mechanism could create new foci of epitope targeting, while the second could increase the action radius of existing foci in epitope space. Both would allow immune targeting to expand gradually into domains of epitope space that used to carry a high risk of autoimmunity, but were becoming safe due to improving specific tolerance.

Mechanisms based on gene assembly also offer multiple 'scalable' solutions for both aspects of expanding epitope coverage. The number of genomic segments is freely scalable, and the set and probability of possible combinations can also be regulated. For example, Ig genes of cartilaginous fish are still characterized by the (probably) ancestral cluster organization of V(D)J miniloci, which involves very limited numbers of gene segments within each locus, and rearrangements are allowed only within the miniloci (Hsu *et al.*, 2006) (Fig. 5A). This genomic arrangement constrains the possible foci of epitope targeting. By contrast, most Ig genes in tetrapods feature translocon organization, in which multiple gene segments are allowed to recombine, generating much greater combinatorial diversity (Fig. 5B). Remarkably, teleost fish have both cluster and translocon organization in different Ig genes or in different species, underlining the flexibility of gradual evolution towards increasing (or decreasing) combinatorial diversity (Hsu *et al.*, 2006).

Furthermore, even with translocon organization, 'random' somatic recombination does not necessarily imply that all possible combinations (specificities) are produced with the same probability. The generation of V(D)J recombinants can be skewed (Jackson *et al.*, 2013; Elhanati *et al.*, 2014), and some lymphocyte subsets [e.g. several types of unconventional T cells (Godfrey *et al.*, 2015)] are characterized by a highly focused receptor repertoire with limited gene combinations and diversity. The degrees of freedom in combinatorial diversity might have evolved gradually with improving specific tolerance, and if some parts of 'receptor space' were more likely to be useful, regulatory

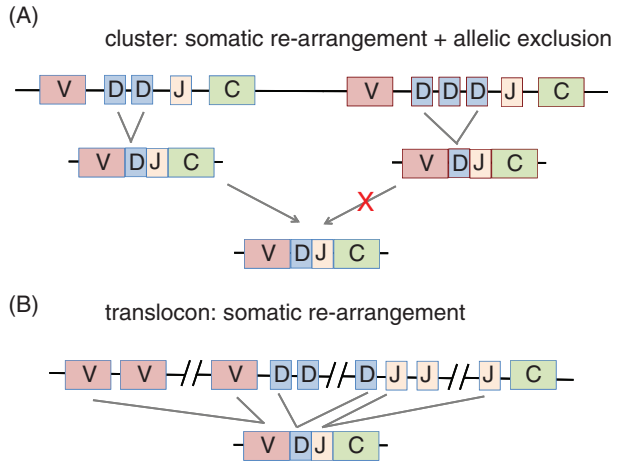

Fig. 5. The schematic representation of cluster and translocon organization in immunoglobulin (Ig) genes. (A) Cluster organization (found in cartilaginous fish and in some genes of bony fish) involves multiple miniloci, each containing limited numbers of variable (V), diversity (D) and joining (J) segments, and a single constant (C) segment: rearrangement is possible only within a minilocus. While several miniloci can be rearranged, most cells seem to express a single transcript (Eason *et al.*, 2004). (B) In the translocon organization characteristic of most tetrapod T- and B-cell receptor genes, a large number of V, D and J segments can be combined freely (although not with equal probability), which yields much greater combinatorial diversity compared with cluster organization.

mechanisms could apparently evolve to ensure the skewed production of these predictably useful specificities.

There is also no reason why additional mechanisms of somatic receptor diversity could not be fine-tuned towards generating broader or more constrained diversity. For example, some vertebrates first generate a limited repertoire relying on somatic recombination only, and switch on the expression of terminal deoxynucleotidyl transferase (TdT; responsible for nucleotide addition diversity) only at later stages of ontogeny (Schwager *et al.*, 1991; Bogue *et al.*, 1992). Knock-out mice lacking TdT display reduced lymphocyte receptor diversity (Gilfillan, Benoist & Mathis, 1995), but are also less prone to autoimmune disease (Conde *et al.*, 1998). The extent of TdT-mediated junctional diversity could probably be flexibly tuned during evolution to match the evolving capabilities of tolerance mechanisms, and the same is likely to be true for the additional mechanisms of hypermutation and gene conversion.

In jawed vertebrates, MHC restriction of adaptive immune responses offers a further scalable solution for the coverage of somatic receptor diversity. Most peptide antigens are able to elicit an immune response only when presented on the surface of a cell bound to an MHC molecule. MHC presentation requires the successive steps of proteasomal cleavage (for class I MHC only), translocation into the lumen of the endoplasmic reticulum (where MHC molecules are loaded), and binding to an MHC molecule. Each of these steps are selective (Hoof *et al.*, 2012), and the degree of selectivity can be fine-tuned by the substrate specificity of cleavage and translocation, and the number and binding specificity of MHC alleles.

The analysis of the highly conserved genome of the elephant shark (*Callorhinchus milii*) (Venkatesh *et al.*, 2014) suggests that MHC alleles were originally in genetic linkage with the genes of the antigen receptors that could bind to them. Such an arrangement might have facilitated the control of the set of peptides involved in MHC presentation, and might also have allowed somatic diversity to get started without thymic positive selection of lymphocytes (because coupled MHC–TCR pairs could be selected for binding over generations).

Tissue-specific restriction offers a further solution to restricting autoimmune collateral damage when specific tolerance is not (yet) efficient. Of note, unconventional T cells tend to recognize antigens in the context of non-polymorphic antigen-presenting molecules, some of which are expressed in a tissue-specific manner (Godfrey *et al.*, 2015).

Clonal amplification and specific memory might also have evolved in incremental steps, contributing to the stepwise co-evolution of the effector and regulatory arms of Darwinian immunity. Clonal amplification can be safe even without specific tolerance for effector cells bearing germline-encoded receptors that are selected for safe targeting across generations (Boehm, 2006), and the genetic circuitry for inducible expansion might have evolved prior to the origins of Darwinian immunity for such cell types (as in NK cells). Then clonal amplification might have been co-opted for cell types using a limited repertoire of somatically diversified receptors [focused on patterns typically associated with pathogens, similar to some classes of unconventional T cells (Godfrey *et al.*, 2015) in extant organisms], and finally also for cells with the broadest diversity of targeting. The evolving genetic circuitry of programmed cell expansion and contraction also incorporated transcription factors that were created in the ancient vertebrate genome duplication event (Rothenberg & Pant, 2004).

In addition to the breadth of targeting involved in clonal amplification, the extent and durability of the amplification could also evolve in gradual steps. In particular, if an immune reaction is short-lived and no memory cells survive, then collateral damage is limited to the time span of the primary immune reaction (launched against an invading pathogen), and this one-time cost might be outweighed by the benefit of efficient defence against the pathogen. That immune effector cells cross-reactive to self would be induced against potentially dangerous non-self antigens, but not to self tissues in the first place, could be ensured by the dependence of clonal amplification on danger signals from the very beginning of Darwinian immunity. The use of danger signals was probably easy to evolve: the new effector mechanisms simply needed to be built on top of the original (innate) decision cascades, co-opting pre-existing inducers of innate immunity as 'danger signals' for evolving Darwinian immunity. This way, self-reactive cells inflicted only limited collateral damage during acute

immune responses, and starting from such a situation, any (initially imperfect) measure of specific tolerance would have been useful and favoured by selection.

The scalability of clonal amplification and specific immune memory can still be observed in the immune systems of extant vertebrates. For example, in sharks 'the memory response is clearly inferior to that of the higher vertebrates' (Flajnik & Kasahara, 2010, p. 50), and repeated challenge with an antigen cannot boost the antibody response beyond the peak of the initial response (Dooley & Flajnik, 2005). Even mammals have several lymphocyte subsets that display limited receptor diversity, tend to target conserved microbial structures, and are able to launch very rapid responses, but generate limited immune memory (Baumgarth, Tung & Herzenberg, 2005; Godfrey *et al.,* 2015). This combination of characteristics may reconstitute (or preserve) the early stages of the evolution of Darwinian immunity, in that restricted somatic diversity and limited memory allow for safe responses without strict check-points (that delay the response of highly diverse classes of lymphocytes) and advanced mechanisms of immune tolerance.

The efficiency of tolerance mechanisms is also likely to have evolved in a stepwise manner. Genome duplication created surplus copies of regulatory factors, but wiring these into a genetic circuitry for regulatory T cells must have taken considerable evolutionary time, and each improvement in regulatory function could further potentiate the evolution of the effector components of Darwinian immunity. Of note, the deletion of Foxp3 in zebrafish results in only a moderate inflammatory phenotype (in contrast to the fatal autoimmune disease observed in Foxp3-deficient mice) (Sugimoto *et al.,* 2017), which is compatible with the view that the capacity of both effector and regulatory immune mechanisms has improved gradually during the evolution of vertebrates.

The action of Treg cells could then also be complemented by the evolution of negative selection, improving not only the reliability, but also the cost efficiency of immune tolerance, by neutralizing autoreactive cells before they had their first chance to expand. In principle, some simple form of negative selection might even have preceded Treg-mediated dominant tolerance in the frame of proto-Darwinian immunity with restricted germline-encoded receptor diversity. In a possible extant analogy, mammalian NK cells go through a period of 'education' early in their development, during which they are able to tune their responsiveness according to the level of inhibitory and stimulatory ligands in their environment (Orr & Lanier, 2010). However, we note that although NK cells are traditionally regarded as components of innate immunity, they are still embedded in the higher regulatory complexity of vertebrates, and it is unclear whether such fine-tuned regulation had been possible before the 'Big Bang' of the WGD event. We cannot rule out that the 'Big Bang' of increasing regulatory complexity opened the way simultaneously to both Treg-mediated dominant tolerance and recessive tolerance by negative selection; remarkably, Foxn1 transcription factor, a marker of the thymopoietic microenvironment, also originated at the WGD event (Singh, Arora & Isambert, 2015). Then, at least in jawed vertebrates, the evolution of the intricate mechanism of promiscuous gene expression in dedicated cells of the thymus (Derbinski *et al.,* 2001) could extend the education of thymocytes (and thereby improve the efficiency of tolerance) to self antigens that are normally restricted to specific tissues. The gene of the transcription factor Aire, the central orchestrator of promiscuous gene expression in the thymus, has been found in the elephant shark (Venkatesh *et al.,* 2014), an ancestral jawed vertebrate, but not yet in lamprey (Smith *et al.,* 2013). Recent studies in mice indicate that promiscuous gene expression promotes the generation of Treg cells involved in dominant tolerance to tissue-specific antigens (Aschenbrenner *et al.,* 2007; Yang *et al.,* 2015); in a recurring theme of immune evolution, new components of immunity tend to evolve interdependencies with pre-existing components.

As the increasing capacity for somatic receptor variability and clonal amplification allowed for increasing repertoire diversity, targeting could also evolve towards higher specificity. This allowed the targeting of variable (not evolutionarily conserved) patterns of potential pathogens, and facilitated the differential recognition of not only self and non-self (Borghans, Noest & De Boer, 1999), but also of distinct pathogens that can be controlled by different effector mechanisms (Borghans & De Boer, 2002).

The evolutionary scenario (including pre-'Big Bang' pre-adaptations and selection pressures) for the evolution of Darwinian immunity in vertebrates is shown in Fig. 6.

(7) Darwinian immunity as a key driver of vertebrate evolution

Finally, we argue that the origin and evolution of Darwinian immunity might have played a crucial role at several stages in the evolution of vertebrates. There are no known vertebrates without Darwinian immunity. Thus, either the innovation was necessary for the subsequent evolution of the vertebrate body plan, or the evolutionary advantage was so large that all other forms without it were outcompeted and went extinct without descendants. The latter possibility becomes highly unlikely once considerable adaptive radiation has occurred, so the emergence of the fundamental framework of vertebrate Darwinian immunity must have happened either shortly after the adaptive radiation of early vertebrates, or even before it, possibly contributing to the evolutionary success of vertebrates.

The situation is somewhat analogous to the origin of eukaryotes and mitochondria. All extant eukaryotes either possess mitochondria or are derived from ancestors that had them. While it is unclear whether it was the acquisition of mitochondria that triggered the burst of evolutionary innovations that led to the last common eukaryotic ancestor (Poole & Gribaldo, 2014), the symbiogenetic event conferred sufficient selective advantage to drive all other protoeukaryotic lineages to extinction. Darwinian immunity evolved along with a whole package of evolutionary

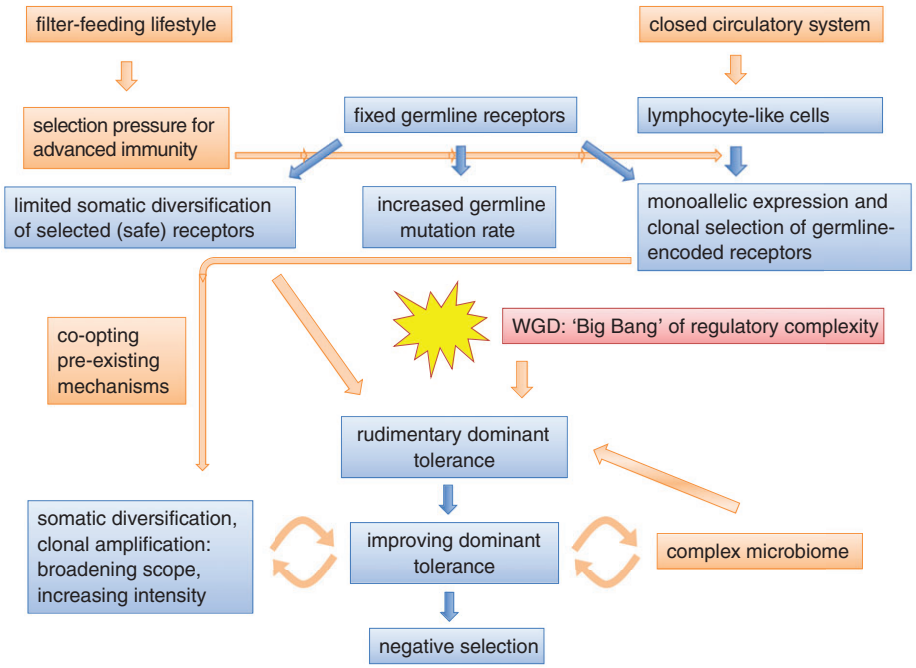

Fig. 6. A proposed scenario for the evolution of Darwinian immunity in vertebrates. Several pre-adaptations have likely paved the way towards Darwinian immunity in the lineage leading to vertebrates. Filter-feeding lifestyle (ancestral in deuterostomes) created selection pressure for immunity with improved (more selective) targeting; closed circulation (ancestral to chordates) allowed for the evolution of lymphocyte-like cells that perform body-wide immune surveillance. Prior to the emergence of specific immune tolerance, improvements in immunity might have involved somatic diversification of selected receptors without feedback and amplification (shotgun immunity), accelerated germline mutations in the receptor genes, and clonal selection limited to a fixed set of safe germline-encoded receptors (proto-Darwinian immunity). Remnants of this stage can be found in extant non-vertebrate deuterostomes (echinoderms, cephalochordates) and in some components of vertebrate immunity. The whole-genome duplication (WGD; followed by segmental duplication of immune-related genes) at the origin of vertebrates initiated a quantum leap in regulatory complexity that allowed, among others, the evolution of regulatory T cells responsible for specific dominant tolerance. Nascent rudimentary forms of dominant tolerance might have provided immediate benefit by mitigating the limited autoimmune damage from shotgun immunity, and possibly by allowing specific tolerance to symbiotic microbes. Dominant tolerance allowed the extension of clonal selection to lymphocytes bearing somatically diversified receptors, and gradual increases in the scope and extent of diversification and amplification were then both driving and allowed by parallel improvements in the tolerance mechanisms in a self-reinforcing cycle (possibly complemented by the benefit of an increasingly complex microbiome). Somatic diversification and clonal amplification might have evolved relatively quickly by co-opting pre-existing molecular mechanisms (although we cannot exclude that efficient clonal selection might also have depended on increased regulatory complexity). Finally, the mechanisms of negative selection might have evolved driven by selection pressure to improve the cost efficiency of specific self-tolerance.

innovations triggered by the WGD event. While the exact contribution of Darwinian immunity to the evolutionary success of vertebrates cannot be directly estimated, the apparent extinction of several intermediate stages of its evolutionary trajectory argues that it must have been a major driver of vertebrate evolution. As argued in previous sections, the selective advantage provided by Darwinian immunity might have included improved cost efficiency of defence against pathogens and/or improved microbiome management.

In addition, the pattern of two alternative implementations of Darwinian immunity in jawless and jawed vertebrates is far from straightforward to explain, and may have further implications for the evolution of vertebrates. Assuming that the evolution of Darwinian immunity was indeed initiated by the establishment of a framework for specific immune tolerance in the common ancestor of all vertebrates, two alternative scenarios can explain the extant pattern of two unrelated implementations of receptor diversity. In the first scenario, one of the two systems (VLR in jawless fish; TCR/BCR in jawed vertebrates) evolved first in the common ancestor of both lineages, but was then replaced by the other system in one of the lineages. It has been speculated that VLR might have evolved first, because all the required genes seem to have been present in the last common vertebrate ancestor, while the horizontal gene transfer that inserted RAG genes into an ancestral TCR/BCR-like gene locus occurred after the split, in the jawed vertebrate lineage (Kato *et al.*, 2012; Kasahara & Sutoh, 2014). Alternatively, the two systems might have arisen independently, each in the common ancestor of one of the lineages, over the background of some form of shotgun and/or proto-Darwinian immunity.

The first scenario (replacement) would imply that the more recent of the two systems had, already in its early

rudimentary form, a selective advantage over the more ancient system, which at that time had already undergone some period of adaptive evolution. If VLR is indeed more ancient (Kato *et al.*, 2012; Kasahara & Sutoh, 2014), then the BCR/TCR system must be more efficient, and it is tempting to speculate that it might have contributed to the much greater evolutionary success of jawed *versus* jawless vertebrates. Under the replacement scenario (irrespective of which system appeared first), the evolution of the second, more powerful system might have been helped by the presence of the tolerance mechanisms that co-evolved with the first system of somatic diversity.

In turn, the alternative scenario of independent origins of both systems from shotgun immunity would imply that two vertebrate lineages that acquired Darwinian immunity remained successful to this day, while all ancestral lineages without it have (apparently) been lost.

Finally, we note that while at the moment, Darwinian immunity is practically synonymous with vertebrate adaptive immunity, independently evolved systems of Darwinian immunity might yet be found in invertebrates. The lessons from vertebrates suggest that higher developmental complexity and, in particular, extensive genome duplications might be prerequisites for the emergence of Darwinian immunity, while filter-feeding and/or reliance on symbiotic microorganisms might give rise to particularly strong selection pressure for improved immunity: invertebrate groups displaying combinations of these traits should be investigated with particular scrutiny. If the last decade has taught us anything, it was that the diversity and ingenuity of invertebrate immune systems is far greater than previously thought: we are certain that the explosive growth of comparative immunology will not fail to deliver further surprises. We list some of the outstanding questions below.

(8) Outstanding questions of Darwinian immunity

(1) What conditions (life-history traits) favour Darwinian immunity over other types of adaptive and innate immunity?
(2) Has Darwinian immunity evolved in any invertebrate taxa? Is the monoallelic expression of variable immune receptors in sea urchins associated with clonal selection?
(3) Is there clonal selection (based on clonally stable receptor identity) in NK cells?
(4) Has Darwinian immunity been lost completely in any vertebrate?
(5) How exactly did the genomic duplication(s) at the origin of vertebrates facilitate the emergence of specific immune tolerance? What drove the exceptional increase in regulatory complexity, in contrast to other genome duplication events?
(6) How do species that appear to have no homologues of Foxp3 [some birds (Andersen *et al.*, 2012); possibly sea lamprey] operate dominant immune tolerance? Do they have a divergent form of the gene (Denyer *et al.*, 2016), or an alternative mechanism has taken over its function?
(7) Is the LRR-based somatic receptor diversity of jawless vertebrates the ancestral vertebrate condition, or did both LRR-based and Ig/RAG-based somatic diversity evolve after the split of jawless and jawed vertebrates?
(8) Is MHC restriction a fortuitous 'complication' in jawed vertebrates, or is this function necessary (inevitable) beyond some level of complexity or potency of Darwinian immunity? In the latter case, are jawless fish below this level, or do they have an analogous system to perform this function?

V. CONCLUSIONS

(1) The various systems of adaptive immunity operate on principles resembling the different modalities of evolution. Template-guided acquisition of heritable targets drives Lamarckian processes in populations of hosts that harbour nucleic acid-based genomic defence, while the clonal selection of somatically diversified lymphocytes in vertebrates relies on Darwinian processes that occur within individual hosts.

(2) The evolution of Darwinian immunity created a controlled arena of cell-level selection within the individual, i.e. selection acting at a higher level opened up, rather than suppressed, selection at a lower level. This is a radical evolutionary innovation, unprecedented in the established examples of major evolutionary transitions. In addition to this embedded transition in individuality, the origin of Darwinian immunity also involved the emergence of a new information system that enabled a transition from limited to practically unlimited capacity to store information about the antigenic environment. We therefore propose adding it to the list of major evolutionary transitions.

(3) So far, Darwinian immunity has been discovered only in vertebrates; however, it is difficult to identify selection pressures that would have favoured Darwinian immunity in vertebrates, but not in any other group. The origin of Darwinian immunity is therefore likely to be constrained by a difficult evolutionary innovation that emerged in vertebrates in a low-probability event, and then triggered a 'Big Bang' of rapid subsequent evolution in adaptive immunity.

(4) Antigen-specific dominant immune tolerance (mediated by regulatory T cells) is a plausible candidate for the tight bottleneck in the evolution of Darwinian immunity. It is likely to be crucial for the safe functioning of Darwinian immunity; it involves a very complicated genetic regulatory circuitry; and many of its components date back to the root of the vertebrate lineage (while most other components of Darwinian immunity apparently pre-date the origin of vertebrates).

(5) Developmental flexibility and complexity is a general hallmark of vertebrates, brought about by a rare burst of genomic duplications at the origin of the group. This abrupt increase in regulatory complexity might have been the direct

trigger that started the evolution of Darwinian immunity: while the role of genomic duplications in vertebrate immune evolution has long been suggested, we here propose that it was the complexity of dominant immune tolerance, in particular, that required the simultaneous availability of a large number of duplicated regulatory elements. The pre-adaptations and selection pressures associated with the lifestyle of ancestral vertebrates might also have contributed to the rare constellation of factors that gave rise to Darwinian immunity (however, these factors are not exclusive to vertebrates).

(6) The ubiquitous presence of Darwinian immunity in vertebrates suggests a major role of this innovation in the evolution of the group. At the minimum, Darwinian immunity must have conferred substantial selection advantage in the competition within the ancestral populations of vertebrates; potentially, it might also have contributed to the general evolutionary success of this lineage.

VI. ACKNOWLEDGEMENTS

The ideas presented herein have greatly benefited from extensive discussions with Louis Du Pasquier; we thank József Burgyán for his advice and feedback on RNAi, and we also thank Martin Flajnik for feedback and insightful comments on an earlier version of the manuscript. We apologize to the authors of all those papers that we were unable to cite due to space constraints from the immense field of comparative immunology. This work was funded by the European Union Seventh Framework Programme (FP7/2007–2013) under ERC grant agreement no. 294332 (Project EvoEvo), and by the grant GINOP-2.3.2-15-2016-00057 (Az evolúció fényében: elvek és megoldások).

We dedicate this paper to Sir Frank Macfarlane Burnet, who not only formulated the theory of clonal selection, but was also the first to recognize the genuinely Darwinian nature of the process.

VII. REFERENCES

Abi-Rached, L., Gilles, A., Shiina, T., Pontarotti, P. & Inoko, H. (2002). Evidence of en bloc duplication in vertebrate genomes. *Nature Genetics* **31**, 100–105.

Adelman, M. K., Schluter, S. F. & Marchalonis, J. J. (2004). The natural antibody repertoire of sharks and humans recognizes the potential universe of antigens. *Protein Jounal* **23**, 103–118.

Adema, C. M. (2015). Fibrinogen-related proteins (FREPs) in mollusks. In *Pathogen-Host Interactions: Antigenic Variation v. Somatic Adaptations*, Volume 57. Results and Problems in Cell Differentiation (eds E. Hsu and L. Du Pasquier), pp. 111–129. Springer International Publishing, Cham.

Agrawal, A., Eastman, Q. M. & Schatz, D. G. (1998). Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature* **394**, 744–751.

Alder, M. N., Herrin, B. R., Sadlonova, A., Stockard, C. R., Grizzle, W. E., Gartland, L. A., Gartland, G. L., Boydston, J. A., Turnbough, C. L. & Cooper, M. D. (2008). Antibody responses of variable lymphocyte receptors in the lamprey. *Nature Immunology* **9**, 319–327.

Altman, M. O., Bennink, J. R., Yewdell, J. W. & Herrin, B. R. (2015). Lamprey VLRB response to influenza virus supports universal rules of immunogenicity and antigenicity. *eLife* **4**, e07467.

Andersen, K. G., Nissen, J. K. & Betz, A. G. (2012). Comparative genomics reveals key gain-of-function events in Foxp3 during regulatory T cell evolution. *Frontiers in Immunology* **3**, 113.

Aschenbrenner, K., D'Cruz, L. M., Vollmann, E. H., Hinterberger, M., Emmerich, J., Swee, L. K., Rolink, A. & Klein, L. (2007). Selection of Foxp3+ regulatory T cells specific for self antigen expressed and presented by Aire+ medullary thymic epithelial cells. *Nature Immunology* **8**, 351–358.

Ashe, A., Sarkies, P., Le Pen, J., Tanguy, M. & Miska, E. A. (2015). Antiviral RNA interference against Orsay virus is neither systemic nor transgenerational in *Caenorhabditis elegans*. *Journal of Virology* **89**, 12035–12046.

Bajoghli, B., Guo, P., Aghaallaei, N., Hirano, M., Strohmeier, C., McCurley, N., Bockman, D. E., Schorpp, M., Cooper, M. D. & Boehm, T. (2011). A thymus candidate in lampreys. *Nature* **470**, 90–94.

Barrangou, R. & Marraffini, L. A. (2014). CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity. *Molecular Cell* **54**, 234–244.

Bartl, S., Baish, M., Weissman, I. L. & Diaz, M. (2003). Did the molecules of adaptive immunity evolve from the innate immune system? *Integrative and Comparative Biology* **43**, 338–346.

Baumgarth, N., Tung, J. & Herzenberg, L. (2005). Inherent specificities in natural antibodies: a key to immune defense against pathogen invasion. *Springer Seminars in Immunopathology* **26**, 347–362.

Bergstrom, C. T., McKittrick, E. & Antia, R. (2003). Mathematical models of RNA silencing: unidirectional amplification limits accidental self-directed reactions. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 11511–11516.

Bernstein, R. M., Schluter, S. F., Bernstein, H. & Marchalonis, J. J. (1996). Primordial emergence of the recombination activating gene 1 (RAG1): sequence of the complete shark gene indicates homology to microbial integrases. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 9454–9459.

Blair, J. E. & Hedges, S. B. (2005). Molecular phylogeny and divergence times of deuterostome animals. *Molecular Biology and Evolution* **22**, 2275–2284.

Boehm, T. (2006). Quality control in self/nonself discrimination. *Cell* **125**, 845–858.

Boehm, T. (2012). Evolution of vertebrate immunity. *Current Biology* **22**, R722–R732.

von Boehmer, H. & Melchers, F. (2010). Checkpoints in lymphocyte development and autoimmune disease. *Nature Immunology* **11**, 14–20.

Bogue, M., Gilfillan, S., Benoist, C. & Mathis, D. (1992). Regulation of N-region diversity in antigen receptors through thymocyte differentiation and thymus ontogeny. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 11011–11015.

Borghans, J. A. & De Boer, R. J. (2002). Memorizing innate instructions requires a sufficiently specific adaptive immune system. *International Immunology* **14**, 525–532.

Borghans, J. A., Noest, A. J. & De Boer, R. J. (1999). How specific should immunological memory be? *The Journal of Immunology* **163**, 569–575.

Bosch, T. C. G. (2014). Rethinking the role of immunity: lessons from Hydra. *Trends in Immunology* **35**, 495–502.

Boudinot, P., Boubekeur, S. & Benmansour, A. (2001). Rhabdovirus infection induces public and private T cell responses in teleost fish. *The Journal of Immunology* **167**, 6202–6209.

Brites, D. & Du Pasquier, L. (2015). Somatic and germline diversification of a putative immunoreceptor within one phylum: dscam in arthropods. In *Pathogen-host Interactions: Antigenic Variation v. Somatic Adaptations*, *Volume 57*. *Results and Problems in Cell Differentiation* (eds E. Hsu and L. Du Pasquier), pp. 131–158. Springer International Publishing, Cham.

Buckley, K. M. & Rast, J. P. (2015). Diversity of animal immune receptors and the origins of recognition complexity in the deuterostomes. *Developmental and Comparative Immunology* **49**, 179–189.

Burnet, F. M. (1957). A modification of Jerne's theory of antibody production using the concept of clonal selection. *The Australian Journal of Science* **20**, 67–69.

Burnet, F. M. (1964). A Darwinian approach to immunity. *Nature* **203**, 451–454.

Burnet, F. M. (1968). Evolution of the immune process in vertebrates. *Nature* **218**, 426–430.

Burnet, F. M. (1970). *Immunological Surveillance*. Pergamon, Oxford.

Buss, L. W. (1987). *The Evolution of Individuality*. Princeton University Press, Princeton.

Castel, S. E. & Martienssen, R. A. (2013). RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nature Reviews Genetics* **14**, 100–112.

Castellanos-Martínez, S. & Gestal, C. (2013). Pathogens and immune response of cephalopods. *Journal of Experimental Marine Biology and Ecology* **447**, 14–22.

Castro, R., Jouneau, L., Pham, H.-P., Bouchez, O., Giudicelli, V., Lefranc, M.-P., Quillet, E., Benmansour, A., Cazals, F., Six, A., Fillatreau, S., Sunyer, O. & Boudinot, P. (2013). Teleost fish mount complex clonal IgM and IgT responses in spleen upon systemic viral infection. *PLoS Pathogens* **9**, e1003098.

Chai, J. N., Zhou, Y. W. & Hsieh, C.-S. (2014). T cells and intestinal commensal bacteria-ignorance, rejection, and acceptance. *FEBS Letters* **588**, 4167–4175.

Cichocki, F., Miller, J. S. & Anderson, S. K. (2011). Killer immunoglobulin-like receptor transcriptional regulation: a fascinating dance of multiple promoters. *Journal of Innate Immunity* **3**, 242–248.

Conde, C., Weller, S., Gilfillan, S., Marcellin, L., Martin, T. & Pasquali, J. L. (1998). Terminal deoxynucleotidyl transferase deficiency reduces the incidence of autoimmune nephritis in (New Zealand Black x New Zealand White)F1 mice. *The Journal of Immunology* **161**, 7023–7030.

Coutinho, A., Hori, S., Carvalho, T., Caramalho, I. & Demengeot, J. (2001). Regulatory T cells: the physiology of autoreactivity in dominant tolerance and "quality control" of immune responses. *Immunological Reviews* **182**, 89–98.

Davis, M. M. & Bjorkman, P. J. (1988). T-cell antigen receptor genes and T-cell recognition. *Nature* **334**, 395–402.

Degnan, S. M. (2015). The surprisingly complex immune gene repertoire of a simple sponge, exemplified by the NLR genes: a capacity for specificity? *Developmental and Comparative Immunology* **48**, 269–274.

Denyer, M. P., Pinheiro, D. Y., Garden, O. A. & Shepherd, A. J. (2016). Missed, not missing: phylogenomic evidence for the existence of avian foxp3. *PLoS One* **11**, e0150988.

Derbinski, J., Schulte, A., Kyewski, B. & Klein, L. (2001). Promiscuous gene expression in medullary thymic epithelial cells mirrors the peripheral self. *Nature Immunology* **2**, 1032–1039.

Derényi, I. & Szöllősi, G. J. (2017). Hierarchical tissue organization as a general mechanism to limit the accumulation of somatic mutations. *Nature Communications* **8**, 14545.

Deshpande, N. R., Parrish, H. L. & Kuhns, M. S. (2015). Self-recognition drives the preferential accumulation of promiscuous CD4+ T-cells in aged mice. *eLife* **4**, e05949.

De Silva, N. S. & Klein, U. (2015). Dynamics of B cells in germinal centres. *Nature Reviews Immunology* **15**, 137–148.

Ding, S.-W. (2010). RNA-based antiviral immunity. *Nature Reviews Immunology* **10**, 632–644.

Dooley, H. & Flajnik, M. F. (2005). Shark immunity bites back: affinity maturation and memory response in the nurse shark, *Ginglymostoma cirratum*. *European Journal of Immunology* **35**, 936–945.

Du Pasquier, L. (2006). Germline and somatic diversification of immune recognition elements in Metazoa. *Immunology Letters* **104**, 2–17.

Du Pasquier, L., Wilson, M., Greenberg, A. S. & Flajnik, M. F. (1998). Somatic mutation in ectothermic vertebrates: musings on selection and origins. *Current Topics in Microbiology and Immunology* **229**, 199–216.

Eason, D. D., Litman, R. T., Luer, C. A., Kerr, W. & Litman, G. W. (2004). Expression of individual immunoglobulin genes occurs in an unusual system consisting of multiple independent loci. *European Journal of Immunology* **34**, 2551–2558.

Eberl, G., Di Santo, J. P. & Vivier, E. (2015). The brave new world of innate lymphoid cells. *Nature Immunology* **16**, 1–5.

Elhanati, Y., Murugan, A., Callan, C. G., Mora, T. & Walczak, A. M. (2014). Quantifying selection in immune receptor repertoires. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 9875–9880.

Elmer, B. M. & McAllister, A. K. (2012). Major histocompatibility complex class I proteins in brain development and plasticity. *Trends in Neurosciences* **35**, 660–670.

Feng, X., Ippolito, G. C., Tian, L., Wiehagen, K., Oh, S., Sambandam, A., Willen, J., Bunte, R. M., Maika, S. D., Harriss, J. V., Caton, A. J., Bhandoola, A., Tucker, P. W. & Hu, H. (2010). Foxp1 is an essential transcriptional regulator for the generation of quiescent naive T cells during thymocyte development. *Blood* **115**, 510–518.

Flajnik, M. F. (1998). Churchill and the immune system of ectothermic vertebrates. *Immunological Reviews* **166**, 5–14.

Flajnik, M. F. (2014). Re-evaluation of the immunological Big Bang. *Current Biology* **24**, R1060–R1065.

Flajnik, M. F. & Kasahara, M. (2010). Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nature Reviews Genetics* **11**, 47–59.

Frank, S. A. (1996). The design of natural and artificial adaptive systems. In *Adaptation* (eds M. R. Rose and G. V. Lauder), pp. 451–505. Academic Press, San Diego.

Fugmann, S. D. (2010). The origins of the Rag genes–from transposition to V(D)J recombination. *Seminars in Immunology* **22**, 10–16.

Fugmann, S. D., Messier, C., Novack, L. A., Cameron, R. A. & Rast, J. P. (2006). An ancient evolutionary origin of the Rag1/2 gene locus. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 3728–3733.

Gans, C. & Northcutt, R. G. (1983). Neural crest and the origin of vertebrates: a new head. *Science* **220**, 268–273.

Gattinoni, L., Lugli, E., Ji, Y., Pos, Z., Paulos, C. M., Quigley, M. F., Almeida, J. R., Gostick, E., Yu, Z., Carpenito, C., Wang, E., Douek, D. C., Price, D. A., June, C. H., Marincola, F. M., Roederer, M. & Restifo, N. P. (2011). A human memory T cell subset with stem cell-like properties. *Nature Medicine* **17**, 1290–1297.

Gerdol, M. & Venier, P. (2015). An updated molecular basis for mussel immunity. *Fish & Shellfish Immunology* **46**, 17–38.

Ghosh, J., Buckley, K. M., Nair, S. V., Raftos, D. A., Miller, C., Majeske, A. J., Hibino, T., Rast, J. P., Roth, M. & Smith, L. C. (2010). Sp185/333: a novel family of genes and proteins involved in the purple sea urchin immune response. *Developmental and Comparative Immunology* **34**, 235–245.

Ghosh, J., Lun, C. M., Majeske, A. J., Sacchi, S., Schrankel, C. S. & Smith, L. C. (2011). Invertebrate immune diversity. *Developmental and Comparative Immunology* **35**, 959–974.

Gilfillan, S., Benoist, C. & Mathis, D. (1995). Mice lacking terminal deoxynucleotidyl transferase: adult mice with a fetal antigen receptor repertoire. *Immunological Reviews* **148**, 201–219.

Gillis, W. Q., St John, J., Bowerman, B. & Schneider, S. Q. (2009). Whole genome duplications and expansion of the vertebrate GATA transcription factor gene family. *BMC Evolutionary Biology* **9**, 207.

Godfrey, D. I., Uldrich, A. P., McCluskey, J., Rossjohn, J. & Moody, D. B. (2015). The burgeoning family of unconventional T cells. *Nature Immunology* **16**, 1114–1123.

Godfrey-Smith, P. (2009). *Darwinian Populations and Natural Selection*. Oxford University Press, Oxford.

Guo, P., Hirano, M., Herrin, B. R., Li, J., Yu, C., Sadlonova, A. & Cooper, M. D. (2009). Dual nature of the adaptive immune system in lampreys. *Nature* **459**, 796–801.

Havran, W. L. & Jameson, J. M. (2010). Epidermal T cells and wound healing. *The Journal of Immunology* **184**, 5423–5428.

Hedrick, S. M. (2009). Immune system: not so superior. *Science* **325**, 1623–1624.

Heimberg, A. M., Cowper-Sal-lari, R., Semon, M., Donoghue, P. C. & Peterson, K. J. (2010). microRNAs reveal the interrelationships of hagfish, lampreys, and gnathostomes and the nature of the ancestral vertebrate. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 19379–19383.

Heimberg, A. M., Sempere, L. F., Moy, V. N., Donoghue, P. C. J. & Peterson, K. J. (2008). MicroRNAs and the advent of vertebrate morphological complexity. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 2946–2950.

Hibino, T., Loza-Coll, M., Messier, C., Majeske, A. J., Cohen, A. H., Terwilliger, D. P., Buckley, K. M., Brockton, V., Nair, S. V., Berney, K., Fugmann, S. D., Anderson, M. K., Pancer, Z., Cameron, R. A., Smith, L. C. & Rast, J. P. (2006). The immune gene repertoire encoded in the purple sea urchin genome. *Developmental Biology* **300**, 349–365.

Hirano, M., Guo, P., McCurley, N., Schorpp, M., Das, S., Boehm, T. & Cooper, M. D. (2013). Evolutionary implications of a third lymphocyte lineage in lampreys. *Nature* **501**, 435–438.

Hoof, I., van Baarle, D., Hildebrand, W. H. & Keşmir, C. (2012). Proteome sampling by the HLA class I antigen processing pathway. *PLoS Computational Biology* **8**, e1002517.

Hori, S., Nomura, T. & Sakaguchi, S. (2003). Control of regulatory T cell development by the transcription factor Foxp3. *Science* **299**, 1057–1061.

Horvath, P. & Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**, 167–170.

Hsu, E., Pulham, N., Rumfelt, L. L. & Flajnik, M. F. (2006). The plasticity of immunoglobulin gene systems in evolution. *Immunological Reviews* **210**, 8–26.

Hu, H., Wang, B., Borde, M., Nardone, J., Maika, S., Allred, L., Tucker, P. W. & Rao, A. (2006). Foxp1 is an essential transcriptional regulator of B cell development. *Nature Immunology* **7**, 819–826.

Huang, G., Xie, X., Han, Y., Fan, L., Chen, J., Mou, C., Guo, L., Liu, H., Zhang, Q., Chen, S., Dong, M., Liu, J. & Xu, A. (2007). The identification of lymphocyte-like cells and lymphoid-related genes in amphioxus indicates the twilight for the emergency of adaptive immune system. *PLoS One* **2**, e206.

Huang, S., Chen, Z., Yan, X., Yu, T., Huang, G., Yan, Q., Pontarotti, P. A., Zhao, H., Li, J., Yang, P., Wang, R., Li, R., Tao, X., Deng, T., Wang, Y., et al. (2014). Decelerated genome evolution in modern vertebrates revealed by analysis of multiple lancelet genomes. *Nature Communications* **5**, 5896.

Huang, S., Yuan, S., Guo, L., Yu, Y., Li, J., Wu, T., Liu, T., Yang, M., Wu, K., Liu, H., Ge, J., Yu, Y., Huang, H., Dong, M., Yu, C., et al. (2008). Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. *Genome Research* **18**, 1112–1126.

Iwasaki, A. & Medzhitov, R. (2015). Control of adaptive immunity by the innate immune system. *Nature Immunology* **16**, 343–353.

Jackson, K. J., Kidd, M. J., Wang, Y. & Collins, A. M. (2013). The shape of the lymphocyte receptor repertoire: lessons from the B cell receptor. *Frontiers in Immunology* **4**, 263.

Janeway, C. A. Jr. (2001). How the immune system works to protect the host from infection: a personal view. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 7461–7468.

Jerne, N. K. (1985). The generative grammar of the immune system. *The EMBO Journal* **4**, 847–852.

John, L. B., Yoong, S. & Ward, A. C. (2009). Evolution of the Ikaros gene family: implications for the origins of adaptive immunity. *The Journal of Immunology* **182**, 4792–4799.

Kallenbach, S., Doyen, N., Fanton d'Andon, M. & Rougeon, F. (1992). Three lymphoid-specific factors account for all junctional diversity characteristic of somatic assembly of T-cell receptor and immunoglobulin genes. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 2799–2803.

Kamiya, T., O'Dwyer, K., Westerdahl, H., Senior, A. & Nakagawa, S. (2014). A quantitative review of MHC-based mating preference: the role of diversity and dissimilarity. *Molecular Ecology* **23**, 5151–5163.

Kasahara, M. (1997). New insights into the genomic organization and origin of the major histocompatibility complex: role of chromosomal (genome) duplication in the emergence of the adaptive immune system. *Hereditas* **127**, 59–65.

Kasahara, M. (1998). What do the paralogous regions in the genome tell us about the origin of the adaptive immune system? *Immunological Reviews* **166**, 159–175.

Kasahara, M., Nakaya, J., Satta, Y. & Takahata, N. (1997). Chromosomal duplication and the emergence of the adaptive immune system. *Trends in Genetics* **13**, 90–92.

Kasahara, M. & Sutoh, Y. (2014). Two forms of adaptive immunity in vertebrates: similarities and differences. *Advances in Immunology* **122**, 59–90.

Kato, L., Stanlie, A., Begum, N. A., Kobayashi, M., Aida, M. & Honjo, T. (2012). An evolutionary view of the mechanism for immune and genome diversity. *The Journal of Immunology* **188**, 3559–3566.

Kim, H.-J., Barnitz, R. A., Kreslavsky, T., Brown, F. D., Moffett, H., Lemieux, M. E., Kaygusuz, Y., Meissner, T., Holderried, T. A. W., Chan, S., Kastner, P., Haining, W. N. & Cantor, H. (2015). Stable inhibitory activity of regulatory T cells requires the transcription factor Helios. *Science* **350**, 334–339.

Klein, J. (1989). Are invertebrates capable of anticipatory immune responses? *Scandinavian Journal of Immunology* **29**, 499–505.

Klein, J. & Nikolaidis, N. (2005). The descent of the antibody-based immune system by gradual evolution. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 169–174.

Klein, L., Kyewski, B., Allen, P. M. & Hogquist, K. A. (2014). Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). *Nature Reviews Immunology* **14**, 377–391.

Koonin, E. V. & Krupovic, M. (2015). Evolution of adaptive immunity from transposable elements combined with innate immune systems. *Nature Reviews Genetics* **16**, 184–192.

Koonin, E. V. & Wolf, Y. I. (2009). Is evolution Darwinian or/and Lamarckian? *Biology Direct* **4**, 42.

Koonin, E. V. & Wolf, Y. I. (2016). Just how Lamarckian is CRISPR-Cas immunity: the continuum of evolvability mechanisms. *Biology Direct* **11**, 1–9.

Krupovic, M., Makarova, K. S., Forterre, P., Prangishvili, D. & Koonin, E. V. (2014). Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity. *BMC Biology* **12**, 36.

Lee, K. A. (2006). Linking immune defenses and life history at the levels of the individual and the species. *Integrative and Comparative Biology* **46**, 1000–1015.

Lee, S. S., Fitch, D., Flajnik, M. F. & Hsu, E. (2000). Rearrangement of immunoglobulin genes in shark germ cells. *The Journal of Experimental Medicine* **191**, 1637–1648.

Lee, S. S., Tranchina, D., Ohta, Y., Flajnik, M. F. & Hsu, E. (2002). Hypermutation in shark immunoglobulin light chain genes results in contiguous substitutions. *Immunity* **16**, 571–582.

Lee, Y. K. & Mazmanian, S. K. (2010). Has the microbiota played a critical role in the evolution of the adaptive immune system? *Science* **330**, 1768–1773.

Lemke, H., Coutinho, A. & Lange, H. (2004). Lamarckian inheritance by somatically acquired maternal IgG phenotypes. *Trends in Immunology* **25**, 180–186.

Levasseur, A., Bekliz, M., Chabriere, E., Pontarotti, P., La Scola, B. & Raoult, D. (2016). MIMIVIRE is a defence system in mimivirus that confers resistance to virophage. *Nature* **531**, 249–252.

Litman, G. W., Rast, J. P. & Fugmann, S. D. (2010). The origins of vertebrate adaptive immunity. *Nature Reviews Immunology* **10**, 543–553.

Loker, E. S., Adema, C. M., Zhang, S. M. & Kepler, T. B. (2004). Invertebrate immune systems–not homogeneous, not simple, not well understood. *Immunological Reviews* **198**, 10–24.

Lowe, C. B., Kellis, M., Siepel, A., Raney, B. J., Clamp, M., Salama, S. R., Kingsley, D. M., Lindblad-Toh, K. & Haussler, D. (2011). Three periods of regulatory innovation during vertebrate evolution. *Science* **333**, 1019–1024.

Lowe, C. J., Clarke, D. N., Medeiros, D. M., Rokhsar, D. S. & Gerhart, J. (2015). The deuterostome context of chordate origins. *Nature* **520**, 456–465.

Lucht, J. M., Mauch-Mani, B., Steiner, H.-Y., Metraux, J.-P., Ryals, J. & Hohn, B. (2002). Pathogen stress increases somatic recombination frequency in Arabidopsis. *Nature Genetics* **30**, 311–314.

Majeske, A. J., Oren, M., Sacchi, S. & Smith, L. C. (2014). Single sea urchin phagocytes express messages of a single sequence from the diverse Sp185/333 gene family in response to bacterial challenge. *The Journal of Immunology* **193**, 5678–5688.

Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. & Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biology Direct* **1**, 7.

Malone, C. D. & Hannon, G. J. (2009). Small RNAs as guardians of the genome. *Cell* **136**, 656–668.

Marchalonis, J. J. & Schluter, S. F. (1990). On the relevance of invertebrate recognition and defence mechanisms to the emergence of the immune response of vertebrates. *Scandinavian Journal of Immunology* **32**, 13–20.

Martinez-Sanchez, M. E., Mendoza, L., Villarreal, C. & Alvarez-Buylla, E. R. (2015). A minimal regulatory network of extrinsic and intrinsic factors recovers observed patterns of CD4+ T cell differentiation and plasticity. *PLoS Computational Biology* **11**, e1004324.

Maynard Smith, J. & Szathmáry, E. (1995). *The Major Transitions in Evolution*. W.H. Freeman, Oxford.

McFall-Ngai, M. (2007). Adaptive immunity – care for the community. *Nature* **445**, 153.

Mehta, A. & Baltimore, D. (2016). MicroRNAs as regulatory elements in immune system logic. *Nature Reviews Immunology* **16**, 279–294.

Michod, R. E. (2003). Cooperation and conflict mediation during the origin of multicellularity. In *Genetic and Cultural Evolution of Cooperation*, (ed. P. Hammerstein), pp. 291–307. MIT Press, Cambridge.

Mojica, F. J. M., Díez-Villaseñor, C. S., García-Martínez, J. & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution* **60**, 174–182.

Moné, Y., Gourbal, B., Duval, D., Du Pasquier, L., Kieffer-Jaquinod, S. & Mitta, G. (2010). A large repertoire of parasite epitopes matched by a large repertoire of host immune receptors in an invertebrate host/parasite model. *PLoS Neglected Tropical Diseases* **4**, e813.

Nag, A., Savova, V., Fung, H. L., Miron, A., Yuan, G. C., Zhang, K. & Gimelbrant, A. A. (2013). Chromatin signature of widespread monoallelic expression. *eLife* **2**, e01256.

Nagawa, F., Kishishita, N., Shimizu, K., Hirose, S., Miyoshi, M., Nezu, J., Nishimura, T., Nishizumi, H., Takahashi, Y., Hashimoto, S., Takeuchi, M., Miyajima, A., Takemori, T., Otsuka, A. J. & Sakano, H. (2007). Antigen-receptor genes of the agnathan lamprey are assembled by a process involving copy choice. *Nature Immunology* **8**, 206–213.

van Niekerk, G., Davis, T. & Engelbrecht, A. M. (2015). Was the evolutionary road towards adaptive immunity paved with endothelium? *Biology Direct* **10**, 47.

Nunez, J. K., Lee, A. S., Engelman, A. & Doudna, J. A. (2015). Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature* **519**, 193–198.

Oettinger, M. A., Schatz, D. G., Gorka, C. & Baltimore, D. (1990). RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science* **248**, 1517–1523.

Ohno, S. (1990). The first set of antigens confronted by the emerging immune system. *Chemical Immunology* **49**, 21–34.

Olovnikov, I., Chan, K., Sachidanandam, R., Newman, D. K. & Aravin, A. A. (2013). Bacterial argonaute samples the transcriptome to identify foreign DNA. *Molecular Cell* **51**, 594–605.

Orr, M. T. & Lanier, L. L. (2010). Natural killer cell education and tolerance. *Cell* **142**, 847–856.

O'Sullivan, T. E., Sun, J. C. & Lanier, L. L. (2015). Natural killer cell memory. *Immunity* **43**, 634–645.

Otto, S. P. & Whitton, J. (2000). Polyploid incidence and evolution. *Annual Review of Genetics* **34**, 401–437.

Pak, J., Maniar, J. M., Mello, C. C. & Fire, A. (2012). Protection from feed-forward amplification in an amplified RNAi mechanism. *Cell* **151**, 885–899.

Palmer, E. (2003). Negative selection–clearing out the bad apples from the T-cell repertoire. *Nature Reviews Immunology* **3**, 383–391.

Pancer, Z., Amemiya, C. T., Ehrhardt, G. R., Ceitlin, J., Gartland, G. L. & Cooper, M. D. (2004). Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. *Nature* **430**, 174–180.

Pancer, Z. & Cooper, M. D. (2006). The evolution of adaptive immunity. *Annual Review of Immunology* **24**, 497–518.

Poole, A. M. & Gribaldo, S. (2014). Eukaryotic origins: how and when was the mitochondrion acquired? *Cold Spring Harbor Perspectives in Biology* **6**, a015990.

Pradeu, T. (2009). Immune system: "Big Bang" in question. *Science* **325**, 393.

Pradeu, T. (2013). Immunity and the emergence of individuality. In *From Groups to Individuals: Evolution and Emerging Individuality* (eds F. Bouchard and P. Huneman), pp. 77–96. MIT Press, Cambridge.

Qi, Q., Liu, Y., Cheng, Y., Glanville, J., Zhang, D., Lee, J.-Y., Olshen, R. A., Weyand, C. M., Boyd, S. D. & Goronzy, J. J. (2014). Diversity and clonal selection in the human T-cell repertoire. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 13139–13144.

Queller, D. C. (1997). Cooperators since life began. *The Quarterly Review of Biology* **72**, 184–188.

Rechavi, O., Minevich, G. & Hobert, O. (2011). Transgenerational inheritance of an acquired small RNA-based antiviral response in *C. elegans*. *Cell* **147**, 1248–1256.

Reynaud, C.-A., Anquez, V., Grimal, H. & Weill, J.-C. (1987). A hyperconversion mechanism generates the chicken light chain preimmune repertoire. *Cell* **48**, 379–388.

Rimer, J., Cohen, I. R. & Friedman, N. (2014). Do all creatures possess an acquired immune system of some sort? *BioEssays* **36**, 273–281.

Rogozin, I. B., Iyer, L. M., Liang, L., Glazko, G. V., Liston, V. G., Pavlov, Y. I., Aravind, L. & Pancer, Z. (2007). Evolution and diversification of lamprey antigen receptors: evidence for involvement of an AID-APOBEC family cytosine deaminase. *Nature Immunology* **8**, 647–656.

Rolff, J. (2007). Why did the acquired immune system of vertebrates evolve? *Developmental & Comparative Immunology* **31**, 476–482.

Rothenberg, E. V. & Pant, R. (2004). Origins of lymphocyte developmental programs: transcription factor evidence. *Seminars in Immunology* **16**, 227–238.

Rubtsov, Y. P., Niec, R. E., Josefowicz, S., Li, L., Darce, J., Mathis, D., Benoist, C. & Rudensky, A. Y. (2010). Stability of the regulatory T cell lineage in vivo. *Science* **329**, 1667–1671.

Rudra, D., deRoos, P., Chaudhry, A., Niec, R. E., Arvey, A., Samstein, R. M., Leslie, C., Shaffer, S. A., Goodlett, D. R. & Rudensky, A. Y. (2012). Transcription factor Foxp3 and its protein partners form a complex regulatory network. *Nature Immunology* **13**, 1010–1019.

Sadtler, K., Estrellas, K., Allen, B. W., Wolf, M. T., Fan, H., Tam, A. J., Patel, C. H., Luber, B. S., Wang, H., Wagner, K. R., Powell, J. D., Housseau, F., Pardoll, D. M. & Elisseeff, J. H. (2016). Developing a pro-regenerative biomaterial scaffold microenvironment requires T helper 2 cells. *Science* **352**, 366–370.

Sakaguchi, S. (2004). Naturally arising CD4+ regulatory T cells for immunologic self-tolerance and negative control of immune responses. *Annual Review of Immunology* **22**, 531–562.

Sandmeier, F. C. & Tracy, R. C. (2014). The metabolic pace-of-life model: incorporating ectothermic organisms into the theory of vertebrate ecoimmunology. *Integrative and Comparative Biology* **54**, 387–395.

Santos, M. E., Athanasiadis, A., Leitao, A. B., Du Pasquier, L. & Sucena, E. (2011). Alternative splicing and gene duplication in the evolution of the FoxP gene subfamily. *Molecular Biology and Evolution* **28**, 237–247.

Saunders, W. B. (1984). Nautilus growth and longevity: evidence from marked and recaptured animals. *Science* **224**, 990–992.

Schwager, J., Burckert, N., Courtet, M. & Du Pasquier, L. (1991). The ontogeny of diversification at the immunoglobulin heavy chain locus in Xenopus. *The EMBO Journal* **10**, 2461–2470.

Sefik, E., Geva-Zatorsky, N., Oh, S., Konnikova, L., Zemmour, D., McGuire, A. M., Burzyn, D., Ortiz-Lopez, A., Lobera, M., Yang, J., Ghosh, S., Earl, A., Snapper, S. B., Jupp, R., Kasper, D., et al. (2015). Individual intestinal symbionts induce a distinct population of RORgamma(+) regulatory T cells. *Science* **349**, 993–997.

Shahriyari, L. & Komarova, N. L. (2013). Symmetric vs. asymmetric stem cell divisions: an adaptation against cancer? *PLoS One* **8**, e76195.

Shu, D. G., Luo, H. L., Conway Morris, S., Zhang, X. L., Hu, S. X., Chen, L., Han, J., Zhu, M., Li, Y. & Chen, L. Z. (1999). Lower Cambrian vertebrates from south China. *Nature* **402**, 42–46.

Simpson, L. J. & Ansel, K. M. (2015). MicroRNA regulation of lymphocyte tolerance and autoimmunity. *The Journal of Clinical Investigation* **125**, 2242–2249.

Singh, P. P., Arora, J. & Isambert, H. (2015). Identification of ohnolog genes originating from whole genome duplication in early vertebrates, based on synteny comparison across multiple genomes. *PLoS Computational Biology* **11**, e1004394.

Smith, J. J. & Keinath, M. C. (2015). The sea lamprey meiotic map improves resolution of ancient vertebrate genome duplications. *Genome Research* **25**, 1081–1090.

Smith, J. J., Kuraku, S., Holt, C., Sauka-Spengler, T., Jiang, N., Campbell, M. S., Yandell, M. D., Manousaki, T., Meyer, A., Bloom, O. E., Morgan, J. R., Buxbaum, J. D., Sachidanandam, R., Sims, C., Garruss, A. S., et al. (2013). Sequencing of the sea lamprey (*Petromyzon marinus*) genome provides insights into vertebrate evolution. *Nature Genetics* **45**, 415–421.

Stach, T. (2008). Chordate phylogeny and evolution: a not so simple three-taxon problem. *Journal of Zoology* **276**, 117–141.

Star, B., Nederbragt, A. J., Jentoft, S., Grimholt, U., Malmstrom, M., Gregers, T. F., Rounge, T. B., Paulsen, J., Solbakken, M. H., Sharma, A., Wetten, O. F., Lanzen, A., Winer, R., Knight, J., Vogel, J. H., et al. (2011). The genome sequence of Atlantic cod reveals a unique immune system. *Nature* **477**, 207–210.

Sugimoto, K., Hui, S. P., Sheng, D. Z., Nakayama, M. & Kikuchi, K. (2017). Zebrafish FOXP3 is required for the maintenance of immune tolerance. *Developmental & Comparative Immunology* **73**, 156–162.

Swarts, D. C., Jore, M. M., Westra, E. R., Zhu, Y., Janssen, J. H., Snijders, A. P., Wang, Y., Patel, D. J., Berenguer, J., Brouns, S. J. & van der Oost, J. (2014*a*). DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* **507**, 258–261.

Swarts, D. C., Makarova, K., Wang, Y., Nakanishi, K., Ketting, R. F., Koonin, E. V., Patel, D. J. & van der Oost, J. (2014*b*). The evolutionary journey of Argonaute proteins. *Nature Structural and Molecular Biology* **21**, 743–753.

Számadó, S. & Szathmáry, E. (2006). Selective scenarios for the emergence of natural language. *Trends in Ecology & Evolution* **21**, 555–561.

Szathmáry, E. (2015). Toward major evolutionary transitions theory 2.0. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 10104–10111.

Szathmáry, E., Jordán, F. & Pál, C. (2001). Molecular biology and evolution. Can genes explain biological complexity? *Science* **292**, 1315–1316.

Szathmáry, E. & Maynard Smith, J. (1995). The major evolutionary transitions. *Nature* **374**, 227–232.

Szilágyi, A., Zachar, I., Fedor, A., de Vladar, H. & Szathmáry, E. (2016). Breeding novel solutions in the brain: a model of Darwinian neurodynamics [version 1; referees: 1 approved, 2 approved with reservations]. *F1000Research* **5**, 2416.

Ucar, O., Tykocinski, L.-O., Dooley, J., Liston, A. & Kyewski, B. (2013). An evolutionarily conserved mutual interdependence between Aire and microRNAs in promiscuous gene expression. *European Journal of Immunology* **43**, 1769–1778.

Venkatesh, B., Lee, A. P., Ravi, V., Maurya, A. K., Lian, M. M., Swann, J. B., Ohta, Y., Flajnik, M. F., Sutoh, Y., Kasahara, M., Hoon, S., Gangu, V., Roy, S. W., Irimia, M., Korzh, V., et al. (2014). Elephant shark genome provides unique insights into gnathostome evolution. *Nature* **505**, 174–179.

Walker, J. A., Barlow, J. L. & McKenzie, A. N. J. (2013). Innate lymphoid cells – how did we miss them? *Nature Reviews Immunology* **13**, 75–87.

Wassenegger, M. & Krczal, G. (2006). Nomenclature and functions of RNA-directed RNA polymerases. *Trends in Plant Science* **11**, 142–151.

Watson, F. L., Puttmann-Holgado, R., Thomas, F., Lamar, D. L., Hughes, M., Kondo, M., Rebel, V. I. & Schmucker, D. (2005). Extensive diversity of Ig-superfamily proteins in the immune system of insects. *Science* **309**, 1874–1878.

Watson, R. A. & Szathmáry, E. (2016). How can evolution learn? *Trends in Ecology & Evolution* **31**, 147–157.

Weaver, C. T. & Hatton, R. D. (2009). Interplay between the Th17 and Treg cell lineages: a (co-)evolutionary perspective. *Nature Reviews Immunology* **9**, 883–889.

Weiss, A. (2015). Lamarckian illusions. *Trends in Ecology & Evolution* **30**, 566–568.

Xiao, C. & Rajewsky, K. (2009). MicroRNA control in the immune system: basic principles. *Cell* **136**, 26–36.

Xie, X., Stubbington, M. J. T., Nissen, J. K., Andersen, K. G., Hebenstreit, D., Teichmann, S. A. & Betz, A. G. (2015). The regulatory T cell lineage factor Foxp3 regulates gene expression through several distinct mechanisms mostly independent of direct DNA binding. *PLoS Genetics* **11**, e1005251.

Yang, S., Fujikado, N., Kolodin, D., Benoist, C. & Mathis, D. (2015). Immune tolerance. Regulatory T cells generated early in life play a distinct role in maintaining self-tolerance. *Science* **348**, 589–594.

Yu, C., Dong, M., Wu, X., Li, S., Huang, S., Su, J., Wei, J., Shen, Y., Mou, C., Xie, X., Lin, J., Yuan, S., Yu, X., Yu, Y., Du, J., Zhang, S., Peng, X., Xiang, M. & Xu, A. (2005). Genes "waiting" for recruitment by the adaptive immune system: the insights from amphioxus. *The Journal of Immunology* **174**, 3493–3500.

Yu, J. K. & Holland, L. Z. (2009). Cephalochordates (amphioxus or lancelets): a model for understanding the evolution of chordate characters. *Cold Spring Harbor Protocols* **2009**, https://doi.org/10.1101/pdb.emo130.

Zarnitsyna, V. I., Evavold, B. D., Schoettle, L. N., Blattman, J. N. & Antia, R. (2013). Estimating the diversity, completeness, and cross-reactivity of the T cell repertoire. *Frontiers in Immunology* **4**, 485.

Zhang, S.-M., Adema, C. M., Kepler, T. B. & Loker, E. S. (2004). Diversification of Ig superfamily genes in an invertebrate. *Science* **305**, 251–254.

Zhang, Y., Xu, K., Deng, A., Fu, X., Xu, A. & Liu, X. (2014). An amphioxus RAG1-like DNA fragment encodes a functional central domain of vertebrate core RAG1. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 397–402.

Zheng, Y., Josefowicz, S. Z., Kas, A., Chu, T.-T., Gavin, M. A. & Rudensky, A. Y. (2007). Genome-wide analysis of Foxp3 target genes in developing and mature regulatory T cells. *Nature* **445**, 936–940.

Zhou, X., Jeker, L. T., Fife, B. T., Zhu, S., Anderson, M. S., McManus, M. T. & Bluestone, J. A. (2008). Selective miRNA disruption in T reg cells leads to uncontrolled autoimmunity. *The Journal of Experimental Medicine* **205**, 1983–1991.

(*Received 21 December 2016; revised 28 June 2017; accepted 30 June 2017; published online 26 July 2017*)