Review

Not peer-reviewed version

Impact of RAG Transposon Co-Option on the Evolution of the Jawed Vertebrate Immune System

Pierre Pontarotti * and Vivek Keshri

Posted Date: 24 April 2026

doi: 10.20944/preprints202604.1756.v1

Keywords: transposable element domestication; comparative immunology; innate and adaptive immunity in jawed vertebrates; RAG Evolution; V(D)J recombination

 doi:10.20944/preprints202604.1756.v1

Review

Impact of RAG Transposon Co-Option on the Evolution of the Jawed Vertebrate Immune System

Pierre Pontarotti [1,2,*] and Vivek Keshri [3]

[1] UR AMU MEPHI D-258, CNRS SNC5039, France

[2] IHU Méditerranée Infection, 19-21 Boulevard Jean Moulin 13005 Marseille, France

[3] REVA University, Department of Biotechnology, School of Applied Science, Yelahanka, Bengaluru-560064, Karnataka, India

* Correspondence: pierre.pontarotti@univ-amu.fr

Abstract

This article reviews current knowledge in comparative immunology and presents updated hypotheses on the evolution of the immune system in jawed vertebrates. It focuses on the co-option of the RAG transposon in the origin of the V(D)J recombination system, proposed to have occurred in two stages. Initially, the RAG transposon, along with other eukaryote-specific transposon such as HAT, interacted with host genes in early eukaryotes, leading to a new transposition mechanism. Subsequently, RAG and host genes were integrated into the V(D)J recombination system, representing a major evolutionary innovation. The broader implications of this events are also considered. Earlier hypothesis suggest that the establishment of the V(D)J recombination system contributed to MHC polymorphism. Phylogenomic evidence indicates that key immune components, including MHC, T-cell receptors (α,β and γ,δ), and immunoglobulins, existed in ancestors and later expanded through gene duplication, forming multigene families with diverse functions. Their proteins products interact with other immune molecules to regulate immune responses. While some retained original functions, others evolved new roles through neo-functionalization. Overall, the co-option of the RAG transposon played a critical role in shaping the immune system of jawed vertebrates by driving innovation in both adaptive and innate immunity.

Keywords: transposable element domestication; comparative immunology; innate and adaptive immunity in jawed vertebrates; RAG Evolution; V(D)J recombination

Introduction

1. A Brief Overview of the Vertebrate Immune System

Vertebrates are divided into two groups: jawed vertebrates and cyclostomes. Both groups share certain features in their immune systems. In particular, they can generate diverse immune receptors through distinct somatic recombination systems. The immune systems of both cyclostomes and jawed vertebrates are organized around two main cell lineages: lymphoid and myeloid cells. Lymphoid cells are unique in their ability to somatically rearrange some of their immune receptors, creating a wide diversity of clonally expressed receptors. In jawed vertebrates, certain lymphocyte populations, such as innate lymphoid cells (ILCs), do not express rearranged receptors. Instead, these cells interact with receptor-bearing lymphocytes or with myeloid cells through various ligand - receptor systems, which may involve transmembrane or secreted molecules, such as cytokines.

Myeloid cells similarly communicate with lymphoid cells via ligand–receptor interactions. These interactions are especially prominent with lymphocytes that possess rearranged receptors, as seen in antigen-presenting cells (APCs) of jawed vertebrates. Importantly, myeloid cells can also initiate immune responses independently, a capability not shared by lymphoid cells [1].

 doi:10.20944/preprints202604.1756.v1

1.1. Variable Lymphocyte Receptor (VLR) - Based Recombination in Cyclostomes

Lamprey, used here as a representative cyclostome, possesses five types of variable lymphocyte receptors (VLRs) composed of leucine-rich repeat (LRR) domains. Four T-like lymphocyte lineages express VLRA, VLRC, VLRD, and VLRE, whereas B-like lymphocytes secrete VLRB antibodies [1,2]. Functional VLR genes are generated from incomplete germline genes through a gene-conversion–like mechanism involving flanking LRR sequences and lineage-specific cytidine deaminases. B-like lymphocytes develop in the hematopoietic typhlosole and kidneys, while T-like lymphocytes develop in the thymoid tissue located at the tips of the gill folds. The presence of distinct T- and B-like lymphocytes and specialized lymphoid tissues represents ancestral features of vertebrate immunity. The germline VLR gene contains invariant regions encoding the 5′ end of the N-terminal LRR (5′LRRNT) and the 3′ end of the C-terminal LRR (3′LRRCT) and stalk, separated by a noncoding region. Numerous LRR cassettes are located upstream and downstream of this gene. During VLR assembly, the noncoding region is replaced by donor LRR cassettes through a cytidine deaminase (CDA) initiated process. Cytosine (C) residues are converted to uracil (U), which is removed by uracil-DNA glycosylase (UNG), creating an abasic (AP) site that is cleaved by apurinic/apyrimidinic endonuclease (APE). The resulting DNA break is repaired by homologous recombination, leading to the formation of a mature and diverse VLR gene. The assembled VLR diversity region includes LRRNT, LRR1, multiple variable LRRs, a connecting peptide (CP), and LRRCT [1]. Although some aspects of innate immunity in cyclostomes have been described, the relationship between innate and adaptive immunity in these organisms remains poorly understood. In contrast, immune mechanisms in jawed vertebrates are well characterized and are discussed in the following section.

1.1. The Immune System of Jawed Vertebrates

As discussed above, the immune system of jawed vertebrates is partly based on lymphocytes that express highly diverse antigen receptors. This diversity is generated through somatic genetic recombination. Immune cells bearing specific receptors can undergo clonal expansion after encountering their corresponding antigen. In T cell receptors (TCRs), the initial genetic rearrangement is generally not followed by further somatic mutations, although rare exceptions have been reported in some species [3–5]. In contrast, B cell receptors (BCRs) undergo additional diversification through somatic hypermutation, a process mediated by Activation-Induced Deaminase (AID). This process promotes the selection and expansion of B cell clones with the highest antigen affinity. This essay focuses on the co-option of Recombination-Activating Genes (RAGs) and their interaction with cellular DNA repair enzymes during V(D)J recombination. The impact of RAG transposon co-option on the evolution of both adaptive and innate immunity is then examined. The innate immune system relies on receptors that do not undergo somatic rearrangement. These receptors recognize conserved microbial structures, known as Pathogen-Associated Molecular Patterns (PAMPs), as well as endogenous danger signals, termed Damage-Associated Molecular Patterns (DAMPs) [6]. Such recognition is mediated by Pattern Recognition Receptors (PRRs), including Toll-like receptors (TLRs), which are conserved across metazoans. For example, TLRs detect lipopolysaccharides (LPS) from Gram-negative bacteria and activate conserved signaling pathways such as NF-κB. However, NF-κB regulates different immune responses depending on the species, inducing antimicrobial peptides in Drosophila and cytokines such as tumor necrosis factor-α (TNFα) in jawed vertebrates [7].

Some PRRs and their signaling pathways are lineage-specific. For instance, the OAS–RNase L antiviral pathway emerged prior to the evolution of jawed vertebrates through molecular innovation [8]. These observations challenge the view that innate immunity is entirely ancient and static, instead highlighting its substantial diversification over evolutionary time.

Finally, genes originally associated with T cell and immunoglobulin receptors were later repurposed as non-rearranging receptors. These receptors enable rapid immune responses independent of immunological memory, contributing to the emergence of a distinct form of innate immunity in jawed vertebrates and complementing other innate immune innovations.

 doi:10.20944/preprints202604.1756.v1

1.1. Jawed Vertebrate RAG and the V(D)J Recombination

V(D)J recombination mediated by recombination-activating gene (RAG) proteins is a central mechanism of the adaptive immune system, generating the extensive diversity of B-cell receptors (immunoglobulins) and T-cell receptors (TCRs) required for the recognition of a broad spectrum of antigens. Immunoglobulins and γδ TCRs are capable of directly recognizing native antigens in their intact structural forms. In contrast, αβ TCRs typically recognize antigenic peptides that have been processed and presented by major histocompatibility complex (MHC) molecules on the cell surface. Notably, a subset of αβ TCRs can also recognize non-classical MHC class I molecules that present non-peptidic small molecules, such as lipids.

1.1.1. Basic Mechanisms

During B- and T-cell development, variable (V), diversity (D) - present only in immunoglobulin heavy chains, TCR β, and TCR γ - and joining (J) gene segments are randomly selected from large genomic repertoires. V(D)J recombination is initiated by the recombination-activating gene products RAG1 and RAG2, which form a complex that recognizes and binds Recombination Signal Sequences (RSSs) flanking the V, D, and J segments. This complex introduces site-specific double-strand breaks at the junctions between the coding segments and their associated RSSs. Following RAG-mediated cleavage, the coding ends adopt hairpin structures. High-mobility group (HMG) proteins contribute structurally by promoting DNA bending and unwinding, thereby enhancing accessibility of the RAG complex to the RSSs and increasing recombination efficiency. The hairpin coding ends are subsequently opened by the endonuclease Artemis, which is activated by the DNA-dependent protein kinase catalytic subunit (DNA-PKcs). Artemis-mediated cleavage generates staggered DNA ends, introducing additional junctional variability. After hairpin opening, terminal deoxynucleotidyl transferase (TdT) adds non-templated nucleotides to the exposed DNA ends, a process known as N-region diversification. In some instances, DNA polymerase μ (Pol μ) also contributes to nucleotide addition, although less efficiently than TdT. Concurrently, nucleotide deletion can occur at the coding ends prior to rejoining, further enhancing junctional diversity. Together, random nucleotide addition and deletion substantially expand the potential diversity of antigen receptors [9].

DNA polymerases Pol μ and Pol λ subsequently fill remaining gaps in a largely template-independent manner, further contributing to junctional diversity. The modified coding ends are then rejoined through the non-homologous end joining (NHEJ) DNA repair pathway. The Ku70/Ku80 heterodimer stabilizes the DNA ends, while DNA-PKcs facilitates end alignment. Finally, the XRCC4–DNA ligase IV complex ligates the DNA ends, completing the recombination process.

This overview provides a simplified description of the mechanism for clarity. For a comprehensive and detailed account, readers are referred to [9].

1.1.1. Immunoglobulin and TCR Receptors

Immunoglobulins (antibodies) and γδ T-cell receptors (TCRs) recognize and bind native antigens directly in their intact conformations, whereas αβ TCRs typically recognize antigenic peptides that have been processed intracellularly and presented on the cell surface by Major Histocompatibility Complex (MHC) molecules. In antibodies, antigen recognition is mediated by Complementarity-Determining Regions (CDRs) within the variable domains of the heavy and light chains. CDR1 and CDR2 are germline-encoded and contribute to the overall shape and charge of the antigen-binding site, while CDR3 particularly in the heavy chain - is highly variable and plays a dominant role in determining binding specificity and affinity. γδ TCRs recognize a wide range of ligands, including conformational epitopes and MHC class I–like molecules such as CD1 and MR1, in a manner reminiscent of antibody recognition [10]. In αβ TCRs, CDR1 and CDR2 primarily interact with the MHC molecule, whereas the highly variable CDR3, generated during V(D)J recombination, directly contacts the presented peptide and is central to antigen specificity [11]. Notably, certain αβ TCRs recognize non-classical MHC class I–like molecules, including CD1, which presents lipid

 doi:10.20944/preprints202604.1756.v1

antigens to invariant natural killer T (iNKT) cells, and MR1, which presents small-molecule metabolites to mucosal-associated invariant T (MAIT) cells. These unconventional TCR–ligand complexes can differ substantially from classical peptide-MHC interactions, as exemplified by MR1, whose antigen-binding cleft accommodates small metabolites derived from vitamin B biosynthesis.

2. Comparison of Cellular Immune Systems in Cyclostomes and Jawed Vertebrates and the Replacement Hypothesis

The conservation of B cell–like, αβ T cell–like, and γδ T cell–like lineages between jawed vertebrates and cyclostomes suggests that these three lymphocyte types were already present in the last common ancestor of vertebrates. This implies that key regulatory mechanisms governing their differentiation had evolved early. Both groups also possess myeloid cells. In jawed vertebrates, myeloid cells interact with lymphocytes through ligand–receptor systems involving membrane-bound and secreted factors, such as cytokines, and are capable of mounting autonomous immune responses, a feature absent in lymphocyte lineages. Although myeloid–lymphoid interactions are well characterized in jawed vertebrates, they remain less explored in cyclostomes; nonetheless, the identification of multiple cytokines and their receptors in lampreys suggests that such interactions were already established in their common ancestor [12].

Additional core immune mechanisms, including the elimination of self-reactive receptors and allelic exclusion, must also have been present in this ancestor. Moreover, an organ supporting immune cell differentiation likely existed, as both vertebrate lineages possess specialized though distinct structures where these processes occur [1].

A central unresolved question is whether the foundational elements governing lymphocyte and myeloid cell ontogeny and regulation emerged concurrently with somatic receptor diversification mediated by either RAG or cytosine deaminase (CDA) or whether they predated the evolution of somatic recombination altogether. Both RAG and CDA based systems are highly complex and likely evolved from simpler ancestral mechanisms. These systems may have arisen independently through convergent evolution, or alternatively via a replacement model, in which RAG replaced CDA in the jawed vertebrate lineage or CDA replaced RAG in cyclostomes [13]. Such replacement likely occurred within a primitive immune framework comprising distinct lymphoid and myeloid lineages. Resolving these questions will require comparative analyses of vertebrate immune systems alongside those of their closest sister phyla.

3. Origin of the Recombination Systems: CDA and RAG Based Recombination

3.1. CDA Based Recombination

Current knowledge of the origins of somatic recombination systems remains limited. Cytosine deaminase–based mechanisms involve enzymes of the AID/APOBEC family, which belong to an ancient deaminase system that emerged early in metazoan evolution and is coupled to conserved DNA repair pathways [14,15]. Members of the AID/APOBEC family have since diversified across metazoans into multiple clades with broad and conserved roles in immunity. The somatic recombination observed in cyclostomes is therefore thought to reflect the co-option of this ancestral deaminase-based system, although the precise mechanisms underlying this co-option remain unresolved.

With respect to variable lymphocyte receptors (VLRs), complete VLR-like genes that likely do not undergo somatic rearrangement have been identified in amphioxus. These genes appear to function in innate immunity and may represent the plesiomorphic, non-rearranging precursors of cyclostome VLRs [16]. Collectively, these findings highlight the need for further comparative and functional studies to clarify the origin and evolutionary trajectory of the VLR recombination system.

3.2. RAG Based Recombination

 doi:10.20944/preprints202604.1756.v1

Although a substantial body of relevant data is available, much of it requires careful reanalysis to establish a robust framework for hypothesis generation. In this context, we outline the analytical approach adopted in this review. The emergence of evolutionary novelty, such as the acquisition of a new physiological function, is a complex process that typically involves the coordinated recruitment of multiple genes and molecular systems. In the case of the V(D)J recombination machinery, several components appear to have been co-opted while retaining their ancestral functions, including Artemis and core DNA repair enzymes. In contrast, other factors, such as terminal deoxynucleotidyl transferase (TdT), acquired specialized functions following gene duplication. Notably, a transposon-derived element was also co-opted and subsequently functionally specialized to become an integral component of the system. These distinct modes of co-option can be resolved through evolutionary-based analyses, as demonstrated in previous studies [13,17,18].

3.2.1. Before RAG Co-Option/ Description of the Proto IG/TCR Complex

It is plausible that a proto–immunoglobulin/T-cell receptor (proto-Ig/TCR) complex existed before the insertion of the RAG transposon. Reconstruction of such a proto-complex requires comparative analysis of the B-cell receptor (BCR) and T-cell receptor (TCR) systems, including both αβ and γδ TCRs. The BCR is a multiprotein transmembrane complex composed of a membrane-bound immunoglobulin, consisting of two identical heavy and two identical light chains that form antigen-binding sites recognizing native antigens. Signal transduction is mediated by the associated accessory proteins Igα (CD79A) and Igβ (CD79B), which contain immunoreceptor tyrosine-based activation motifs (ITAMs) that initiate intracellular signaling upon antigen engagement. This organization is conserved across jawed vertebrates [19].

Similarly, the TCR complex comprises a heterodimeric antigen-recognition unit (αβ or γδ) associated with the CD3 signaling subunits (CD3γ, CD3δ, CD3ε, CD3ζ, and CD3η). These accessory proteins, which contain ITAMs within their cytoplasmic domains, are essential for signal transduction following antigen recognition. The TCR itself possesses only a short cytoplasmic tail and relies on the CD3 complex—organized as distinct dimers—for intracellular signaling. This architecture is likewise conserved throughout jawed vertebrates [20,21].

Notably, the CD3γ, CD3δ, and CD3ε subunits of the TCR complex and the CD79α and CD79β subunits of the BCR share common structural features, including extracellular IgV-like domains and intracellular ITAM-containing regions that are phosphorylated by conserved tyrosine kinases [22,23]. These shared structural and functional properties suggest that the proto-Ig/TCR complex likely consisted of Ig-like antigen receptors associated with accessory signaling molecules bearing IgV-like domains and ITAM motifs, coupled to conserved downstream signaling pathways [24].

Together, these observations provide a framework for reconstructing the evolutionary origins of the BCR and TCR complexes and support the hypothesis that RAG transposon insertion occurred into a pre-existing proto-Ig/TCR gene, followed by diversification into the modern immunoglobulin and TCR gene families.

3.2.2. Co-Option of the RAG Transposon

Multiple studies have demonstrated that the RAG recombinase evolved from an ancestral transposon, a process accompanied by loss of integration capacity and a concomitant increase in excision activity (Huang, Tao, Yuan, Zhang, Li, Helen A Beilinson, et al., 2016a; Tsakou-Ngouafo et al., 2020; Zhang et al., 2019). The initial step in this evolutionary transition likely involved domestication of the RAG transposon, marked by loss of integrase function and enhanced excisase activity. A subsequent event entailed the loss of the surrounding terminal inverted repeats (TIRs), resulting in fixation of the RAG locus on ancestral chromosome 11 in the common ancestor of jawed vertebrates. This scenario is supported by conserved synteny observed in non-vertebrate chordates such as amphioxus. . third key event likely involved reinsertion of TIR-like elements into an Ig/TCR-like gene, giving rise to the ancestral antigen receptor locus. This was followed by duplication and diversification of the ancestral TCR/Ig locus, leading to the emergence of distinct immunoglobulin

 doi:10.20944/preprints202604.1756.v1

and TCR gene families. Notably, RAG domestication may not have occurred directly for V(D)J recombination. [28] proposed that the early domesticated RAG complex initially functioned as a defense mechanism against Transib transposons, supported by the high sequence similarity between Transib TIRs and V(D)J recombination signal sequences (RSSs). In this model, RAG-mediated excision would have neutralized Transib elements and prevented their reintegration, consistent with the apparent absence of Transib family transposons in jawed vertebrate genomes.

Collectively, the emergence of the V(D)J recombination system can be viewed as a multistep co-option process. First, an ancestral DNA repair machinery was recruited to resolve hairpin intermediates generated during transposon excision. Second, the transposon-derived RAG system itself was co-opted to mediate programmed V(D)J recombination, potentially following or incorporating an intermediate anti-transposon defense phase.

3.2.2.1. First Step: Co-Option of an Ancestral DNA Repair System to Manage Transposons That Form Hairpins in Flanking Regions During Excision

Transposons that generate hairpin structures in the flanking regions (HFRs) during excision, notably those of the hAT and Transib/RAG families, are restricted to eukaryotes . Following excision, these flanking-region hairpins are cleaved by the Artemis–DNA-PKcs endonuclease complex, after which the DNA ends are repaired by the canonical non-homologous end-joining (C-NHEJ) machinery, including Ku70, Ku80, XRCC4, DNA ligase IV, and X-family DNA polymerases [31].

It is likely that the DNA repair system, originally involved in general repair mechanisms, predates the emergence of HFR transposon and that it was co-opted to repair the fragment after the hairpin cut. It has to be noted that DNA hairpin termini are found in both eukaryotes and prokaryotes:

1. Hairpins at the Transposon Ends: For example, in the IS4 family of bacteria and in PiggyBac transposons in eukaryotes. [31].
2. Hairpins on the Flanking Regions Left after HFR transposon excision (as already discussed examples include HAT and Transib transposons, as well as RAG-mediated rearrangements [31].
3. Eukaryotic Viral DNA Hairpin Termini: Some double-stranded viral genomes are linear but feature covalently closed hairpin termini. Examples include:Kaumoebavirus [32] Vaccinia virus [33] Chlorella virus [34] African swine fever virus [35].
4. Hairpin Telomeres in Prokaryotes: Hairpin telomeres are also found in the prokaryotic world, including in some bacteriophages, plastids, mitochondria, and bacteria. For example, see the review by [36]. Artemis and DNA-PKcs are eukaryote-specific factors that mediate the cleavage of DNA hairpins generated in flanking regions following the excision of Transib or hAT transposons. These proteins also participate in the resolution of viral DNA hairpin termini, as observed in adeno-associated viruses, which lack their own resolvase [37]. In contrast, several other eukaryotic viruses encode dedicated hairpin-resolving enzymes [38]. In prokaryotes, hairpin resolution is carried out by specialized hairpin telomere resolvases [36]. The widespread and conserved role of the Artemis–DNA-PKcs complex in repairing transposon-induced hairpins across eukaryotes suggests that this function represents its ancestral activity, whereas its involvement in viral hairpin processing likely reflects a secondary and less common co-option.

3.2.2.1.1. Cooption of Artemis and DNA-PKcs Ability to Open Flanking Hairpins

Either Artemis or DNA-PKcs may have ancestrally possessed hairpin-opening activity, potentially facilitating the emergence of hAT- and Transib-type transposons (HFR transposons) in eukaryotes. Alternatively, both proteins may have initially exhibited only weak hairpin-cleaving activity, which could have been strengthened under selective pressure from HFR transposons. Concurrently, transposons may have evolved enhanced excision capabilities as Artemis and DNA-PKcs improved their hairpin-processing efficiency, reflecting a possible co-evolutionary relationship. Failure to resolve hairpins following transposon excision would prevent DNA repair, resulting in cell death and inhibition of transposon propagation.

 doi:10.20944/preprints202604.1756.v1

Beyond transposon-related functions, Artemis in complex with DNA-PKcs plays a central role in double-strand break (DSB) repair. Germline mutations in Artemis cause radiosensitive severe combined immunodeficiency (RS-SCID), characterized by the near-complete absence of circulating B and T lymphocytes and heightened sensitivity to ionizing radiation [39,40]. Artemis-deficient cells also exhibit increased susceptibility to DSB-inducing agents and contribute to the maintenance of cell cycle checkpoints and the repair of stalled replication forks. Mechanistically, Artemis/DNA-PKcs facilitates DSB repair, in part, by removing 3′-phosphoglycolate termini generated by ionizing radiation at roughly half of all DNA breaks [41].

These observations suggest that the ancestral function of the Artemis/DNA-PKcs complex may have been general DNA break repair, with subsequent specialization for hairpin opening in the context of ancestral HFR transposons. Consistently, the absence of DNA-PKcs in yeast correlates with the lack of HFR transposons [42]. However, the presence of hAT-family Hobo transposons in Drosophila [43] raises questions regarding the mechanisms of hairpin resolution in species lacking DNA-PKcs [44].

3.2.2.1.2. Transposon Excision and P-Diversity: Generation of Junctional Diversity through Hairpin Cleavage

During RAG-mediated recombination, excision of the RAG transposon involves cleavage of DNA hairpins, resulting in the formation of single-stranded overhangs at the DNA ends. During subsequent repair, a DNA polymerase, such as polymerase X, adds palindromic (P) nucleotides to these overhangs. This process plays a critical role in the generation of junctional diversity (Colot et al., 1998; Huang, Tao, Yuan, Zhang, Li, Helen A Beilinson, et al., 2016b).

3.2.2.2. Second Step: Co-Option of the RAG Transposon System for V(D)J Recombination

The RAG V(D)J recombination system co-opted conserved proteins originally involved in the RAG transposon life cycle, including C-NHEJ factors such as KU70, KU80, XRCC4, and ligase 4 (for joining), as well as Artemis and DNA-PKcs (for repair) and polymerase X-like proteins. Additionally, it incorporated proteins that underwent functional shifts during co-option, such as RAG (for excision) and TDT (for joining).

3.2.2.2.1 Joining and N addition diversity

3.2.2.2.2. In addition to the co-option of the RAG transposon and the DNA repair machinery, several other genes were specifically recruited for the V(D)J recombination process. A prominent example is terminal deoxynucleotidyl transferase (TdT), which acquired novel functional properties following a gene duplication event. TdT and polymerase μ (Polμ) originated from a duplication within the jawed vertebrate lineage [47]. However, TdT exhibits substantially higher terminal transferase activity than both jawed vertebrate Polμ and non-vertebrate Polμ-like proteins [48]. Site-directed mutagenesis studies identified a key DNA-binding residue, Arg387, in human Polμ as a major determinant limiting its terminal transferase activity relative to TdT [49]. Substitution of this residue (R387K) to mimic TdT resulted in a marked increase in Polμ terminal transferase activity, approaching that of TdT. Notably, this R387K substitution is conserved across all jawed vertebrate TdT proteins, whereas Arg387 is retained in jawed vertebrate Polμ and non-vertebrate Polμ-like proteins. These findings indicate that TdT specialization as an efficient terminal transferase arose in the jawed vertebrate lineage following gene duplication.

3.2.2.2.3. The case of HMGB1, HMGB2, and HMGB3

HMGB1, HMGB2, and HMGB3 participate in diverse biological processes through their interactions with DNA and chromatin [50]. Among these paralogs, HMGB1 and HMGB2 play key roles in V(D)J recombination by promoting DNA bending, enhancing RAG protein binding and activity, and stabilizing the RAG–DNA complex [9]. In both mouse and human systems, HMGB1 has also been shown to facilitate DNA transposon excision, as demonstrated for the amphioxus RAG transposon (Huang et al., 2016) and the Sleeping Beauty transposon [52]. However, it remains unclear whether this activity is specific to jawed vertebrates, which uniquely possess three HMGB paralogs - HMGB1, HMGB2, and HMGB3 [53]. Substitution of human HMGB1 with HMGB proteins from non-

 doi:10.20944/preprints202604.1756.v1

jawed vertebrates in RAG transposon excision assays may help determine whether this excision-promoting function represents a specialized adaptation of HMGB1 and HMGB2 in the jawed vertebrate lineage.

Notably, canonical HMGB proteins are restricted to metazoans [53], suggesting that they did not co-evolve with DNA transposons that generate hairpin structures upon excision, as such transposons are also found in non-metazoan eukaryotes. Consequently, the evolutionary history of HMGB proteins remains difficult to resolve, particularly with respect to whether they were initially co-opted for transposition or later recruited for V(D)J recombination.

Finally, although the biochemical mechanisms of V(D)J recombination have been outlined, this process represents only one component of a highly complex system. Additional evolutionary innovations in the ancestors of jawed vertebrates were essential for shaping immunoglobulin and T-cell receptor loci and for regulating recombination activity [9]. A comprehensive understanding of the origin and evolution of V(D)J recombination therefore remains an ongoing challenge, with many aspects yet to be elucidated.

4. Possible Impact of the Co-Option of the Somatic Recombination Mechanism

How was the immune system organized prior to the emergence of receptor diversity generated by somatic recombination mechanisms such as CDA or RAG mediated processes? This question extends beyond antigen receptor variability to encompass the complexity of cellular interactions, cytokine networks, and cytokine receptor systems. Moreover, to what extent did the advent of somatic recombination contribute to the increasing complexity of the immune system, including the evolution and diversification of innate immune components?

4.1. Avoiding Self-Recognition After Establishing the Somatic Recombination System

Antigen receptors generated through somatic recombination enable the immune system to recognize a vast diversity of foreign pathogens. However, because this process is inherently stochastic, it can also produce receptors that recognize self-antigens. If not properly controlled, such self-reactive receptors may lead to autoimmune disease, in which the immune system attacks host tissues. To prevent this, the immune system employs negative selection. During this process, immature B cells that recognize self-antigens [54] and αβ T cells that either bind too strongly or fail to appropriately interact with self–MHC complexes [55] are eliminated through apoptosis. This selective removal ensures that only non-self-reactive lymphocytes mature and participate in immune responses. Negative selection thus represents a critical safeguard against autoimmunity, maintaining immune tolerance while preserving effective pathogen recognition. In the following section, we focus on the RAG recombination system, for which the most detailed mechanistic insights are currently available.

4.2. RAG Transposon Co-Option and Its Possible Impact on the Origin of the MHC Polymorphism

4.2.1. MHC Origin

The major histocompatibility complex (MHC) is essential for peptide presentation to αβ T cell receptors (TCRs), a function mediated by MHC class I and class II molecules. This role is conserved across all jawed vertebrates, indicating that the MHC-based antigen presentation system was already established in their common ancestor. However, it remains unclear whether ancestral MHC-like molecules were capable of presenting peptides to a proto-TCR lacking somatic rearrangement, or whether this interaction emerged only after the evolution of TCR rearrangement capacity. Regardless, as discussed below, the co-option of the RAG transposon played a major role in shaping MHC class I and class II polymorphism. MHC class I and class II proteins share a common ancestor, likely resembling present-day class II molecules [56]. This ancestral protein is thought to have functioned as a homodimer, with each monomer composed of an IgC1 domain and a peptide-binding

 doi:10.20944/preprints202604.1756.v1

domain (PBD). The emergence of this monomer likely resulted from domain shuffling, involving the fusion of exons encoding IgC1 and PBD domains [56]. The evolutionary origin of the PBD remains unresolved, as this domain has so far been identified exclusively in jawed vertebrates.

In addition to MHC molecules, multiple other proteins contribute to antigen presentation. As previously described [17], some of the genes involved were recruited while retaining ancestral functions, whereas others arose through gene duplication and subsequently acquired novel roles. The αβ TCR interacts with MHC class I and class II molecules in conjunction with the CD8 and CD4 co-receptors, respectively. Both CD4 and CD8 are members of the immunoglobulin superfamily, share structural similarities, and likely originated from a common ancestral gene.

It is therefore plausible that MHC molecules and CD4/CD8-like co-receptors existed prior to the emergence of the αβ TCR and the integration of the RAG transposon into the ancestral TCR/Ig locus. Within this evolutionary framework, it remains difficult to determine whether RAG-mediated diversification of the αβ TCR directly drove the differentiation and functional specialization of the CD4 and CD8 co-receptors.

4.2.2. Origin of the MHC Polymorphism

A previously proposed hypothesis suggests that an ancestral MHC-like molecule bound pathogen-associated molecular patterns (PAMPs) and presented them to ancestral TCR-like receptors [25,57].

(A) In an early stage, the ancestral MHC-like molecule likely possessed a limited binding repertoire, interacting with only a small number of pathogen-derived ligands. Correspondingly, ancestral TCR-like receptors may have recognized specific classes of PAMPs presented by these MHC-like molecules. Mutations expanding the PAMP-binding capacity of the ancestral MHC-like molecule may not have been accommodated by the existing, non-diverse TCR-like receptors, potentially leading to the loss of such variants through genetic drift.

(B) The integration of the recombination-activating gene (RAG) transposon into an ancestral TCR-like gene likely increased receptor diversity and broadened antigen recognition. This diversification would have enabled mutated MHC-like molecules presenting novel PAMPs to be recognized by a heterogeneous TCR repertoire, facilitating their evolutionary retention. These MHC variants would then be subject to selection through multiple mechanisms, including allelic polymorphism, increased peptide-binding promiscuity, and expansion into multigene families.

An alternative, but not mutually exclusive, scenario proposes that the ancestral αβ TCR initially recognized native antigens, as observed for immunoglobulins and γδ TCRs, and later evolved the capacity to recognize peptides presented by MHC molecules, a process described by the MHC capture hypothesis [58]. This transition would similarly promote peptide-binding promiscuity, allelic polymorphism, and expansion of the MHC gene family.

4.3. Beyond the Origins of MHC Polymorphism and Mechanisms to Avoid Self-Recognition, What Has Been the Impact of the Co-Option of the RAG Transposon on the Evolution of the Immune System in Jawed Vertebrates?

As outlined above, the principal proteins underlying the adaptive immune response include the Major Histocompatibility Complex (MHC), T cell receptors (TCRs), and immunoglobulins. The evolution of these key components was driven largely by gene duplication events, giving rise to multigene families with functional diversification. These core proteins do not act independently rather, they operate within complex interaction networks involving ligand–receptor interactions or direct protein - protein contacts that result in co-activation or co-inhibition (co-regulation). In addition, numerous associated proteins play essential roles in modulating both protein activity and gene expression. Some of these interacting components have retained ancestral functions and were subsequently co-opted during the evolution of the immune system, whereas others arose through gene duplication followed by neofunctionalization or through exon shuffling, enabling the acquisition of novel functions. Importantly, this evolutionary process remains ongoing and continues

 doi:10.20944/preprints202604.1756.v1

to shape adaptive immunity across diverse lineages of jawed vertebrates (see, for example, (Akula et al., 2014; Flajnik, 2016) Figure 1-2.

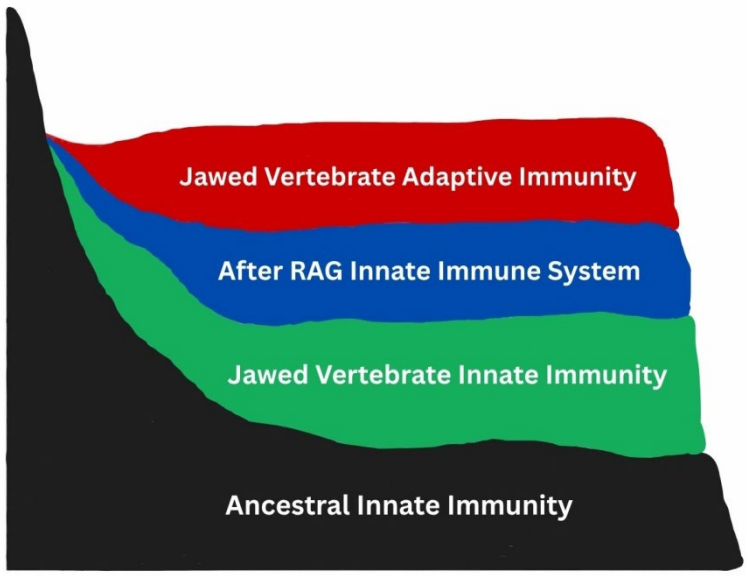

Figure 1. Schematic representation of the evolution of vertebrate immune systems. The black region indicates ancestral innate immunity. The green layer represents the development of vertebrate innate immunity. The blue layer shows the enhancement of innate immune responses following RAG-mediated mechanisms. The red layer illustrates the emergence of vertebrate adaptive immunity, highlighting increasing complexity over evolutionary time.

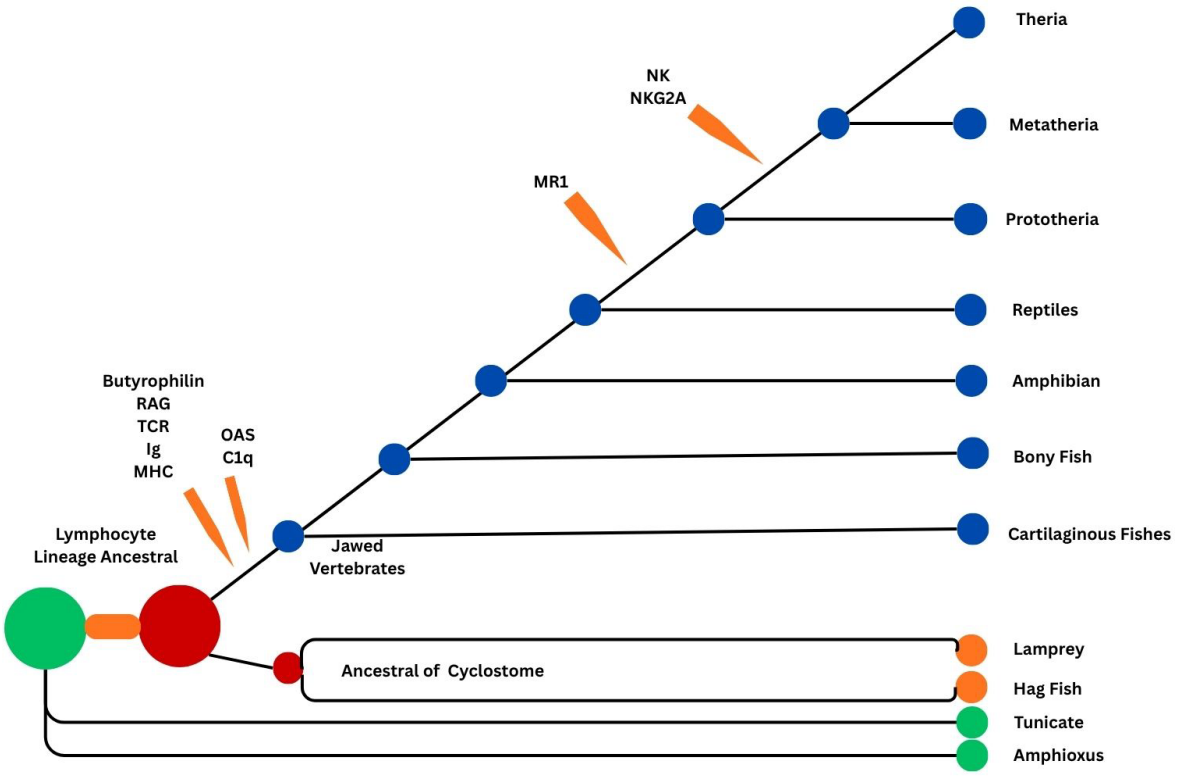

Figure 2. Phylogenetic tree depicting the evolutionary emergence of some immune components across jawed vertebrates.. Large green node = lymphocyte lineage ancestor; large red node = last common vertebrate ancestor;

 doi:10.20944/preprints202604.1756.v1

blue nodes = jawed vertebrate ancestor and descendant taxa; orange nodes = cyclostomes (Lamprey, Hagfish); green nodes = invertebrate chordates (Tunicate, Amphioxus). Orange arrows indicate the branch point at which key immune genes first appeared: RAG, TCR, Ig, MHC, and Butyrophilin emerged at the jawed vertebrate ancestor marking the origin of adaptive immunity; OAS and C1q were acquired at the same node; MR1 arose prior to the reptile divergence; NK receptors and NKG2A appeared prior to mammalian divergence.

4.3.1. Origin and Evolution of Divergent Class I Families and Their Receptors

Classical MHC class I molecules, which present peptide antigens to αβ T cell receptors (TCRs), are conserved across all examined jawed vertebrates. In contrast, multiple forms of non-classical MHC class I molecules exhibit lineage-specific distributions and are restricted to particular subphyla of jawed vertebrates. Phylogenetic analyses indicate that non-classical MHC class I genes originated from classical MHC class I genes through gene duplication events that occurred at distinct stages during the evolution of jawed vertebrates.

Each non-classical MHC class I family clusters as a monophyletic group with the classical MHC class I molecules of its corresponding clade, including mammals, primates, eutherian mammals, bony fishes, and other lineages. These duplication events arose independently in different branches of the jawed vertebrate phylogeny, followed by lineage-specific functional divergence. Over evolutionary time, the duplicated genes underwent functional modifications that resulted in the specialized immunological roles characteristic of non-classical MHC class I molecules within each lineage. The following sections describe these non-classical MHC class I genes and their associated receptors in detail.

4.3.1. 1Non-Classical Class I Genes Interacting with Semi-Rearranging αβ TCRs and Certain γδ T Cells

Semi-rearranging αβ T cell receptors (TCRs), as well as γδ T cells (see above), participate in innate immune responses and recognize MHC-like proteins either in association with small organic ligands or, in some cases, in the absence of bound ligands. In mammals, two well-characterized populations of innate αβ T cells have been described: mucosal-associated invariant T (MAIT) cells and invariant natural killer T (iNKT) cells. Together with γδ T cells, these populations constitute the three major classes of innate T cells.

Invariant natural killer T (iNKT) cells represent a specialized T cell subset that recognizes lipid antigens presented by the MHC-like molecule CD1d. The MAIT cell compartment similarly exhibits limited TCR diversity in humans and recognizes a restricted range of microbially derived vitamin B metabolites presented by the MHC-like molecule MR1.

The interaction between αβ TCRs and peptide-presenting MHC molecules is conserved across all jawed vertebrates, with the exception of certain species in which key components of this system have been pseudogenized, as exemplified by anglerfish [62]. In contrast, interactions between αβ TCRs and non-peptidic antigens presented by MHC-like molecules are restricted to specific vertebrate subphyla. The CD1 and MR1 families emerged later during jawed vertebrate evolution; the CD1 family originated within the tetrapod lineage through duplication of a classical MHC class I gene [63], whereas the MR1 family arose from another classical MHC family member in the ancestors of mammals [63,64].

These observations suggest that αβ TCR recognition of non-peptidic antigens represents a derived evolutionary state resulting from the co-evolution of MHC–antigen complexes and αβ TCRs. Alternatively, ancestral TCRs may initially have recognized MHC molecules presenting lipid antigens or microbially derived vitamin B metabolites. In this scenario, subsequent structural evolution of the TCR binding interface enabled peptide recognition, while somatic recombination further expanded TCR diversity and facilitated the recognition of novel antigenic targets. γδ T cell receptors recognize a wide range of native structures (see Table 1 in [65] including non-classical MHC class I molecules, but do so through mechanisms distinct from those employed by αβ TCRs. γδ TCRs may recognize non-classical MHC class I molecules in the absence of a ligand, recognize ligand-

 doi:10.20944/preprints202604.1756.v1

bound MHC molecules without directly contacting the ligand, or engage both the ligand and the MHC molecule in a manner distinct from αβ TCR recognition (Adams et al., 2015; Harly et al., 2022.; Le Nours et al., 2019a; Pellicci et al., 2020; Rice et al., 2021). These interaction modes are likely to represent derived states.

Gamma-delta (γδ) T cells play a critical role in sensing danger signals and pathogen-associated molecular patterns (PAMPs) and appear to opportunistically utilize MHC molecules, either with or without bound ligands. Consequently, the emergence of the RAG-mediated recombination system may have facilitated not only the expansion of TCR diversity but also the diversification of MHC recognition and function, thereby contributing to the functional complexity of adaptive and innate-like immune responses.

4.3.1.2. Classical and Non-Classical MHC Molecules in Interaction with Non-TCR Receptors

As described above, interactions between classical MHC molecules and αβ T cell receptors (TCRs) likely emerged early during the evolution of jawed vertebrates, as they are conserved across nearly all extant members of this group. γδ T cell receptors are also capable of interacting with members of the butyrophilin family, a protein family that is conserved throughout jawed vertebrates. However, outside of mammals such as humans and mice, such interactions have not yet been documented in other vertebrate lineages.

In addition to TCR-mediated recognition, both classical and non-classical MHC molecules engage receptors other than TCRs. Prominent examples include interactions with natural killer (NK) cell receptors such as NKG2A, NKG2D, leukocyte immunoglobulin-like receptors (LILRs), and killer cell immunoglobulin-like receptors (KIRs), which are largely restricted to mammals. In rodents, KIRs are functionally replaced by Ly49 receptors. These mammal-specific receptor systems are thought to have originated in the ancestors of modern mammals [70,71].

The majority of these receptors are expressed on the surface of natural killer cells, which survey target cells through mechanisms of missing-self and induced-self recognition, including the detection of MHC–peptide complexes (see [72].

4.3.1.2.1. NK receptors that interact with MHC / peptide complex

Natural killer (NK) cells are lymphocytes that play critical roles in innate immunity, contribute to the regulation of adaptive immune responses, and participate in placental development. In mice and humans, NK cell education and functional regulation are mediated by diverse receptor systems that recognize MHC molecules expressed on tissue cells. These NK cell receptors interact with both classical and non-classical MHC class I molecules, thereby shaping NK cell responsiveness and effector functions.

NK receptors that interacts with classical MHC class I /peptide complex

In humans, polymorphic epitopes of classical MHC class I molecules namely HLA-A, HLA-B, and HLA-C—are recognized by a diverse repertoire of killer cell immunoglobulin-like receptors (KIRs). In mice, an analogous function is mediated by Ly49 receptors, which interact primarily with H-2K and H-2D molecules. KIRs and Ly49 receptors are structurally distinct cell-surface glycoproteins that serve as receptors for classical MHC class I molecules on natural killer (NK) cells. In species such as mice, rats, and horses, the Ly49 gene family encodes multiple inhibitory and activating receptors with specificities for different MHC class I molecules, which are expressed in diverse combinations on NK cells [73–75]. In contrast, in primates, cattle, and goats, functional Ly49 genes are absent, and NK cell recognition of MHC class I is mediated instead by KIRs [76,77]. This complementary distribution indicates a functional replacement between the Ly49 and KIR receptor families across mammalian lineages.

Both KIR and Ly49 receptors contribute to NK cell activation through “missing-self” recognition. When MHC class I expression is reduced or lost, for example as a consequence of viral infection or cellular stress, inhibitory receptor signaling is diminished, resulting in NK cell activation and elimination of the affected cells. Although the downstream signaling pathways are broadly conserved, the mechanisms governing activating receptor function remain less well understood.

 doi:10.20944/preprints202604.1756.v1

Notably, activating receptors are thought to have arisen through convergent evolution from inhibitory receptor homologs [78].

KIRs and Ly49 receptors differ markedly in their modes of MHC class I recognition. KIRs engage HLA class I molecules at the C-terminal end of the peptide-binding groove, with residues at positions 7 and 8 of the bound nonamer peptide exerting a major influence on receptor binding (see, for example [72]. In contrast, murine Ly49 receptors bind beneath the peptide-binding platform, contacting the MHC α3 domain and regions adjacent to the peptide-binding groove, without direct interaction with the presented peptide [79,80].

An exception within the Ly49 family is Ly49C, which exhibits peptide-dependent sensitivity to the MHC class I molecule H-2K^b [81]. This sensitivity does not result from direct peptide–receptor contact; instead, dynamic allosteric effects within the MHC molecule have been shown to modulate Ly49C recognition [82].

Evolutionarily, KIR genes arose in the ancestors of eutherian mammals through exon shuffling [83], whereas the Ly49 family likely originated in the common ancestor of mammals [84]**.** Subsequent lineage-specific expansions led to diversification of the KIR gene family in some mammalian lineages and of the Ly49 family in others, reinforcing their complementary roles in NK cell recognition of MHC class I molecules.

NK receptors that interacts with non classical MHC class I peptide complex

NKG2A/CD94 receptors are present in eutherian mammals [84]. In humans and mice, for which functional characterization is most extensive, the NKG2A/CD94 heterodimer interacts with HLA-E in humans and with H2-Qa1 in mice [85,86]. In both species, NKG2A/CD94 recognizes MHC class I–derived leader sequence peptides presented by HLA-E and H2-Qa1, respectively, resulting in the delivery of inhibitory signals to natural killer (NK) cells.

HLA-E and H2-Qa1 are non-classical MHC class Ib molecules encoded within the MHC region. These molecules may have arisen independently from classical MHC class I genes in primate and rodent lineages, respectively. Consistent with this possibility, phylogenetic analyses indicate that HLA-E in humans and H2 class I genes in mice form distinct monophyletic groups. However, this pattern may alternatively reflect concerted evolution within the MHC locus. Under such a scenario, HLA-E and H2-Qa1 could represent orthologous genes rather than independently derived paralogs [87].

4.3.1.2.2. Interaction with MHC/ MHC like with non T cell receptor independently of the PBD and the peptide bound to it

Interaction of MHC non classical class I with Leukocyte immunoglobulin-like receptors (LILR)

Leukocyte immunoglobulin-like receptors (LILRs) can be classified into two major groups based on ligand specificity. Group 1 LILRs, which include LILRA1, LILRA2, LILRA3, LILRB1, and LILRB2, possess highly conserved HLA-binding sites and interact with the most conserved regions of MHC class I molecules, namely the α3 domain and β2-microglobulin In contrast, Group 2 LILRs - including LILRA4, LILRA5, LILRA6, LILRB3, LILRB4, and LILRB5 recognize a broader range of ligands and generally do not bind classical MHC class I molecules (Shiroishi, Kuroki, et al., 2006). An exception within this group is LILRB5, which is capable of binding both angiopoietin-like proteins and HLA class I molecules. LILRs are present in all eutherian mammals [88] and share a common evolutionary ancestor with killer cell immunoglobulin-like receptors (KIRs) in the common ancestor of mammals. The LILR and KIR families subsequently diverged and became distinct receptor lineages in the ancestors of eutherian mammals.

Interaction of non MHC classical class I gene (MICA/MICB, ULBP) with NKG2D

NKG2D (KLRK1) is a major activating receptor expressed by natural killer (NK) cells. Its ligands comprise a diverse set of MHC class I–like molecules whose expression is induced in response to cellular stress, including infection and tumorigenesis. In humans, two principal families of NKG2D ligands (NKG2DLs) have been identified: the MHC class I–related chains (MIC), which are encoded within the MHC region, and the UL16-binding proteins (ULBPs), which are encoded outside the

 doi:10.20944/preprints202604.1756.v1

MHC locus. In mice, ULBP-like ligands are also present, together with an additional ligand, MILL, which is not found in humans (see [89].

The designation NKG2D is somewhat misleading, as the NKG2D gene exhibits limited sequence similarity to the closely related NKG2A, NKG2C, NKG2E, and NKG2F genes, which form a highly homologous gene cluster and are thought to have arisen through gene duplication events.

All known NKG2D ligands contain α1 and α2 domains that form an MHC class I–like fold; however, none appear to possess an open peptide-binding groove. In humans, MICA and MICB, unlike ULBPs or murine NKG2D ligands, additionally contain an α3 domain but do not associate with β2-microglobulin, distinguishing them from classical HLA class I molecules. Despite the low sequence similarity between human MIC and ULBP proteins and their murine counterparts (RAE-1, H60, and MULT1), all of these ligands bind NKG2D with relatively high affinity. Structural analyses of NKG2D complexes with ULBP3, MICA, and RAE-1β reveal a conserved receptor–ligand interaction mode, in which NKG2D engages the top of the α1–α2 platform domain, despite the use of distinct amino acid residues across ligand families [90–92].

NKG2D (KLRK1) is present in therian mammals, including both eutherians and marsupials [93–95]. Evolutionary analyses indicate that MIC-, ULBP-, and MILL-like genes also emerged in therian mammals, consistent with co-evolution of the NKG2D receptor–ligand system [96]. Notably, the genomic loci encoding NKG2D, Ly49, and NKG2A/CD94 are maintained in conserved synteny in eutherian mammals [97]. In contrast, among marsupials—the sister group of eutherian mammals only NKG2D is retained within a conserved syntenic context.

Other MHC class I-like molecules that do not interact with either αβ TCRs or NK receptors.

Endothelial Protein C Receptor (EPCR)

Similar to members of the CD1 family, the endothelial protein C receptor (EPCR) harbors a tightly bound phospholipid within its antigen-presenting groove. This phospholipid is required for EPCR to interact with protein C, an interaction mediated by contacts between the α1 domain of EPCR and the γ-carboxyglutamic acid (Gla) domain of protein C [98]. This binding facilitates the activation of protein C and is therefore critical for the protein C–dependent anticoagulant pathway [99]

In addition to its role in coagulation, EPCR can directly engage the T cell receptor (TCR) of a human Vγ4Vδ5 γδ T cell clone, thereby enabling γδ T cells to recognize cytomegalovirus-infected endothelial cells as well as epithelial tumor cells. This mode of recognition is mechanistically distinct from protein C binding, as it involves the β-sheet face of EPCR rather than the α-helical face used for interaction with protein C [100]. The specificity of this interaction is largely determined by the complementarity-determining region 3 (CDR3) loops of both the γ and δ chains, which exhibit extensive variability generated by RAG-mediated recombination.

This structural and genetic variability endows γδ TCRs with the capacity to recognize a broad spectrum of ligands with diverse conformations, thereby contributing to their functional versatility. Such recognition is consistent with the opportunistic binding strategy of γδ TCRs, which preferentially engage native structural motifs, including those present in MHC and MHC-like molecules (see above).

HFE interacts with transferrin receptor role on the iron uptake

HFE, a non-classical MHC class I–like molecule, interacts with transferrin receptor 1 (TFR1), a homodimeric type II transmembrane glycoprotein responsible for cellular iron uptake through binding of iron-loaded transferrin. Structural and biochemical analyses indicate that the HFE–TFR1 interaction involves the α1 and α2 helical domains of HFE and the helical region of TFR1, which also participates in transferrin binding. The resulting HFE–TFR1 complex exhibits two-fold symmetry with a 2:2 stoichiometry [101].

ZAG (zinc-α2-glycoprotein)

The structure of zinc-α2-glycoprotein (ZAG) closely resembles that of classical MHC class I molecules at the level of backbone conformation, sharing approximately 30–40% amino acid sequence identity [102]. However, ZAG exhibits notable modifications within the putative antigen-binding groove, which is adapted for the binding of small hydrophobic molecules, including compounds

 doi:10.20944/preprints202604.1756.v1

structurally similar to polyethylene glycols (PEGs) or fatty acids**.** Despite these structural insights, the identity of the physiologically relevant ligand(s) for ZAG remains unknown, as does its precise biological function in vivo.

The neonatal Fc receptor (FcRn)

The neonatal Fc receptor (FcRn) is an MHC class I–like molecule that has evolved specialized functions distinct from antigen presentation. FcRn primarily binds immunoglobulin G (IgG) derived from maternal milk, mediating its transcytosis across mucosal epithelia and subsequent release into the circulation. This process is essential for passive immunity in neonates, enabling the transfer of maternal antibodies that provide early immune protection.

In addition to its role in neonatal immunity, FcRn plays a central role in maintaining systemic IgG homeostasis by protecting IgG molecules from lysosomal degradation, thereby regulating both mucosal and systemic immune responses. Unlike classical MHC class I molecules, which possess an open antigen-binding groove, FcRn contains a closed groove that does not participate in ligand binding. Instead, FcRn engages a single Fc chain through interactions at the edge of its platform domain, with key contact residues contributed by the α2 domain and the β2-microglobulin (β2m) subunit [103–105]**.**

MHC class I–like genes are present across jawed vertebrate lineages beyond mammals [106]**.** As discussed above, the functions and molecular interactions of these molecules have been extensively characterized in eutherian mammals. In contrast, comparable functional data remain scarce for non-eutherian mammals and other jawed vertebrates. A notable exception has been described in Xenopus, where interactions between MHC class I–like molecules and an invariant T cell lineage have been reported [107,108]**.**

4.3.2. Immunoglobulin Co-Option

4.3.2.1. C1q Binding to Immunoglobulins

The Classical and Lectin Pathways of the complement system are both present in jawed vertebrates. These pathways converge on the formation of the same C3 convertase (C4b2a), which plays a central role in pathogen defense. Here's an overview of their activation and shared mechanisms.

1. The Classical Pathway:

This pathway is initiated when antibodies (IgG or IgM) bind to a pathogen, activating the C1 complex, which consists of C1q and the proteases C1r and C1s. Upon activation, C1q triggers C1r and C1s to cleave C4 and C2, forming the C3 convertase.

2. The Lectin Pathway:

The lectin pathway is activated when mannose-binding lectin (MBL) or ficolins recognize specific carbohydrate patterns on pathogens. This recognition activates MASP-1 and MASP-2 proteases, which then cleave C4 and C2, generating the same C3 convertase as the classical pathway. In both pathways, the formation of C3 convertase leads to:Opsonization: Marking pathogens for phagocytosis., Inflammation: Through the release of C3a. and Membrane Attack Complex (MAC) Formation: Leading to pathogen lysis [109]. Interestingly, the lectin pathway is present even in non-jawed vertebrates (reviewed by [110]), and C1q itself is found in species outside the jawed vertebrate phyla [111]. This suggests that the interaction between C1q and immunoglobulins represents an evolutionary innovation, combining an ancient protein (C1q) with a newer protein family (immunoglobulins). This adaptation highlights a significant step in the evolution of the vertebrate immune system, linking pre-existing components with novel mechanisms of pathogen recognition and defense.

4.3.2.2. FC Receptors Immunoglobulin Interaction

Fc receptors (FcRs) are members of the immunoglobulin superfamily and are expressed on the surface of diverse immune cells, including macrophages, neutrophils, and natural killer cells. These

 doi:10.20944/preprints202604.1756.v1

receptors bind the Fc region of immunoglobulins and play central roles in antibody-mediated immune responses. Through interactions with antibodies, FcRs mediate a range of effector functions, including phagocytosis, antibody-dependent cellular cytotoxicity (ADCC), and modulation of inflammatory responses.

FcRs exhibit isotype specificity and recognize distinct Fc classes, such as IgG, IgA, and other immunoglobulin isotypes. Accordingly, FcRs are classified into different receptor families, including Fcγ receptors (FcγRs), which bind IgG, and Fcα receptors (FcαRs), which bind IgA. Comparative genomic analyses indicate that the expansion and diversification of FcR gene families vary among jawed vertebrate lineages, reflecting co-evolution with immunoglobulin Fc isotypes. This coordinated evolution of receptors and antibody classes is thought to enhance the functional versatility and effectiveness of antibody-dependent immune mechanisms [60].

4.3.3. Gamma Delta TCRs

4.3.4. Gamma delta (γδ)

T lymphocytes identified to date exhibit a notable bias against self-reactive recognition. Consistent with the immunoglobulin-like structure of γδ T-cell928 receptors (TCRs) and the diversity of their repertoire (see above), the self-antigens directly recognized by γδ TCRs are structurally diverse, encompassing both MHC-related and MHC-unrelated molecules [65]. Many of these self-antigens are constitutively expressed on cells and healthy tissues, suggesting the presence of regulatory mechanisms that ensure γδ T-cell responses remain tightly controlled, thereby preventing autoimmunity.

This highlights the specialization of γδ T cells in recognizing danger signals, a function shared with cells expressing non-rearranging receptors [112].

In humans, this recognition can occur either through clonal expansion, as seen predominantly in Vδ1 T cells, or independently of clonal expansion, as observed in Vγ9Vδ2 T cells [113].

It has to be noted that some Gamma Delta TCRs interact with member of the butyrophilin families , this has well be described in the cases of mammalians The interaction between TCR gamma delta (TCRγδ) and butyrophilin (BTN) heterodimers exhibits significant variation across mammalian species. In most mammals, Vγ9Vδ2+ T cells interact with BTN3/BTN2, whereas in humans, Vγ4+ T cells engage with BTNL3/BTNL8, and in mice, Vγ7+ T cells interact with Btnl1/Btnl6. The BTN sub-families originated in the common ancestor of mammals, but evolutionary divergence has led to species-specific losses. For instance, BTNL3/BTNL8 have been lost in the mouse lineage, Btnl1/Btnl6 in the human lineage, and BTN3 has also been lost in the mouse lineage [114,115]. The biology of the butyrophilin-TCRγδ system is best understood in the context of BTN3/BTN2 and Vγ9Vδ2+ T cells in humans [114,116].

Vγ9Vδ2+ T cells interact with BTN2/BTN3 through a mechanism involving

phosphoantigens. These phosphoantigens bind to the intracellular B30.2 domain of BTN3, triggering a conformational change that enables BTN3 to interact with BTN2A1. BTN2A1 subsequently binds directly to the TCR Vγ9 chain on Vγ9Vδ2+ T cells. Phosphoantigens recognized by the B30.2 domain can originate from microorganisms via the non-mevalonate isoprenoid synthesis pathway. A notable example is the high-affinity , binding to B30.2, phosphoantigen (E)-4-hydroxy-3-methyl-but-2-enyl pyrophosphate (HMBPP), produced by bacteria and apicomplexan [117]. Alternatively, phosphoantigens can be generated by mammalian cells through the mevalonate pathway, which produces the lower-affinity binding to B30.2 phosphoantigen isopentenyl pyrophosphate (IPP) [118].

The phosphoantigen (E)-4-hydroxy-3-methyl-but-2-enyl pyrophosphate (HMBPP) could therefore considered as PAMP and phosphoantigen isopentenyl pyrophosphate (IPP) as danger signal. We do not have information for the other members of the butyrophilin family in human mice even if it has been shown that they interact with gamma delta cell [114] and regulate their

 doi:10.20944/preprints202604.1756.v1

activities.971 The butyrophilin family is conserved across all vertebrate species [115]. However, functional data on these families in non-mammalian have not been so far investigated.

The presence of TCRγδ has been described in jawed vertebrates [119] with exception of squamate (Morrissey et al. 2022 shown) this suggests that the interaction between TCRγδ and butyrophilin may have been present in the common ancestor of jawed vertebrates.

An important question arises: does the interaction between butyrophilin and TCR occur with prearranged TCRs, predating the co-option of the RAG transposon? This possibility could parallel the hypothesized interaction between MHC molecules and prearranged TCRαβ. In any cases the co-option of RAG allow the diversification of the Gamma Delta TCR family and their role on detecting danger signal and PAMP.

Conclusion

In conclusion, the co-option of the RAG transposon represents a transformative event in the evolution of the immune system in jawed vertebrates. This key innovation laid the foundation for the adaptive immune system, enabling the recognition of a vast repertoire of antigens. Additionally, the domestication of RAG significantly broadened the capacity of organisms to detect and respond to danger signals and pathogen-associated molecular patterns (PAMPs), highlighting its pivotal role in shaping immune defence mechanisms.:

Author Contributions: Pierre Pontarotti conceived the idea for this review and prepared the manuscript first fraft. Vivek Keshri contributed by editing the manuscript and preparing the figures.

Funding: This research was supported by the French Government under the "Investissements D'avenir" (Investments for the Future) program, managed by the Agence Nationale de la Recherche Française (ANR National Agency for Research), (reference: Méditerranée Infection 10-IAHU-03).

Acknowledgments: The author sincerely thanks all the MSc students who contributed to this project, as well as colleagues in comparative immunology, immunology, and evolutionary biology for their valuable discussions and insights.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Boehm, T.; Hirano, M.; Holland, S.J.; Das, S.; Schorpp, M.; Cooper, M.D. Evolution of Alternative Adaptive Immune Systems in Vertebrates. *Annu. Rev. Immunol.* 2018, *36*, 19–42, doi:10.1146/annurev-immunol-042617-053028.
2. Das, S.; Boehm, T.; Holland, S.J.; Rast, J.P.; Fontenla-Iglesias, F.; Morimoto, R.; Valadez, J.G.; Heimroth, R.D.; Hirano, M.; Cooper, M.D. Evolution of Two Distinct Variable Lymphocyte Receptors in Lampreys: VLRD and VLRE. *Cell Rep.* 2023, *42*, doi:10.1016/j.celrep.2023.112933.
3. Ott, J.A.; Castro, C.D.; Deiss, T.C.; Ohta, Y.; Flajnik, M.F.; Criscitiello, M.F. Somatic Hypermutation of T Cell Receptor α Chain Contributes to Selection in Nurse Shark Thymus. *Elife* 2018, *7*, doi:10.7554/eLife.28477.
4. Antonacci, R.; Mineccia, M.; Lefranc, M.P.; Ashmaoui, H.M.E.; Lanave, C.; Piccinni, B.; Pesole, G.; Hassanane, M.S.; Massari, S.; Ciccarese, S. Expression and Genomic Analyses of Camelus Dromedarius T Cell Receptor Delta (TRD) Genes Reveal a Variable Domain Repertoire Enlargement Due to CDR3 Diversification and Somatic Mutation. *Mol. Immunol.* 2011, *48*, 1384–1396, doi:10.1016/j.molimm.2011.03.011.
5. Chen, H.; Bernstein, H.; Ranganathan, P.; Schluter, S.F. Somatic Hypermutation of TCR γ V Genes in the Sandbar Shark. *Dev. Comp. Immunol.* 2012, *37*, 176–183, doi:10.1016/j.dci.2011.08.018.
6. Pradeu, T.; Thomma, B.P.H.J.; Girardin, S.E.; Lemaitre, B. The Conceptual Foundations of Innate Immunity: Taking Stock 30 Years Later. *Immunity* 2024, *57*, 613–631, doi:10.1016/j.immuni.2024.03.007.
7. Brubaker, S.W.; Bonham, K.S.; Zanoni, I.; Kagan, J.C. Innate Immune Pattern Recognition: A Cell Biological Perspective. *Annu. Rev. Immunol.* 2015, *33*, 257–290, doi:10.1146/annurev-immunol-032414-112240.

 doi:10.20944/preprints202604.1756.v1

8. Chu, L.; Gong, Z.; Wang, W.; Han, G.Z. Origin of the OAS-RNase L Innate Immune Pathway before the Rise of Jawed Vertebrates via Molecular Tinkering. *Proc. Natl. Acad. Sci. U. S. A.* 2023, *120*, doi:10.1073/pnas.2304687120.
9. Schatz, D.G.; Zhang, Y.; Xiao, J.; Zha, S.; Zhang, Y.; Alt, F.W. The Mechanism, Regulation and Evolution of V(D)J Recombination. *Molecular Biology of B Cells, Third Edition* 2024, 13–57, doi:10.1016/B978-0-323-95895-0.00004-0.
10. Sok, C.L.; Rossjohn, J.; Gully, B.S. The Evolving Portrait of Γδ TCR Recognition Determinants. *J. Immunol.* 2024, *213*, 543–552, doi:10.4049/jimmunol.2400114.
11. Rossjohn, J.; Gras, S.; Miles, J.J.; Turner, S.J.; Godfrey, D.I.; McCluskey, J. T Cell Antigen Receptor Recognition of Antigen-Presenting Molecules. *Annu. Rev. Immunol.* 2015, *33*, 169–200, doi:10.1146/annurev-immunol-032414-112334.
12. Boulay, J.-L.; Du Pasquier, L.; Cooper, M.D. Cytokine Receptor Diversity in the Lamprey Predicts the Minimal Essential Cytokine Networks of Vertebrates. *J. Immunol.* 2022, *209*, 1013–1020, doi:10.4049/jimmunol.2200274.
13. Morales Poole, J.R.; Paganini, J.; Pontarotti, P. Convergent Evolution of the Adaptive Immune Response in Jawed Vertebrates and Cyclostomes: An Evolutionary Biology Approach Based Study. *Dev. Comp. Immunol.* 2017, *75*, 120–126, doi:10.1016/j.dci.2017.02.011.
14. Krishnan, A.; Iyer, L.M.; Holland, S.J.; Boehm, T.; Aravind, L. Diversification of AID/APOBEC-like Deaminases in Metazoa: Multiplicity of Clades and Widespread Roles in Immunity. *Proc. Natl. Acad. Sci. U. S. A.* 2018, *115*, E3201–E3210, doi:10.1073/PNAS.1720897115/-/DCSUPPLEMENTAL.
15. Liu, M.C.; Liao, W.Y.; Buckley, K.M.; Yang, S.Y.; Rast, J.P.; Fugmann, S.D. AID/APOBEC-like Cytidine Deaminases Are Ancient Innate Immune Mediators in Invertebrates. *Nature Communications 2018 9:1* 2018, *9*, 1948-, doi:10.1038/s41467-018-04273-x.
16. Cao, D.D.; Liao, X.; Cheng, W.; Jiang, Y.L.; Wang, W.J.; Li, Q.; Chen, J.Y.; Chen, Y.; Zhou, C.Z. Structure of a Variable Lymphocyte Receptor-like Protein from the Amphioxus Branchiostoma Floridae. *Sci. Rep.* 2016, *6*, doi:10.1038/srep19951.
17. Danchin, E.; Vitiello, V.; Vienne, A.; Richard, O.; Gouret, P.; McDermott, M.F.; Pontarotti, P. The Major Histocompatibility Complex Origin. *Immunol. Rev.* 2004, *198*, 216–232, doi:10.1111/j.0105-2896.2004.00132.x.
18. Danchin, E.; Vitiello, V.; Vienne, A.; Richard, O.; Gouret, P.; McDermott, M.F.; Pontarotti, P. The Major Histocompatibility Complex Origin. *Immunol. Rev.* 2004, *198*, 216–232, doi:10.1111/J.0105-2896.2004.00132.X.
19. Li, R.; Wang, T.; Bird, S.; Zou, J.; Dooley, H.; Secombes, C.J. B Cell Receptor Accessory Molecule CD79α: Characterisation and Expression Analysis in a Cartilaginous Fish, the Spiny Dogfish (Squalus Acanthias). *Fish Shellfish Immunol.* 2013, *34*, 1404–1415, doi:10.1016/j.fsi.2013.02.015.
20. Pettinello, R.; Redmond, A.K.; Secombes, C.J.; Macqueen, D.J.; Dooley, H. Evolutionary History of the T Cell Receptor Complex as Revealed by Small-Spotted Catshark (Scyliorhinus Canicula). *Dev. Comp. Immunol.* 2017, *74*, 125–135, doi:10.1016/j.dci.2017.04.015.
21. Gully, B.S.; Ferreira Fernandes, J.; Gunasinghe, S.D.; Vuong, M.T.; Lui, Y.; Rice, M.T.; Rashleigh, L.; Lay, C.S.; Littler, D.R.; Sharma, S.; et al. Structure of a Fully Assembled Γδ T Cell Antigen Receptor. *Nature* 2024, *634*, 729–736, doi:10.1038/s41586-024-07920-0.
22. Radaev, S.; Zou, Z.; Tolar, P.; Nguyen, K.; Nguyen, A.T.; Krueger, P.D.; Stutzman, N.; Pierce, S.; Sun, P.D. Structural and Functional Studies of Igαβ and Its Assembly with the B Cell Antigen Receptor. *Structure* 2010, *18*, 934–943, doi:10.1016/j.str.2010.04.019.
23. Gold, D.P.; Clevers, H.; Alarcon, B.; Dunlap, S.; Novotny, J.; Williams, A.F.; Terhorst, C. Evolutionary Relationship between the T3 Chains of the T-Cell Receptor Complex and the Immunoglobulin Supergene Family. *Proc. Natl. Acad. Sci. U. S. A.* 1987, *84*, 7649–7653, doi:10.1073/pnas.84.21.7649.
24. Cantrell, D. Signaling in Lymphocyte Activation. *Cold Spring Harb. Perspect. Biol.* 2015, *7*, 1–14, doi:10.1101/cshperspect.a018788.
25. Tsakou-Ngouafo, L.; Paganini, J.; Kaufman, J.; Pontarotti, P. Origins of the RAG Transposome and the MHC. *Trends Immunol.* 2020, *41*, 561–571, doi:10.1016/j.it.2020.05.002.

 doi:10.20944/preprints202604.1756.v1

26. Zhang, Y.; Cheng, T.C.; Huang, G.; Lu, Q.; Surleac, M.D.; Mandell, J.D.; Pontarotti, P.; Petrescu, A.J.; Xu, A.; Xiong, Y.; et al. Transposon Molecular Domestication and the Evolution of the RAG Recombinase. *Nature* 2019, *569*, 79–84, doi:10.1038/s41586-019-1093-7.
27. Huang, S.; Tao, X.; Yuan, S.; Zhang, Y.; Li, P.; Beilinson, H.A.; Zhang, Y.; Yu, W.; Pontarotti, P.; Escriva, H.; et al. Discovery of an Active RAG Transposon Illuminates the Origins of V(D)J Recombination. *Cell* 2016, *166*, 102–114, doi:10.1016/j.cell.2016.05.032.
28. Yakovenko, I.; Agronin, J.; Smith, L.C.; Oren, M. Guardian of the Genome: An Alternative RAG/Transib Co-Evolution Hypothesis for the Origin of V(D)J Recombination. *Front. Immunol.* 2021, *12*, doi:10.3389/fimmu.2021.709165.
29. Atkinson, P.W. HAT Transposable Elements. *Microbiol. Spectr.* 2015, *3*, doi:10.1128/microbiolspec.mdna3-0054-2014.
30. Kapitonov, V. V.; Jurka, J. RAG1 Core and V(D)J Recombination Signal Sequences Were Derived from Transib Transposons. *PLoS Biol.* 2005, *3*, 0998–1011, doi:10.1371/journal.pbio.0030181.
31. Hickman, A.B.; Dyda, F. DNA Transposition at Work. *Chem. Rev.* 2016, *116*, 12758–12784, doi:10.1021/acs.chemrev.6b00003.
32. Geballa-Koukoulas, K.; Andreani, J.; La Scola, B.; Blanc, G. The Kaumoebavirus LCC10 Genome Reveals a Unique Gene Strand Bias among "Extended Asfarviridae." *Viruses* 2021, *13*, doi:10.3390/v13020148.
33. Baroudy, B.M.; Venkatesan, S.; Moss, B. Incompletely Base-Paired Flip-Flop Terminal Loops Link the Two DNA Strands of the Vaccinia Virus Genome into One Uninterrupted Polynucleotide Chain. *Cell* 1982, *28*, 315–324, doi:10.1016/0092-8674(82)90349-X.
34. Zhang, Y.; Strasser, P.; Grabherr, R.; Van Etten, J.L. Hairpin Loop Structure at the Termini of the Chlorella Virus PBCV-1 Genome. *Virology* 1994, *202*, 1079–1082, doi:10.1006/viro.1994.1444.
35. González, A.; Talavera, A.; Almendral, J.M.; Viñuela, E. Hairpin Loop Structure of African Swine Fever Virus DNA. *Nucleic Acids Res.* 1986, *14*, 6835–6844, doi:10.1093/nar/14.17.6835.
36. Kobryn, K.; Chaconas, G. Hairpin Telomere Resolvases. *Microbiol. Spectr.* 2014, *2*, doi:10.1128/microbiolspec.mdna3-0023-2014.
37. Inagaki, K.; Ma, C.; Storm, T.A.; Kay, M.A.; Nakai, H. The Role of DNA-PKcs and Artemis in Opening Viral DNA Hairpin Termini in Various Tissues in Mice. *J. Virol.* 2007, *81*, 11304–11321, doi:10.1128/jvi.01225-07.
38. Garcia, A.D.; Aravind, L.; Koonin, E. V.; Moss, B. Bacterial-Type DNA Holliday Junction Resolvases in Eukaryotic Viruses. *Proc. Natl. Acad. Sci. U. S. A.* 2000, *97*, 8926–8931, doi:10.1073/pnas.150238697.
39. Moshous, D.; Callebaut, I.; De Chasseval, R.; Corneo, B.; Cavazzana-Calvo, M.; Le Deist, F.; Tezcan, I.; Sanal, O.; Bertrand, Y.; Philippe, N.; et al. Artemis, a Novel DNA Double-Strand Break Repair/V(D)J Recombination Protein, Is Mutated in Human Severe Combined Immune Deficiency. *Cell* 2001, *105*, 177–186, doi:10.1016/S0092-8674(01)00309-9.
40. Van Der Burg, M.; IJspeert, H.; Verkaik, N.S.; Turul, T.; Wiegant, W.W.; Morotomi-Yano, K.; Mari, P.O.; Tezcan, I.; Chen, D.J.; Zdzienicka, M.Z.; et al. A DNA-PKcs Mutation in a Radiosensitive T-B- SCID Patient Inhibits Artemis Activation and Nonhomologous End-Joining. *J. Clin. Invest.* 2009, *119*, 91–98, doi:10.1172/JCI37141.
41. Povirk, L.F.; Zhou, T.; Zhou, R.; Cowan, M.J.; Yannone, S.M. Processing of 3′-Phosphoglycolate-Terminated DNA Double Strand Breaks by Artemis Nuclease. *Journal of Biological Chemistry* 2007, *282*, 3547–3558, doi:10.1074/jbc.M607745200.
42. Muszewska, A.; Steczkiewicz, K.; Stepniewska-Dziubinska, M.; Ginalski, K. Cut-and-Paste Transposons in Fungi with Diverse Lifestyles. *Genome Biol. Evol.* 2017, *9*, 3463–3477, doi:10.1093/gbe/evx261.
43. Deprá, M.; da Silva Valente, V.L.; Margis, R.; Loreto, E.L.S. The Hobo Transposon and Hobo-Related Elements Are Expressed as Developmental Genes in Drosophila. *Gene* 2009, *448*, 57–63, doi:10.1016/j.gene.2009.08.012.
44. Lees-Miller, J.P.; Cobban, A.; Katsonis, P.; Bacolla, A.; Tsutakawa, S.E.; Hammel, M.; Meek, K.; Anderson, D.W.; Lichtarge, O.; Tainer, J.A.; et al. Uncovering DNA-PKcs Ancient Phylogeny, Unique Sequence Motifs and Insights for Human Disease. *Prog. Biophys. Mol. Biol.* 2021, *163*, 87–108, doi:10.1016/j.pbiomolbio.2020.09.010.

 doi:10.20944/preprints202604.1756.v1

45. Colot, V.; Haedens, V.; Rossignol, J.-L. Extensive, Nonrandom Diversity of Excision Footprints Generated by Ds-like Transposon Ascot-1 Suggests New Parallels with V(D)J Recombination. *Mol. Cell. Biol.* 1998, *18*, 4337–4346, doi:10.1128/mcb.18.7.4337.
46. Huang, S.; Tao, X.; Yuan, S.; Zhang, Y.; Li, P.; Beilinson, H.A.; Zhang, Y.; Yu, W.; Pontarotti, P.; Escriva, H.; et al. Discovery of an Active RAG Transposon Illuminates the Origins of V(D)J Recombination. *Cell* 2016, *166*, 102–114, doi:10.1016/j.cell.2016.05.032.
47. Bienstock, R.J.; Beard, W.A.; Wilson, S.H. Phylogenetic Analysis and Evolutionary Origins of DNA Polymerase X-Family Members. *DNA Repair (Amst).* 2014, *22*, 77–88, doi:10.1016/j.dnarep.2014.07.003.
48. Loc'h, J.; Delarue, M. Terminal Deoxynucleotidyltransferase: The Story of an Untemplated DNA Polymerase Capable of DNA Bridging and Templated Synthesis across Strands. *Curr. Opin. Struct. Biol.* 2018, *53*, 22–31, doi:10.1016/j.sbi.2018.03.019.
49. Andrade, P.; Martín, M.J.; Juárez, R.; De Saro, F.L.; Blanco, L. Limited Terminal Transferase in Human DNA Polymerase Mu Defines the Required Balance between Accuracy and Efficiency in NHEJ. *Proc. Natl. Acad. Sci. U. S. A.* 2009, *106*, 16203–16208, doi:10.1073/pnas.0908492106.
50. Reeves, R. High Mobility Group (HMG) Proteins: Modulators of Chromatin Structure and DNA Repair in Mammalian Cells. *DNA Repair (Amst).* 2015, *36*, 122–136, doi:10.1016/j.dnarep.2015.09.015.
51. Huang, S.; Tao, X.; Yuan, S.; Zhang, Y.; Li, P.; Beilinson, H.A.; Zhang, Y.; Yu, W.; Pontarotti, P.; Escriva, H.; et al. Discovery of an Active RAG Transposon Illuminates the Origins of V(D)J Recombination. *Cell* 2016, *166*, 102–114, doi:10.1016/j.cell.2016.05.032.
52. Zayed, H.; Izsvák, Z.; Khare, D.; Heinemann, U.; Ivics, Z. The DNA-Bending Protein HMGB1 Is a Cellular Cofactor of Sleeping Beauty Transposition. *Nucleic Acids Res.* 2003, *31*, 2313–2322, doi:10.1093/nar/gkg341.
53. Pang, Y.; Xiao, R.; Liu, X.; Li, Q. Identification and Characterization of the Lamprey High-Mobility Group Box 1 Gene. *PLoS One* 2012, *7*, doi:10.1371/journal.pone.0035755.
54. Nemazee, D. Mechanisms of Central Tolerance for B Cells. *Nat. Rev. Immunol.* 2017, *17*, 281–294, doi:10.1038/nri.2017.19.
55. Inglesfield, S.; Cosway, E.J.; Jenkinson, W.E.; Anderson, G. Rethinking Thymic Tolerance: Lessons from Mice. *Trends Immunol.* 2019, *40*, 279–291, doi:10.1016/j.it.2019.01.011.
56. Okamura, K.; Dijkstra, J.M.; Tsukamoto, K.; Grimholt, U.; Wiegertjes, G.F.; Kondow, A.; Yamaguchi, H.; Hashimoto, K. Discovery of an Ancient MHC Category with Both Class I and Class II Features. *Proc. Natl. Acad. Sci. U. S. A.* 2021, *118*, doi:10.1073/pnas.2108104118.
57. Paganini, J.; Pontarotti, P. Search for MHC/TCR-Like Systems in Living Organisms. *Front. Immunol.* 2021, *12*, doi:10.3389/fimmu.2021.635521.
58. Hein, W.R. Structural and Functional Evolution of the Extracellular Regions of T Cell Receptors. *Semin. Immunol.* 1994, *6*, 361–372, doi:10.1006/smim.1994.1046.
59. Flajnik, M.F. Evidence of G.O.D.'s Miracle: Unearthing a RAG Transposon. *Cell* 2016, *166*, 11–12.
60. Akula, S.; Mohammadamin, S.; Hellman, L. Fc Receptors for Immunoglobulins and Their Appearance during Vertebrate Evolution. *PLoS One* 2014, *9*, doi:10.1371/journal.pone.0096903.
61. Harly, C.; Robert, J.; Legoux, F.; Lantz, O. Γδ, NKT and MAIT Cells during Evolution: Redundancy or Specialized Functions?, doi:10.4049/jimmunol.2200105.
62. Swann, J.B.; Holland, S.J.; Petersen, M.; Pietsch, T.W.; Boehm, T. The Immunogenetics of Sexual Parasitism. *Science* 2020, *369*, doi:10.1126/science.aaz9445.
63. Dijkstra, J.M.; Yamaguchi, T.; Grimholt, U. Conservation of Sequence Motifs Suggests That the Nonclassical MHC Class I Lineages CD1/PROCR and UT Were Established before the Emergence of Tetrapod Species. *Immunogenetics* 2018, *70*, 459–476, doi:10.1007/s00251-017-1050-2.
64. Boudinot, P.; Mondot, S.; Jouneau, L.; Teyton, L.; Lefranc, M.P.; Lantz, O. Restricting Nonclassical MHC Genes Coevolve with TRAV Genes Used by Innate-like T Cells in Mammals. *Proc. Natl. Acad. Sci. U. S. A.* 2016, *113*, E2983–E2992, doi:10.1073/pnas.1600674113.
65. Deseke, M.; Prinz, I. Ligand Recognition by the Γδ TCR and Discrimination between Homeostasis and Stress Conditions. *Cell. Mol. Immunol.* 2020, *17*, 914–924, doi:10.1038/s41423-020-0503-y.
66. Adams, E.J.; Gu, S.; Luoma, A.M. Human Gamma Delta T Cells: Evolution and Ligand Recognition. *Cell. Immunol.* 2015, *296*, 31–40, doi:10.1016/j.cellimm.2015.04.008.

 doi:10.20944/preprints202604.1756.v1

67. Le Nours, J.; Gherardin, N.A.; Ramarathinam, S.H.; Awad, W.; Wiede, F.; Gully, B.S.; Khandokar, Y.; Praveena, T.; Wubben, J.M.; Sandow, J.J.; et al. A Class of Γδ T Cell Receptors Recognize the Underside of the Antigen-Presenting Molecule MR1. *Science* 2019, *366*, 1522–1527, doi:10.1126/science.aav3900.
68. Rice, M.T.; von Borstel, A.; Chevour, P.; Awad, W.; Howson, L.J.; Littler, D.R.; Gherardin, N.A.; Le Nours, J.; Giles, E.M.; Berry, R.; et al. Recognition of the Antigen-Presenting Molecule MR1 by a Vδ3+ Γδ T Cell Receptor. *Proc. Natl. Acad. Sci. U. S. A.* 2021, *118*, doi:10.1073/pnas.2110288118.
69. Pellicci, D.G.; Koay, H.F.; Berzins, S.P. Thymic Development of Unconventional T Cells: How NKT Cells, MAIT Cells and Γδ T Cells Emerge. *Nat. Rev. Immunol.* 2020, *20*, 756–770, doi:10.1038/S41577-020-0345-Y;SUBJMETA.
70. Vance, R.E.; Tanamachi, D.M.; Hanke, T.; Raulet, D.H. Cloning of a Mouse Homolog of CD94 Extends the Family of C-Type Lectins on Murine Natural Killer Cells. *Eur. J. Immunol.* 1997, *27*, 3236–3241, doi:10.1002/eji.1830271222.
71. MacLachlan, B.J.; Sullivan, L.C.; Brooks, A.G.; Rossjohn, J.; Vivian, J.P. Structure of the Murine CD94-NKG2A Receptor in Complex with Qa-1b Presenting an MHC-I Leader Peptide. *FEBS J.* 2024, *291*, 1530–1544, doi:10.1111/febs.17050.
72. Sim, M.J.W.; Brennan, P.; Wahl, K.L.; Lu, J.; Rajagopalan, S.; Sun, P.D.; Long, E.O. Innate Receptors with High Specificity for HLA Class I-Peptide Complexes. *Sci. Immunol.* 2023, *8*, doi:10.1126/sciimmunol.adh1781.
73. Futas, J.; Horin, P. Natural Killer Cell Receptor Genes in the Family Equidae: Not Only Ly49. *PLoS One* 2013, *8*, doi:10.1371/journal.pone.0064736.
74. Wilhelm, B.T.; Mager, D.L. Rapid Expansion of the Ly49 Gene Cluster in Rat. *Genomics* 2004, *84*, 218–221, doi:10.1016/j.ygeno.2004.01.010.
75. Anderson, S.K.; Ortaldo, J.R.; McVicar, D.W. The Ever-Expanding Ly49 Gene Family: Repertoire and Signaling. *Immunol. Rev.* 2001, *181*, 79–89, doi:10.1034/j.1600-065X.2001.1810106.x.
76. Schwartz, J.C.; Gibson, M.S.; Heimeier, D.; Koren, S.; Phillippy, A.M.; Bickhart, D.M.; Smith, T.P.L.; Medrano, J.F.; Hammond, J.A. The Evolution of the Natural Killer Complex; a Comparison between Mammals Using New High-Quality Genome Assemblies and Targeted Annotation. *Immunogenetics* 2017, *69*, 255–269, doi:10.1007/s00251-017-0973-y.
77. Carrillo-Bustamante, P.; Keşmir, C.; de Boer, R.J. The Evolution of Natural Killer Cell Receptors. *Immunogenetics* 2016, *68*, 3–18, doi:10.1007/s00251-015-0869-7.
78. Abi-Rached, L.; Parham, P. Natural Selection Drives Recurrent Formation of Activating Killer Cell Immunoglobulin-like Receptor and Ly49 from Inhibitory Homologues. *J. Exp. Med.* 2005, *201*, 1319–1332, doi:10.1084/jem.20042558.
79. Deng, L.; Cho, S.; Malchiodi, E.L.; Kerzic, M.C.; Dam, J.; Mariuzza, R.A. Molecular Architecture of the Major Histocompatibility Complex Class I-Binding Site of Ly49 Natural Killer Cell Receptors. *Journal of Biological Chemistry* 2008, *283*, 16840–16849, doi:10.1074/jbc.M801526200.
80. Dam, J.; Guan, R.; Natarajan, K.; Dimasi, N.; Chlewicki, L.K.; Kranz, D.M.; Schuck, P.; Margulies, D.H.; Mariuzza, R.A. Variable MHC Class I Engagement by Ly49 Natural Killer Cell Receptors Demonstrated by the Crystal Structure of Ly49C Bound to H-2K(b). *Nat. Immunol.* 2003, *4*, 1213–1222, doi:10.1038/ni1006.
81. Franksson, L.; Sundbäck, J.; Achour, A.; Bernlind, J.; Glas, R.; Kärre, K. Peptide Dependency and Selectivity of the NK Cell Inhibitory Receptor Ly-49C., doi:10.1002/(SICI)1521-4141(199909)29:09<2748::AID-IMMU2748>3.0.CO;2-C.
82. Ma, J.; Ayres, C.M.; Hellman, L.M.; Devlin, J.R.; Baker, B.M. Dynamic Allostery Controls the Peptide Sensitivity of the Ly49C Natural Killer Receptor. *Journal of Biological Chemistry* 2021, *296*, doi:10.1016/j.jbc.2021.100686.
83. Guselnikov, S. V.; Taranin, A. V. Unraveling the LRC Evolution in Mammals: IGSF1 and A1BG Provide the Keys. *Genome Biol. Evol.* 2019, *11*, 1586–1601, doi:10.1093/gbe/evz102.
84. Schwartz, J.C.; Gibson, M.S.; Heimeier, D.; Koren, S.; Phillippy, A.M.; Bickhart, D.M.; Smith, T.P.L.; Medrano, J.F.; Hammond, J.A. The Evolution of the Natural Killer Complex; a Comparison between Mammals Using New High-Quality Genome Assemblies and Targeted Annotation. *Immunogenetics* 2017, *69*, 255–269, doi:10.1007/s00251-017-0973-y.

 doi:10.20944/preprints202604.1756.v1

85. Petrie, E.J.; Clements, C.S.; Lin, J.; Sullivan, L.C.; Johnson, D.; Huyton, T.; Heroux, A.; Hoare, H.L.; Beddoe, T.; Reid, H.H.; et al. CD94-NKG2A Recognition of Human Leukocyte Antigen (HLA)-E Bound to an HLA Class I Leader Sequence. *J. Exp. Med.* 2008, *205*, 725–735, doi:10.1084/jem.20072525.
86. MacLachlan, B.J.; Sullivan, L.C.; Brooks, A.G.; Rossjohn, J.; Vivian, J.P. Structure of the Murine CD94-NKG2A Receptor in Complex with Qa-1b Presenting an MHC-I Leader Peptide. *FEBS J.* 2024, *291*, 1530–1544, doi:10.1111/febs.17050.
87. Joly, E.; Rouillon, V. The Orthology of HLA-E and H2-Qa1 Is Hidden by Their Concerted Evolution with Other MHC Class I Molecules. *Biol. Direct* 2006, *1*, doi:10.1186/1745-6150-1-2.
88. Guselnikov, S. V.; Taranin, A. V. Unraveling the LRC Evolution in Mammals: IGSF1 and A1BG Provide the Keys. *Genome Biol. Evol.* 2019, *11*, 1586–1601, doi:10.1093/gbe/evz102.
89. Adams, E.J.; Luoma, A.M. The Adaptable Major Histocompatibility Complex (MHC) Fold: Structure and Function of Nonclassical and MHC Class I-like Molecules. *Annu. Rev. Immunol.* 2013, *31*, 529–561, doi:10.1146/annurev-immunol-032712-095912.
90. Radaev, S.; Rostro, B.; Brooks, A.G.; Colonna, M.; Sun, P.D. Conformational Plasticity Revealed by the Cocrystal Structure of NKG2D and Its Class I MHC-like Ligand ULBP3. *Immunity* 2001, *15*, 1039–1049, doi:10.1016/S1074-7613(01)00241-2.
91. Li, P.; McDermott, G.; Strong, R.K. Crystal Structures of RAE-1β and Its Complex with the Activating Immunoreceptor NKG2D. *Immunity* 2002, *16*, 77–86, doi:10.1016/S1074-7613(02)00258-3.
92. Li, P.; Morris, D.L.; Willcox, B.E.; Steinle, A.; Spies, T.; Strong, R.K. Complex Structure of the Activating Immunoreceptor NKG2D and Its MHC Class I-like Ligand MICA. *Nat. Immunol.* 2001, *2*, 443–451, doi:10.1038/87757.
93. Schwartz, J.C.; Gibson, M.S.; Heimeier, D.; Koren, S.; Phillippy, A.M.; Bickhart, D.M.; Smith, T.P.L.; Medrano, J.F.; Hammond, J.A. The Evolution of the Natural Killer Complex; a Comparison between Mammals Using New High-Quality Genome Assemblies and Targeted Annotation. *Immunogenetics* 2017, *69*, 255–269, doi:10.1007/s00251-017-0973-y.
94. Belov, K.; Sanderson, C.E.; Deakin, J.E.; Wong, E.S.W.; Assange, D.; McColl, K.A.; Gout, A.; De Bono, B.; Barrow, A.D.; Speed, T.P.; et al. Characterization of the Opossum Immune Genome Provides Insights into the Evolution of the Mammalian Immune System. *Genome Res.* 2007, *17*, 982–991, doi:10.1101/gr.6121807.
95. Vivier, E.; Tomasello, E.; Paul, P. Lymphocyte Activation via NKG2D: Towards a New Paradigm in Immune Recognition? *Curr. Opin. Immunol.* 2002, *14*, 306–311, doi:10.1016/S0952-7915(02)00337-0.
96. Kondo, M.; Maruoka, T.; Otsuka, N.; Kasamatsu, J.; Fugo, K.; Hanzawa, N.; Kasahara, M. Comparative Genomic Analysis of Mammalian NKG2D Ligand Family Genes Provides Insights into Their Origin and Evolution. *Immunogenetics* 2010, *62*, 441–450, doi:10.1007/s00251-010-0438-z.
97. Van Der Kraan, L.E.; Wong, E.S.W.; Lo, N.; Ujvari, B.; Belov, K. Identification of Natural Killer Cell Receptor Genes in the Genome of the Marsupial Tasmanian Devil (Sarcophilus Harrisii). *Immunogenetics* 2013, *65*, 25–35, doi:10.1007/s00251-012-0643-z.
98. Oganesyan, V.; Oganesyan, N.; Terzyan, S.; Qu, D.; Dauter, Z.; Esmon, N.L.; Esmon, C.T. The Crystal Structure of the Endothelial Protein C Receptor and a Bound Phospholipid. *Journal of Biological Chemistry* 2002, *277*, 24851–24854, doi:10.1074/jbc.C200163200.
99. Rao, L.V.M.; Esmon, C.T.; Pendurthi, U.R. Endothelial Cell Protein C Receptor: A Multiliganded and Multifunctional Receptor. *Blood* 2014, *124*, 1553–1562, doi:10.1182/blood-2014-05-578328.
100. Willcox, C.R.; Pitard, V.; Netzer, S.; Couzi, L.; Salim, M.; Silberzahn, T.; Moreau, J.F.; Hayday, A.C.; Willcox, B.E.; Déchanet-Merville, J. Cytomegalovirus and Tumor Stress Surveillance by Binding of a Human Γδ T Cell Antigen Receptor to Endothelial Protein C Receptor. *Nat. Immunol.* 2012, *13*, 872–879, doi:10.1038/ni.2394.
101. Bennett, M.J.; Lebrón, J.A.; Bjorkman, P.J. Crystal Structure of the Hereditary Haemochromatosis Protein HFE Complexed with Transferrin Receptor. *Nature* 2000, *403*, 46–53, doi:10.1038/47417.
102. LM, S.; AJ, C.; Pj, B. Crystal Structure of Human ZAG, a Fat-Depleting Factor Related to MHC Molecules. *Science* 1999, *283*, 1914–1919, doi:10.1126/science.283.5409.1914.
103. Burmeister, W.P.; Huber, A.H.; Bjorkman, P.J. Crystal Structure of the Complex of Rat Neonatal Fc Receptor with Fc. *Nature* 1994, *372*, 379–383, doi:10.1038/372379a0.

 doi:10.20944/preprints202604.1756.v1

104. Martin, W.L.; West, A.P.; Gan, L.; Bjorkman, P.J. Crystal Structure at 2.8 Å of an FcRn/Heterodimeric Fc Complex: Mechanism of PH-Dependent Binding. *Mol. Cell* 2001, *7*, 867–877, doi:10.1016/S1097-2765(01)00230-1.
105. Baker, K.; Qiao, S.W.; Kuo, T.; Kobayashi, K.; Yoshida, M.; Lencer, W.I.; Blumberg, R.S. Immune and Non-Immune Functions of the (Not so) Neonatal Fc Receptor, FcRn. *Semin. Immunopathol.* 2009, *31*, 223–236, doi:10.1007/s00281-009-0160-9.
106. Dijkstra, J.M.; Yamaguchi, T.; Grimholt, U. Conservation of Sequence Motifs Suggests That the Nonclassical MHC Class I Lineages CD1/PROCR and UT Were Established before the Emergence of Tetrapod Species. *Immunogenetics* 2018, *70*, 459–476, doi:10.1007/s00251-017-1050-2.
107. Edholm, E.S.; Banach, M.; Rhoo, K.H.; Pavelka, M.S.; Robert, J. Distinct MHC Class I-like Interacting Invariant T Cell Lineage at the Forefront of Mycobacterial Immunity Uncovered in Xenopus. *Proc. Natl. Acad. Sci. U. S. A.* 2018, *115*, E4023–E4031, doi:10.1073/pnas.1722129115.
108. Banach, M.; Edholm, E.S.; Robert, J. Exploring the Functions of Nonclassical MHC Class Ib Genes in Xenopus Laevis by the CRISPR/Cas9 System. *Dev. Biol.* 2017, *426*, 261–269, doi:10.1016/j.ydbio.2016.05.023.
109. Dunkelberger, J.R.; Song, W.C. Complement and Its Role in Innate and Adaptive Immune Responses. *Cell Res.* 2010, *20*, 34–50, doi:10.1038/cr.2009.139.
110. Nonaka, M. Evolution of the Complement System. *Subcell. Biochem.* 2014, *80*, 31–43, doi:10.1007/978-94-017-8881-6_3.
111. Matsushita, M. The Complement System of Agnathans. *Front. Immunol.* 2018, *9*, doi:10.3389/fimmu.2018.01405.
112. Gong, T.; Liu, L.; Jiang, W.; Zhou, R. DAMP-Sensing Receptors in Sterile Inflammation and Inflammatory Diseases. *Nat. Rev. Immunol.* 2020, *20*, 95–112, doi:10.1038/s41577-019-0215-7.
113. Le Nours, J.; Gherardin, N.A.; Ramarathinam, S.H.; Awad, W.; Wiede, F.; Gully, B.S.; Khandokar, Y.; Praveena, T.; Wubben, J.M.; Sandow, J.J.; et al. A Class of Γδ T Cell Receptors Recognize the Underside of the Antigen-Presenting Molecule MR1. *Science* 2019, *366*, 1522–1527, doi:10.1126/science.aav3900.
114. Herrmann, T.; Karunakaran, M.M. Phosphoantigen Recognition by Vγ9Vδ2 T Cells. *Eur. J. Immunol.* 2024, *54*, doi:10.1002/eji.202451068.
115. Afrache, H.; Gouret, P.; Ainouche, S.; Pontarotti, P.; Olive, D. The Butyrophilin (BTN) Gene Family: From Milk Fat to the Regulation of the Immune Response. *Immunogenetics* 2012, *64*, 781–794, doi:10.1007/s00251-012-0619-z.
116. Fulford, T.S.; Soliman, C.; Castle, R.G.; Rigau, M.; Ruan, Z.; Dolezal, O.; Seneviratna, R.; Brown, H.G.; Hanssen, E.; Hammet, A.; et al. Vγ9Vδ2 T Cells Recognize Butyrophilin 2A1 and 3A1 Heteromers. *Nat. Immunol.* 2024, *25*, 1355–1366, doi:10.1038/s41590-024-01892-z.
117. Hintz, M.; Reichenberg, A.; Altincicek, B.; Bahr, U.; Gschwind, R.M.; Kollas, A.K.; Beck, E.; Wiesner, J.; Eberl, M.; Jomaa, H. Identification of (E)-4-Hydroxy-3-Methyl-but-2-Enyl Pyrophosphate as a Major Activator for Human Gammadelta T Cells in Escherichia Coli. *FEBS Lett.* 2001, *509*, 317–322, doi:10.1016/S0014-5793(01)03191-X.
118. Tanaka, Y.; Tanaka, Y.; Bloom, B.R.; Morita, C.T.; Brenner, M.B.; Nieves, E. Natural and Synthetic Non-Peptide Antigens Recognized by Human Gamma Delta T Cells. *Nature* 1995, *375*, 155–158, doi:10.1038/375155a0.
119. Rast, J.P.; Anderson, M.K.; Strong, S.J.; Luer, C.; Litman, R.T.; Litman, G.W. α, β, γ, and δ t Cell Antigen Receptor Genes Arose Early in Vertebrate Phylogeny. *Immunity* 1997, *6*, 1–11, doi:10.1016/S1074-7613(00)80237-X.