中国科学: 生命科学 2024 年 第 54 卷 第 9 期: 1581 ~ 1594
SCIENTIA SINICA Vitae lifecn.scichina.com
《中国科学》杂志社 SCIENCE CHINA PRESS
评 述 中国知名大学及研究院所专栏 中山大学生命科学学院专辑

抗原受体基因多样性产生机制的起源与演化

元少春*, 黄子文, 蔡振西, 黄盛丰, 徐安龙*

中山大学生命科学学院, 广东省药用功能基因研究重点实验室, 南海海洋生物技术教育部工程研究中心, 中山大学进化与合成生物学研究中心, 广州 510275

* 联系人, E-mail: yuanshch@mail.sysu.edu.cn; lssxal@mail.sysu.edu.cn

收稿日期: 2024-05-09; 接受日期: 2024-06-17; 网络版发表日期: 2024-08-21

国家自然科学基金(批准号: 32170888)、广州市基础与应用基础研究项目(批准号: 2024A04J3326)和高校基本科研业务费(批准号: 23yxqntd001)资助

摘要 动物免疫系统根据其应答方式及存在形式的不同, 可以分为天然免疫和适应性免疫. 由于适应性免疫的关键器官/分子(比如胸腺//MHC)等只在有颌类脊椎动物中被发现, 因此被认为是生命演化到高级阶段的体现, 关于适应性免疫系统的起源一直是演化生物学和免疫学研究共同关注的重要问题. 本文重点梳理了国内外学者近年来在抗原受体基因多样性产生机制及其演化等方面所取得的进展, 尤其是在V(D)J重排机制的转座子起源及*RAGL*转座子的宿主驯化机制上. 本文还总结了无颌类脊椎动物VLR受体及数个进化阶段特有的免疫多样性受体的研究进展. 这些研究在完善经典免疫学理论的同时, 也为医疗保健领域带来崭新的思路与技术.

关键词 抗原受体基因, RAG, V(D)J重排, 演化

免疫系统由免疫器官、免疫细胞和免疫分子组成. 免疫应答按作用方式的不同可分为天然免疫和适应性免疫. 天然免疫主要利用种系基因编码的模式识别受体(pattern recognition receptors, PRRs)识别病原相关分子模式(pathogen-associated molecular patterns, PAMPs)或者损伤相关分子模式(damage-associated molecular patterns, DAMPs), 启动快速的免疫应答. 适应性免疫则是由T/B淋巴细胞通过V(D)J重排机制产生具有高度多样性且抗原特异性的识别受体, 进而通过B细胞介导的体液免疫和T细胞介导的细胞免疫发挥作用. 天然免疫被认为在物种间保守存在, 而适应性免疫曾一度被认为仅存在于有颌类脊椎动物中. 直到2004年, 美国埃默里大学的Max D. Cooper教授团队[1]在无颌类脊椎动物七鳃鳗中发现了与传统概念不同的适应性免疫新机制——由淋巴样细胞(lymphocyte like cells, LLCs)通过类似于基因转换(gene conversion, GC)的方式产生可变淋巴细胞受体(variable lymphocyte receptor, VLR)进行抗原识别的免疫应答机制. VLR的发现不仅揭示了一种进化阶段特有的适应性免疫机制, 还将适应性免疫的起源往前推进了一亿年. 同期, 中山大学生命科学学院徐安龙教授团队则利用处于无脊椎动物向脊椎动物过渡的物种——被誉为“活化石”的文昌鱼为主要研究对象, 通过整合免疫应答机制与分子进化概念对免疫系统的起源与演化进行了跨物种的比

引用格式: 元少春, 黄子文, 蔡振西, 等. 抗原受体基因多样性产生机制的起源与演化. 中国科学: 生命科学, 2024, 54: 1581–1594
Yuan S C, Huang Z W, Cai Z X, et al. Origin and evolution of antigen receptor gene diversity (in Chinese). Sci Sin Vitae, 2024, 54: 1581–1594, doi: 10.1360/SSV-2024-0149

© 2024 《中国科学》杂志社 www.scichina.com
https://engine.scichina.com/doi/10.1360/SSV-2024-0149

较免疫学研究, 发现了脊椎动物祖先“天然免疫基因大爆炸”的现象, 并证实了V(D)J重排机制的转座子起源假说, 为诺贝尔生理学或医学奖获得者利根川进(Susumu Tonegawa)提出的抗体重排机制的转座子起源假说提供了直接的证据. 本文将就抗原受体基因多样性产生机制的起源与演化, 物种特有的免疫识别多样性机制进行阐述, 并在此基础上, 探讨其应用的潜在发展方向.

1 V(D)J重排机制的起源及演化研究

1.1 V(D)J重排的基本过程

哺乳动物抗原受体基因的多样性由V(D)J重排机制介导. V(D)J重排是发生在体细胞基因组水平的DNA重组事件, 由重组激活基因RAG(recombination-activating gene)复合物所介导. RAG复合物是由RAG1和RAG2两个非同源分子组成的异源四聚体, 它们可特异性识别抗原受体基因座上的V, D和J基因片段侧翼的重组信号序列(recombination signal sequence, RSS)[2]. RSS序列由一段保守的七核苷酸序列(CACAGTG, 称为heptamer), 一段保守的富含A的九核苷酸序列(ACAAAAACC, 称为nonamer)以及中间非保守的间隔序列(spacer)组成[3](图1A). Spacer的长度常为12或23 bp, 因此RSS序列被分别命名为12RSS及23RSS. 一个侧翼是12RSS的基因片段通常只能与一个侧翼是23RSS的基因片段连接, 这在免疫学上被称为抗体重排的12/23规则(12/23 rule)[4].

在V(D)J重排过程中, RAG复合物首先识别并捕获RSS序列, 并在RSS序列和编码区序列的连接处切割DNA链, 形成pair complex(PC). RAG切割DNA后暴露的3′-OH通过亲核攻击, 导致互补链断裂, 产生一对5′磷酸化的RSS平末端DNA序列(称为signal ends)和一对3′带有发夹结构的编码区DNA(称为coding ends)[5]. 随后, 编码区DNA在内切酶Artemis的作用下, 打开发夹结构, 并通过非同源末端修复连接途径(non-homologous end joining, NHEJ)连接在一起, 形成可正常编码抗原受体基因的coding joints[6]; 而带有RSS的signal ends则自身环化形成signal joints, 在细胞分裂的过程中逐渐丢失(图1B).

哺乳动物RAG1和RAG2均包含一个核心的结构域(图1C). RAG1的核心结构域具有介导DNA底物切割及重组的活性, 包含NBD(nonamer binding domain, NBD), DDBD(dimerization and DNA binding domain, DDBD), RNH(RNaseH, RNH), ZnB(zinc finger-binding, ZnB)和CTD(C-terminal domain, CTD)结构域(图1C). NBD结构域主要负责与RSS的nonamer序列结合并促进RAG1的二聚化[7]. DDBD, RNH和ZnB结构域则组成RAG1的催化核心, 可结合RSS中的heptamer序列并执行核酸内切酶功能. 催化核心结构域中的DDE(D600, D708和E962)基序对RAG1的活性至关重要(图1C)[8]. RAG1的非核心序列在生理条件下具有不可或缺的功能, 如RAG1 N端非核心序列可与Ku70/80相互作用[9], 非核心序列中的RING结构域可对组蛋白H3进行泛素化, 提高细胞内V(D)J重排的效率[10,11].

RAG2的核心结构域由六个kelch-like motif组成(图1C), 其空间结构像六个叶片的螺旋桨, 其中第六个叶片与RAG1相互作用[12]. RAG2的非核心序列包括酸性铰链区(acidic hinge, AH)、PHD结构域(plant homeodomain, PHD)和C端结构域(图1C). 与RAG1一样, RAG2非核心序列具有不可或缺的功能, 其中酸性铰链区对V(D)J重排中NHEJ修复途径的使用至关重要[13], 从而在保证基因组稳定性的同时进一步增加coding joint连接处的序列多样性. RAG2的PHD结构域可与磷酸肌醇和H3K4me3(histone H3 trimethylated at lysine 4)结合, 在V(D)J重排过程中发挥重要的作用, 因为破坏RAG2对H3K4me3识别使得V(D)J重排的效率显著下降[14,15].

1.2 V(D)J重排的调控

*RAG*基因的表达异常会促进染色体的不稳定性, 导致细胞癌变等, 因此*RAG*基因在淋巴细胞中的表达受到严格的调控[16]. *RAG1*和*RAG2*基因在基因组上以tail-to-tail的方向相邻排布[17,18]. 在*RAG*基因内部还存在一个演化上保守的基因*NWC*[19], 其可能参与*RAG*基因的表达调控[20,21]. *RAG*在淋巴细胞的发育过程中会出现两次表达量的升高和降低[22]. 以B细胞为例, *RAG*基因会在重链重排时表达量上升, 重链重排后表达量下降; 轻链重排阶段*RAG*基因经历新一轮的表达量上调, 在轻链重排后*RAG*基因表达量又会下降[23]. 因此, *RAG*基因的表达受到多种顺式作用元件和反式作用因子以及细胞周期的协同调控, 使得RAG复合物的活性仅维持在V(D)J重排所需的特定阶段中. 比如, 在

https://engine.scichina.com/doi/10.1360/SSV-2024-0149

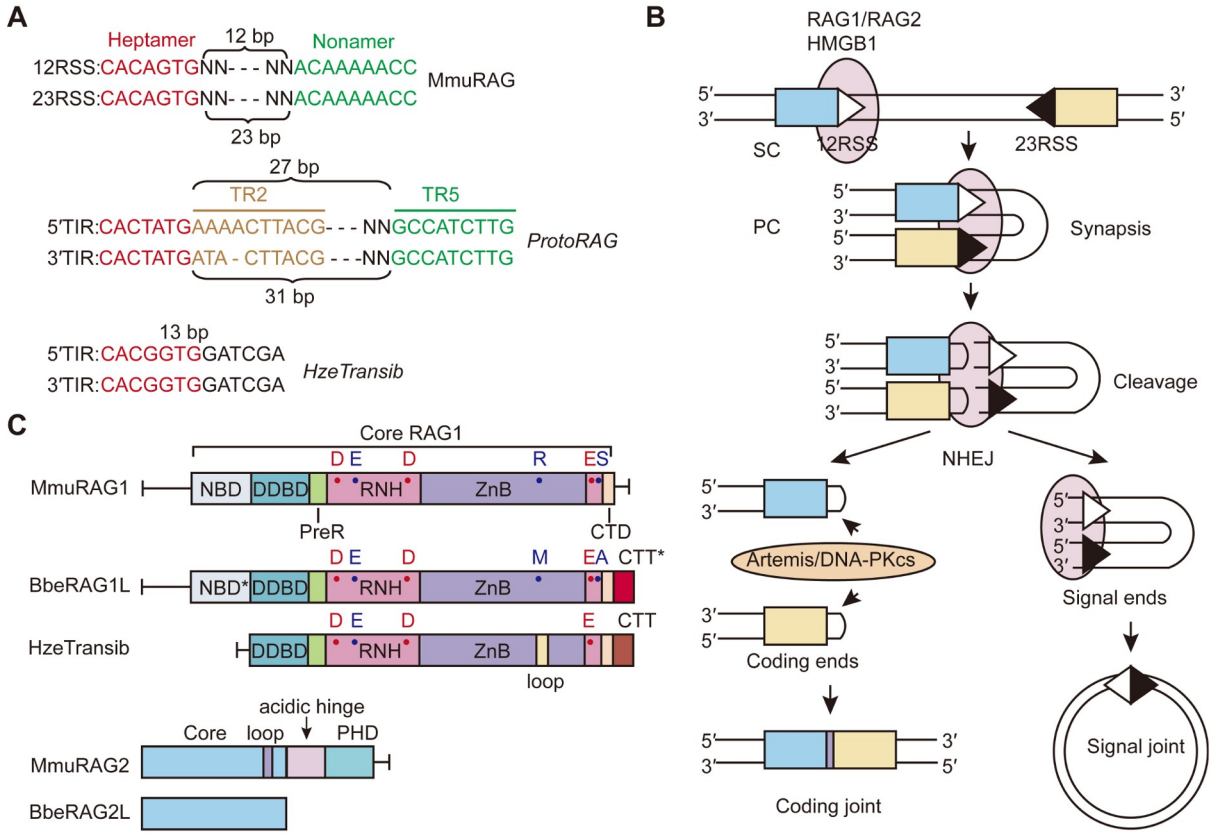

图 1　RAG介导的V(D)J重排机制. A: 小鼠12/23RSS, *ProtoRAG*的5/3TIR和*HzeTransib*的5/3TIR序列比较. 红色的字母表示heptamer序列, 绿色的字母代表nonamer序列; B: RAG介导的V(D)J重排过程; C: 小鼠RAG1/2, 文昌鱼BbeRAG1/2L和棉铃虫HzeTransib蛋白结构域比较. 红色圆点表示保守的DDE基序, 紫色圆点表示RAG被宿主驯化后突变的氨基酸位点. Bbe, 文昌鱼; Hze, 棉铃虫; Mmu, 小鼠

Figure 1　RAG-mediated V(D)J recombination. A: Sequences of 12/23 RSS and TIRs of *ProtoRAG* and *HzeTransib*. Heptamers and nonamers of TIRs and RSSs are depicted in red and green, respectively; B: schematic of V(D)J recombination mediated by the RAG complex; C: protein domain organization of mouse RAG compared with BbeRAGL and HzeTransib. The red and purple dots represent the evolutionarily conserved DDE motif and the residues mutated during the domestication of RAG, respectively. Bbe, *Branchiostoma belcheri*; Hze, *Helicoverpa zea*; Mmu, *Mus musculus*

*RAG1*和*RAG2*基因附近发现的D3, Ep, Erag和ASE (antisilencing element)等顺式作用元件可以调控*RAG*基因的表达[24,25]; 转录因子如FoxO1, Foxp1, Gata3, Pax5和Ikaros等可协同调节*RAG*基因在淋巴细胞发育过程中的表达[26]. 此外, RAG2的翻译后修饰调控使得RAG复合物的活性被限制在细胞周期G1期. 细胞周期相关激酶CDK1可磷酸化RAG2蛋白的490位苏氨酸[27], 磷酸化的RAG2被E3泛素连接酶Skp2-SCF引导进入蛋白酶体降解途径[28]. RAG2被降解后, RAG1虽然可在细胞核内聚集, 但介导V(D)J重排的活性极低[29].

不仅RAG的表达可影响V(D)J重排, RAG本身的酶活偏好性也可影响V(D)J重排的效率. TCRβ基因座由V-23RSS, 12RSS-D-23RSS和12RSS-J基因片段组成, 根据12/23规则, 其V基因片段可直接和J基因片段进行重排. 然而在生理状况下极少见TCRβ的V片段直接和J片段重排, 这是因为RAG通过其对RSS序列选择的偏好性, 使得TCRβ基因倾向于D和J的重排, 这被称为超12/23规则[30]. 此外, RAG对RSS侧翼的序列也有切割偏好性, 进而影响V(D)J重排的效率[31,32]. 最近的研究表明, RAG通过CTCF复合物介导的环挤压结构介导V(D)J重排的发生, 从染色质水平解析了RAG调控抗体重链(IgH)重排的过程[33].

与哺乳动物*RAG*基因的表达受到严格调控不同, 低等脊椎动物*RAG*基因的表达可能具有更高的自由度, 如斑马鱼和非洲爪蟾的*RAG*基因可在其生殖腺中

https://engine.scichina.com/doi/10.1360/SSV-2024-0149

表达[34,35]，另有研究也报道斑马鱼卵巢细胞中可能发生RAG复合物介导的切割[36]. 这些研究提示，RAG复合物的表达调控机制在演化过程中逐渐完善，并最终被限定在淋巴细胞特定的发育阶段.

1.3 V(D)重排的起源与演化研究

在适应性免疫应答的形成过程中，由RAG复合物介导的V(D)J重排机制起着异常关键的作用. 由于V(D)J重排在经典免疫学教科书中被认为仅存在于有颌类脊椎动物中，在低等生物中寻找*RAG1*及*RAG2*的同源基因，并阐述其和V(D)J重排之间的关系，则成为追溯适应性免疫系统起源的关键.

由于V(D)J重排机制与原核生物中的DNA转座机制非常相似，当V(D)J重排机制被发现时，日裔美籍免疫学家Susumu Tonegawa团队[3]即创造性地提出: V(D)J重排可能起源于DNA转座子的转座插入事件. 随后，免疫学界在研究V(D)J重排的分子机制时，也在积极寻找支持“V(D)J重排起源于DNA转座子”的证据(图2). 首先, 在1989~1990年期间, 美国著名生物家David Baltimore团队[17,18]发现了介导V(D)J重排的两个核心分子，RAG1和RAG2，并揭示了它们可通过识别12/23RSS序列介导抗原受体基因多样性的产生. 由于*RAG1*和*RAG2*在基因组中呈对向分布，并且编码区多为单外显子结构，呈现出类似于原核或者病毒类基因的结构，进一步提示了V(D)J重排的转座子起源(图3). 到20世纪末，David G. Schatz团队[37]通过体外转座实验发现，哺乳动物RAG1/RAG2酶复合物可以介导V(D)J重组中被切割下来的RSS信号末端进行非特异位点的转座, 为V(D)J重排的转座子起源假说提供了关键的实验证据. Dreyfus和Gelfand[38]通过序列分析发现哺乳动物RAG蛋白复合物所识别的RSS类似于DNA转座子的TIR(terminal inverted repeat)序列. Kapitonov等人[39]则在无脊椎动物中发现了*Transib*转座子所编码的转座酶与RAG1核心序列具有同源性. Hencken等人[40]通过体外的切割和转座实验发现，Transib转座酶作用于TIR的模式与脊椎动物RAG复合物作用于RSS底物的模式类似, 均需形成末端闭合的中间DNA产物, 进一步提示 RAG1的核心结构域和RSS序列可能起源于*Transib*转座子. 2006年, Jonathan P. Rast团队[41]在棘皮动物紫海胆(*Strongylocentrotus purpuratus*)基因组上发现了成对出现的，并呈tail-to-tail转录的*RAG1-like*和*RAG2-like*基因(图3). 虽然紫海胆*RAG-like*基因在基因组上的分布形式及蛋白结构域组成已与哺乳动物*RAG*基因高度相似，但在紫海胆*RAG-like*基因的附近并未发现转座子特征序列, 比如TSD(target site duplication)及TIR序列. 因此，尽管上述研究都为“V(D)J重排起源

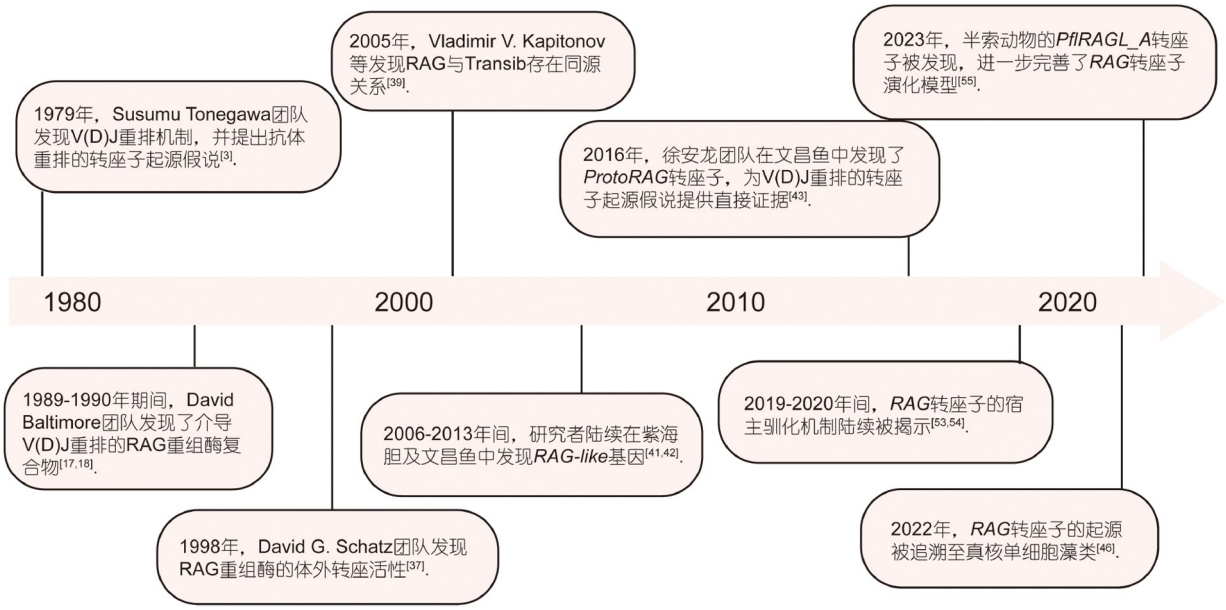

图 2 *RAG*基因起源与演化的研究历程

Figure 2 Timeline of the studies on the origin and evolution of *RAG*

https://engine.scichina.com/doi/10.1360/SSV-2024-0149

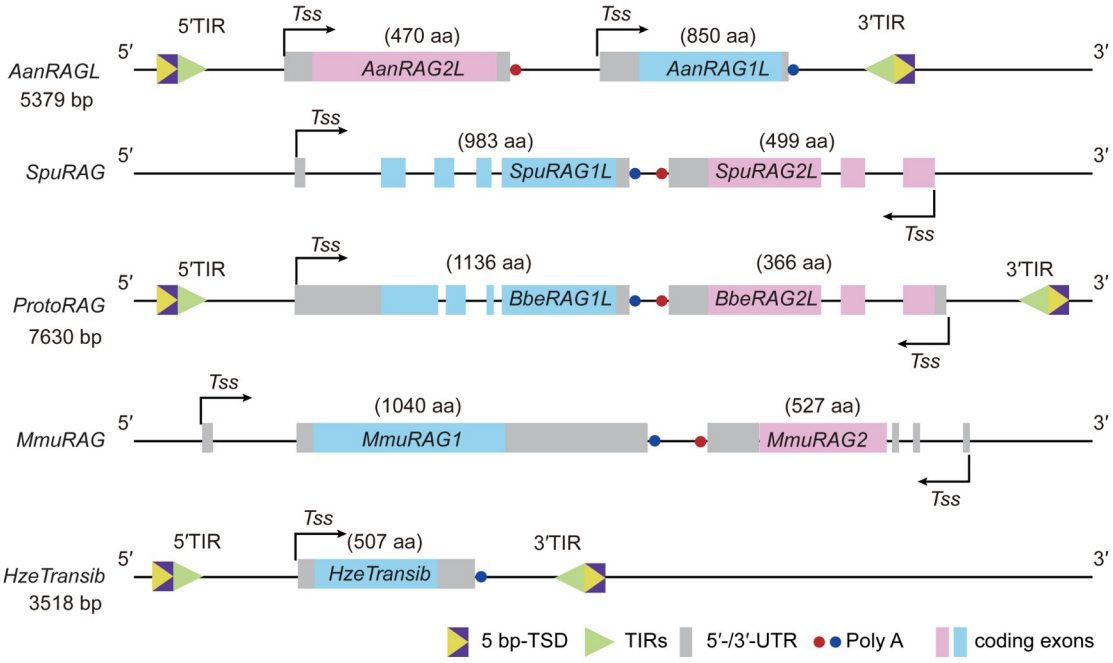

图 3　各物种*RAG*基因在基因组上的分布示意图. Aan: *Aureococcus anophagefferens*; Spu: *Strongylocentrotus purpuratus*; Mmu: *Mus musculus*; Hze: *Helicoverpa zea*; Tss: 转录起始位点; TIR: 末端反向重复序列; TSD: 靶位点重复序列

Figure 3　Schematic of the genomic organization of mouse *RAG* and *RAG-like* genes in several representative species. Aan, *Aureococcus anophagefferens*; Spu, *Strongylocentrotus purpuratus*; Mmu, *Mus musculus*; Hze, *Helicoverpa zea*; Tss, transcription start site; TIR, terminal inverted repeat; TSD, target site duplication

于RAG转座子”的假说提供了一定的支持, 但由于所有脊椎动物编码RAG蛋白的基因座及棘皮动物海胆编码RAG-like(RAGL)蛋白的基因座都不具备转座子的基本特性, 使得“V(D)J重排机制起源于RAG转座子”的假说仍然缺乏直接的有力证据.

2014年, 中国科学院生物化学与细胞生物学研究所的刘小龙教授和中山大学的徐安龙教授团队[42]在国家重点基础研究发展计划(973计划)的支持下, 在佛罗里达文昌鱼(*Branchiostoma floridae*, bf)基因组中发现一段与脊椎动物RAG1核心序列同源的序列, 称为bfRAG1L. 功能实验证明, bfRAG1L与小鼠RAG2结合后可识别RSS序列并介导V(D)J重排, 首次在无脊椎动物中发现RAG1的同源分子具有与哺乳动物RAG1相似的活性[42]. 随后, 在2016年, 徐安龙团队[43]通过自主构建中国白氏文昌鱼(*Branchiostoma belcheri*)基因组草图, 并对文昌鱼全基因组进行深度的转座子挖掘, 发现了一类新型的DNA转座子元件, 称为*ProtoRAG*. 与海胆和哺乳动物*RAG*基因的连锁结构一致, *ProtoRAG*转座子包含一对tail-to-tail转录的*RAG1-like*和*RAG2-like*基因. 重要的是, 在*ProtoRAG*上下游还包含一对反向重复序列TIR, 以及一对5 bp的TSD序列(图3). 机制研究表明, 与哺乳动物RAG复合物介导的V(D)J重排过程高度相似, *ProtoRAG*所编码的RAG1L及RAG2L蛋白也需要形成复合物, 才能识别自身的TIR序列并介导转座及宿主DNA的重连接[43]. *ProtoRAG*的发现及其转座机制的研究为“V(D)J重排机制起源于RAG转座子”的假说提供了最直接的证据. 随后, 研究人员相继在后口动物、原口动物和刺胞动物中均发现完整的*RAGL*转座子序列[44,45]. 在2022年的一项研究中, 徐安龙团队[46]通过搜索近68万个物种的基因组序列, 在单细胞真核生物抑食金球藻(*Aureococcus anophagefferens*)基因组中发现了迄今最为古老且完整的*RAGL*转座子, 从而将V(D)J重排的起源追溯至单细胞真核生物(图3). 在该研究中, 研究人员通过对Transib, RAGL及RAG蛋白进行系统的演化分析, 提出了现存的*Transib*转座子和*RAGL*转座子均起源于细菌*Transib*转座子的观点.

https://engine.scichina.com/doi/10.1360/SSV-2024-0149

1.4 *RAGL*转座子的宿主驯化

尽管Transib与RAG1L及RAG1核心蛋白序列的相似性很低, 但三者都包含保守的DDE催化核心基序. 而且类似于RSS, *Transib*转座子TIR序列也含有保守的七聚体(heptamer), 特别是前三个高度保守的碱基CAC[39]. 不同的是, *Transib*转座子中缺乏编码*RAG2-like*的基因, 其转座过程完全依赖于Transib转座酶(图3)[40]. 目前唯一报道具有转座能力的*Transib*转座子是来源于棉铃虫(*Helicoverpa zea*)的*HzeTransib*[40,47,48]. HzeTransib转座酶与RAG1核心序列相比缺乏N端的NBD结构域, 相对应地, *HzeTransib*的TIR序列中也缺乏nonamer-like序列(图1A和C), 这说明Transib与RAG在底物识别的模式上有所差别. 通过解析HzeTransib与DNA底物结合的晶体结构, 耶鲁大学David G. Schatz团队[48]发现, 相比于RAG1, HzeTransib在结构上缺乏与RAG2相互作用的区域, 但HzeTransib具有一小段独特的loop结构(328~334 aa), 该loop在转座的过程中能够稳定靶DNA, 被认为是Transib缺乏RAG2的补偿.

虽然转座子的移动是物种演化的重要驱动力[49], 但转座子的活跃跳动会造成宿主基因组的不稳定. 因此, 在与宿主长期的博弈过程中, 转座子常常会被驯化为对宿主有利的基因或处于“石化”的失活状态[50]. 为揭示*ProtoRAG*为何能够在长期的演化过程中保持活性, 徐安龙团队首先揭示了*ProtoRAG*是一个转录自激活的转座子. 进一步通过对能够结合*ProtoRAG* TIR序列的反式作用因子进行筛选, 该团队发现文昌鱼YY1(Ying Yang 1)因子可抑制*ProtoRAG* 的转录及转座活性, 从而在维持*ProtoRAG* 低活性的同时促进宿主基因组的稳定性, 提示YY1的调控作用可能是*ProtoRAG*在亿万年演化过程中能够存活至今的原因之一[51]. 该团队还发现, 与RSS的功能类似, *ProtoRAG* TIR序列的碱基保守性及序列长度对*ProtoRAG*的活性至关重要, 尤其是TIR序列中保守的CAC三碱基, 其中任意一个碱基的突变都会严重影响*ProtoRAG*的转座活性[52]. 另一个关于*ProtoRAG*转座子功能演化的重要研究来自David G. Schatz/Yong Xiong/徐安龙团队的合作[53], 他们通过对文昌鱼RAGL转座酶复合物的结构进行解析, 发现哺乳动物RAG1蛋白E649和S963位点的存在使得其介导双边切割(coupled cleavage)的能力显著提升, 且RAG1蛋白R848位点以及RAG2蛋白C端酸性铰链区抑制了哺乳动物RAG复合物介导转座的活性[53]. David G. Schatz团队[54]还进一步发现, RAG2核心结构域中的loop(333~342 aa)也对RAG复合物的转座活性具有显著的抑制作用. 值得注意的是, 在识别DNA底物的过程中, 哺乳动物RAG1通过其NBD结构域识别RSS中的nonamer序列, 而文昌鱼RAG1L在识别底物的过程中, 则由其CTT*结构域扮演底物结合的角色[53]. 尽管目前关于这种DNA底物识别模式的变化机制暂不清晰, 但这些研究均提示, 由RAGL转座酶向RAG重组酶的演化过程中, E649, R848及S963等位点的适应性进化, RAG2酸性铰链区及RAG1 CTT功能域的获得/丢失是RAGL转座酶被宿主驯化为RAG重组酶的关键.

近期, David G. Schatz和徐安龙团队[44,55]进一步合作, 在半索动物柱头虫(*Ptychodera flava*)中发现了A类*RAGL*转座子(*PflRAGL_A*), 进一步为RAG的演化提供了线索(图4). 系统发育分析表明, A类*RAGL*转座子与脊椎动物中的RAG具有更近的亲缘关系. *PflRAGL_A*已具备一些RAG的特征, 比如PflRAG1L_A存在与哺乳动物RAG1一样的E649位点, 同时PflRAG2L_A获得了酸性铰链区. 有趣的是, PfRAG1L_A具有与文昌鱼RAG1L相似的CTT*结构域及与哺乳动物RAG1相似的NBD结构域, 其TIR也含有nonamer-like的序列, 这些特征均表明*PflRAGL_A*可能处于*RAGL*转座子向RAG重组酶演变的中间状态[55]. 后续对*PflRAGL_A*活性和功能的研究将进一步揭示RAGL转座酶向RAG重组酶演化的机制.

在上述研究的基础上, 本团队[46]提出了V(D)J重排机制的起源与演化模型(图4). 该模型认为, 真核生物的*Tranisb*转座子通过捕获一个*RAG2-like*的基因形成原始的*RAGL*转座子. 在随后的进化过程中, *Transib*转座子主要发生水平基因转移(horizontal gene transfer, HGT), 而*RAGL*转座子在多细胞动物中垂直演化, 并在无脊椎动物向脊椎动物演化的阶段, 出现了与脊椎动物RAG具有更高亲缘关系的A类*RAGL*转座子[44]. A类*RAGL*转座子通过整合至一段原始的Ig V基因序列内部, 形成原始的V和J基因片段. 随后, RAG1蛋白E649, R848和S963等位点的适应性进化, RAG2L蛋白酸性铰链序列的获得, 抑制了RAGL转座酶复合物的转座活性, 提高了RAG复合物对底物双边切割的能力, 从而促进了RAG重组酶复合物的宿主驯化.

https://engine.scichina.com/doi/10.1360/SSV-2024-0149

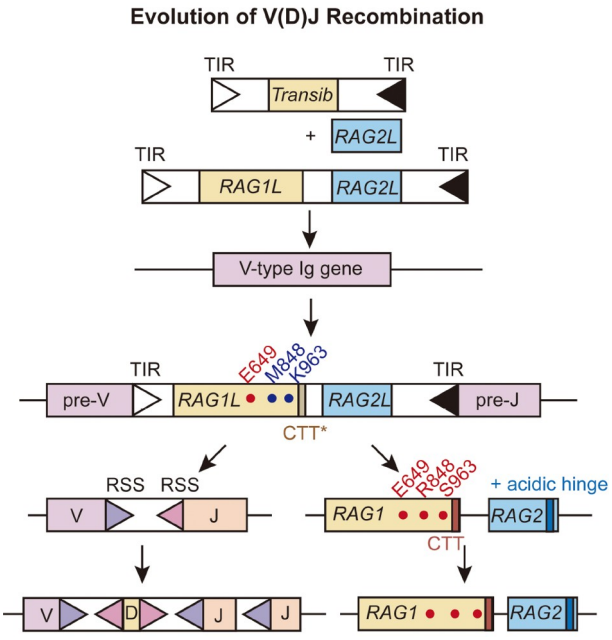

图 4　V(D)J重排机制的起源与演化模型
Figure 4　Model for the origin and evolution of RAG-mediated V(D)J recombination

2　其他具有多样性的免疫识别受体

2.1　无颌类脊椎动物的VLR重排

20世纪60年代，研究者在无颌类脊椎动物七鳃鳗中发现适应性免疫的特征，但一直未发现其拥有与有颌类脊椎动物相似的抗原识别受体[56]．直至2004年，Max D. Cooper团队[1]在七鳃鳗中发现了一种新型的富含LRR(leucin rich repeat)基序的适应性免疫受体VLR．不同于有颌类脊椎动物中介导适应性免疫反应的免疫球蛋白分子，成熟的VLR蛋白由高度多样化的LRR组成．目前，VLR分子可分为VLRA，VLRB，VLRC，VLRD和VLRE五类，其中表达VLRB的细胞被认为是类B淋巴细胞，而表达VLRA/C/D/E细胞则被认为是类T淋巴细胞[57]．与IgM类似，VLRB分子分为膜结合型和分泌型两种形式，膜结合型VLRB以GPI(glycosylphosphatidylinositol)锚定而非跨膜的方式实现膜定位，分泌型的VLRB则类似于Ig抗体，呈多聚体形式，是体液免疫中的主要成分[58]．虽有研究表明VLR的基因结构与脊椎动物特异性的血小板糖蛋白复合物可能具有共同的起源[59]，但关于VLR分子的起源仍缺少系统的研究．

与V(D)J重排类似，VLR的重排同样具有等位基因排斥的现象，即每个淋巴细胞仅表达单个功能性的VLR分子[60]．但不同于*Ig*和*TCR*基因座，*VLR*胚系基因座仅包括保守的LRRNT和LRRCT两部分恒定区，中间插有一段约2 kb的非编码DNA序列．胚系基因的两侧包含大量排列和方向均不相同的LRR供体盒(LRR cassettes)，且无明显重组信号(图5)[61]．已有的研究表明，为了生成高度多样化的抗原受体基因，*VLR*基因座可能是通过一种类似于基因转换的机制来启动侧翼LRR供体盒的初始复制，再通过“复制选择”机制以短同源区域作为锚定点逐步替换恒定区之间的非编码DNA，实现功能性组装[61]．目前认为VLR的多样性除了以不同组合复制各种LRR片段形成嵌合LRR供体盒来实现外，还可以通过在“复制选择”过程中使用不同的短同源性区域进行延伸，产生序列不同的重组*VLR*基因．早在2009年，就有研究指出VLR的重排可能由胞苷脱氨酶CDA1(cytidine deaminase 1)及CDA2(cytidine deaminase 2)分子介导[58]．七鳃鳗*CDA1*及*CDA2*基因属于AID/APOBEC超家族成员，该家族成员AID(activation-induced cytidine deaminase)分子在有颌类脊椎动物中介导了Ig的类别转换及体细胞超突变．近期，通过在七鳃鳗中敲除*CDA2*基因，Thomas Boehm团队[62]发现，类B细胞谱系的*VLRB*基因组装失败，而类T细胞

https://engine.scichina.com/doi/10.1360/SSV-2024-0149

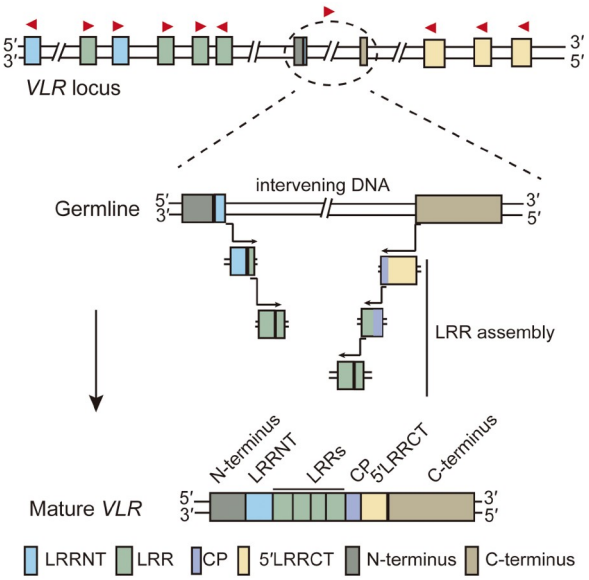

图 5 日本七鳃鳗*VLR*基因通过“复制选择”机制进行重排. 红色箭头表示基因转录的方向, 黑色箭头表示重排的*VLR*基因组装的方向, 绿色、蓝色、紫色和黄色的方框分别表示LRR (富含亮氨酸的重复序列), LRRNT (氨基端的LRR), CP (连接肽)和5′LRRCT (羧基端的LRR)

Figure 5 Assembly of the *VLRs* of *Lampetra japonica* by a process involving copy choice. The red arrows indicate the transcriptional direction of gene segments; the black arrows represent the direction of DNA synthesis in *VLR* assembly. The green, blue, purple, and yellow boxes indicate LRR (leucine-rich repeat), LRRNT (amino-terminal LRR), CP (connecting peptide), and 5′-LRRCT (carboxy-terminal LRR), respectively

谱系的*VLRA*和*VLRC*基因组装正常, 从而明晰了VLRB的体细胞多样化由七鳃鳗CDA2分子介导, 提出在脊椎动物两个姐妹分支中独立进化的抗原受体基因组装和多样化的趋同进化机制[62].

虽然VLRB的重排已被证实与CDA2相关, 研究人员也对VLR重排使用的LRR供体盒进行了使用频率和规律分析, 但是VLRB重排的具体机制以及VLRB的信号传递方式仍然不清楚. 另外, 由于CDA分子在七鳃鳗中发生了种属特有的基因扩增, 并在个体间出现了拷贝数差异, 比如*CDA1*基因在某些七鳃鳗个体中丢失[63,64], 提示类T细胞谱系的VLRA/C/D/E重排可能并非由CDA1分子介导. 关于七鳃鳗VLRA/C/D/E重排由哪类分子通过哪种机制介导, 其细胞内的信号传递机制等也均处于未知的状态.

2.2 其他物种特有的多样化免疫识别分子

虽然在无脊椎动物中并未发现经典的抗原识别受体基因, 但节肢动物Dscam(Down’s syndrome cell adhesion molecule), 蜗牛FREP(fibrinogen-related protein), 海胆185/333及文昌鱼VCBP(V region-containing chitin binding protein)等多样性免疫识别受体的发现, 提示在生物演化中存在物种特有的免疫识别及应答机制.

在果蝇等节肢动物中发现的Dscam分子属于免疫球蛋白超家族(Ig superfamily, IgSF), 其可通过选择性剪切产生高达38016种的mRNA异构体[65,66]. 虽然Dscam可变剪接的功能多样性主要集中在神经细胞的“自我”与“非我”识别中, 但也有多个研究表明Dscam可能是节肢动物特异性免疫反应的基础[67,68]. 比如表达Dscam的昆虫血细胞对细菌有弱的吞噬作用, 而分泌型的Dscam则被推测可结合入侵病原体[69]. 2004年在蜗牛中发现的FREP家族也被认为具有高度多样性并在宿主防御系统中起作用[70]. FREP由一个或两个IgSF结构域和C端的FBG(fibrinogen domain)结构域组成, 其IgSF结构域可通过重组和点突变产生高度的多样性[70]. 关于海胆185/333基因家族的研究表明, 该家族蛋白在成体或胚胎受到免疫刺激后表达[71]. 在快速基因重组、RNA编辑和/或低保真率转录、翻译后修饰的共同作用下, 海胆体腔细胞中的185/333蛋白可产生高度多样性[72]. 为应对不同的PAMP, Sp185/333还可通过转录产生不同的185/333基因组合[73]. 文昌鱼VCBP的研究表明其可通过基因互换和点突变的方式产生高度多样性[74]. VCBP的N端包含两个Ig V结构域, C端则含有几丁质结合结构域(chitin binding domain, CBD). CBD结构域的出现提示VCBP可能参与球蛋白-外源凝集素的免疫应答[75]. 在细菌感染过程中, VCBP在肠道的特异表达也从侧面证明了其可参与免疫应答[76].

除上述可产生高度多样性并含有Ig结构域的免疫识别受体外, 大量PRR及其配体在过去30年间相继被发现. 与BCR/TCR不同的是, PRR是由种系基因编码的. 20世纪90年代, 果蝇Toll及哺乳动物TLR4(Toll like receptor 4)分子的发现[77,78], 在证实模式识别理论的同时, 也在免疫学界掀起了一股寻找PRR的热潮. 随后, TLRs及其配体, 病毒dsRNA识别受体RLR(RIG-I like receptor)[79], 介导炎症小体组装的NOD样受体(NOD-like receptor, NLR)[80], DNA识别受体cGAS(cyclic GMP-AMP synthase)[81], ZBP1(Z-DNA binding protein 1)[82]等相继被发现. 2006~2008年期间, 研究人员通过比较免疫学研究, 在无脊椎动物, 比如头索动物

https://engine.scichina.com/doi/10.1360/SSV-2024-0149

文昌鱼和棘皮动物海胆中发现了天然免疫基因普遍扩增的现象，被称为“天然免疫基因大爆发”[83,84]. 比如, 海胆基因组编码超过200个TLR分子及近200个NLR分子[84,85]; 文昌鱼中虽只有39个*TLR*基因[86], 但却拥有至少40个负责TLR信号转导的胞内TIR(Toll-interleukin-1 receptor)接头分子[83]. 此外，这些TIR接头分子还通过蛋白结构域的全新组合产生了多个新的信号传导分子[87]. 通过对数个天然免疫信号通路开展系统研究, 徐安龙团队进一步指出“天然免疫基因的扩增及结构域动态重组”为脊椎动物免疫系统的形成奠定了重要的分子基础, 比如TNFR(tumor necrosis factor receptor)结构域与DD(death domain)结构域的重组, 产生了死亡受体(death receptor, DR)并促进外源性细胞凋亡通路的出现[88]. 近期, 通过对果蝇cGLR(cGAS-like receptor)分子的研究发现, 与脊椎动物cGAS识别DNA并产生2′3′-cGAMP(cyclic GMP-AMP)不同，果蝇cGLR识别dsRNA并产生3′2′-cGAMP[89,90]. 2023年的两个研究表明，cGLR在刺胞动物和双壳纲动物中存在大量扩增, 并且其合成的第二信使分子也有较大差异[91,92]. 节肢动物及脊椎动物cGLR分子主要合成2′3′-cGAMP, 而软体动物cGLR除合成2′3′-cGAMP外，还可合成2′3′-cUA、3′3′-cUA和3′3′-cAA等[91]. 关于PRR的比较免疫学研究表明, 与适应性免疫相比, 天然免疫虽有更古老的起源，但其并非简单且静止，而是动态和高度复杂的. 因此，进一步在低等生物中挖掘具有高度多样化的免疫识别分子并阐明其作用机制仍然是比较与进化免疫学的前沿方向. 阐明这些多样性是如何产生的将提高人们对免疫系统演化的认识，深入研究引起新功能基因产生的基因重组机制也将给生物学研究带来新的启发.

3　展望

虽然V(D)J重排机制，*RAG*基因的起源与演化, VLR及其重排机制, PRR扩增及识别多样性等研究已取得长足的进展, 但其中仍存在许多有意思的问题待解决. 一是在进化过程中，*Transib*转座子在何时，以何种方式捕获*RAG2-like*基因并形成完整的*RAGL*转座子. 有研究者认为在早期的真核生物中, *RAG2-like*基因可能作为宿主的成员发挥类似抗毒素的功能, 从而抑制*RAG1-like*基因(毒素)引起的基因组不稳定性[93]. 由于RAG2包含Kelch和PHD结构域，通过AlphaFold结构预测和蛋白结构聚类等手段有望在单细胞真核生物中寻找*RAG2*的祖先基因，从而探究其在*RAGL*转座子形成及V(D)J重排机制演化中的作用. 二是*RAG*基因表达调控机制的演化. 有文献报道，斑马鱼*RAG1*和*RAG2*基因在其卵巢中表达, 非洲爪蟾的*RAG2*基因也在其卵巢中有表达[35]. 与之相对应的是, 在鲶形目斑点叉尾鮰(*Ictalurus punctatus*)抗体重链基因座中发现一个胚系基因组上已经重排完成的VDJ基因序列[94]. 在软骨鱼中，胚系基因组中也存在已经重排好的抗体基因[95]. 这些研究提示，在脊椎动物演化的早期，*RAG*可能在生殖细胞中泄露表达，在生殖细胞中发生重排的抗原受体基因可能与发育早期的免疫防御相关[96]. 因此, 跨物种比较*RAG*及*RAGL*基因的表达调控，有望揭示V(D)J重排的调控新机制. 三是*RAGL*及*Transib*转座子潜在应用前景的挖掘. 由于转座子是一类能够在基因组内或不同基因组间移动的遗传元件，它们在生物学研究中具有多种应用, 比如构建突变库、基因标记和多样性分析、转基因等. 来源于昆虫的*piggyBac*转座子系统由于转座效率高，宿主范围广，在哺乳动物细胞中也具有高转座活性等优点，目前已在动物基因组功能研究、基因转移、多能干细胞诱导、细胞治疗、基因治疗等领域得到了广泛应用[97,98]. 与*piggyBac*转座子常插入在TTAA靶点不同，*ProtoRAG*及*Transib*转座子偏好插入在富含GC的位点[43,48]. 因此，对*ProtoRAG*及*Transib*转座系统进行开发和优化, 有望为*piggyBac*等转座系统的应用提供有效的补充. 四是进一步挖掘物种特有的适应性免疫应答机制. 在过去20年中，昆虫Dscam和七鳃鳗VLR等免疫相关基因，以及近年来在细菌中发现的CRISPR-Cas系统等，均是进化阶段特有的适应性免疫应答机制. 如何从丰富的组学数据中挖掘进化阶段特有的免疫识别受体并阐明其作用机制，是另一个富有挑战性的研究工作. 五是揭示无颌类脊椎动物VLR的重排机制并拓展其应用. 作为无颌类脊椎动物特有的抗体分子，VLRB具有分子量小、稳定性高等优点. 然而，由于VLRB的重排机制尚不清楚，极大地限制了VLRB抗体的开发与应用. 因此, 在明晰VLR重排机制的过程中, 利用新型的抗体库构建及筛选策略, 有望在自身抗原检测和特殊抗体制备等领域拓展VLR的应用.

https://engine.scichina.com/doi/10.1360/SSV-2024-0149

参考文献

1 Pancer Z, Amemiya C T, Ehrhardt G R A, et al. Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. Nature, 2004, 430: 174–180

2 Kim M S, Lapkouski M, Yang W, et al. Crystal structure of the V(D)J recombinase RAG1-RAG2. Nature, 2015, 518: 507–511

3 Sakano H, Hüppi K, Heinrich G, et al. Sequences at the somatic recombination sites of immunoglobulin light-chain genes. Nature, 1979, 280: 288–294

4 Tonegawa S. Somatic generation of antibody diversity. Nature, 1983, 302: 575–581

5 Roth D B, Menetski J P, Nakajima P B, et al. V(D)J recombination: Broken DNA molecules with covalently sealed (hairpin) coding ends in scid mouse thymocytes. Cell, 1992, 70: 983–991

6 Ma Y, Pannicke U, Schwarz K, et al. Hairpin opening and overhang processing by an artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. Cell, 2002, 108: 781–794

7 Yin F F, Bailey S, Innis C A, et al. Structure of the RAG1 nonamer binding domain with DNA reveals a dimer that mediates DNA synapsis. Nat Struct Mol Biol, 2009, 16: 499–508

8 Swanson P C. The DDE motif in RAG-1 is contributed in *trans* to a single active site that catalyzes the nicking and transesterification steps of V(D)J recombination. Mol Cell Biol, 2001, 21: 449–458

9 Raval P, Kriatchko A N, Kumar S, et al. Evidence for Ku70/Ku80 association with full-length RAG1. Nucleic Acids Res, 2008, 36: 2060–2072

10 McMahan C J, Difilippantonio M J, Rao N, et al. A basic motif in the N-terminal region of RAG1 enhances V(D)J recombination activity. Mol Cell Biol, 1997, 17: 4544–4552

11 Grazini U, Zanardi F, Citterio E, et al. The RING domain of RAG1 ubiquitylates histone H3: a novel activity in chromatin-mediated regulation of V(D)J joining. Mol Cell, 2010, 37: 282–293

12 Ru H, Chambers M G, Fu T M, et al. Molecular mechanism of V(D)J recombination from synaptic RAG1-RAG2 complex structures. Cell, 2015, 163: 1138–1152

13 Coussens M A, Wendland R L, Deriano L, et al. RAG2's acidic hinge restricts repair-pathway choice and promotes genomic stability. Cell Rep, 2013, 4: 870–878

14 Matthews A G W, Kuo A J, Ramón-Maiques S, et al. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. Nature, 2007, 450: 1106–1110

15 Elkin S K, Ivanov D, Ewalt M, et al. A PHD finger motif in the C terminus of RAG2 modulates recombination activity. J Biol Chem, 2005, 280: 28701–28710

16 Papaemmanuil E, Rapado I, Li Y, et al. RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. Nat Genet, 2014, 46: 116–125

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

18 Schatz D G, Oettinger M A, Baltimore D. The V(D)J recombination activating gene, RAG-1. Cell, 1989, 59: 1035–1048

19 Cebrat M, Miazek A, Kisielow P. Identification of a third evolutionarily conserved gene within the RAG locus and its RAG1-dependent and -independent regulation. Eur J Immunol, 2005, 35: 2230–2238

20 Kisielow P, Miazek A, Cebrat M. *NWC*, a new gene within *RAG* locus: could it keep GOD under control? Int J Immunogenet, 2008, 35: 395–399

21 Sniezewski L, Janik S, Laszkiewicz A, et al. The evolutionary conservation of the bidirectional activity of the *NWC* gene promoter in jawed vertebrates and the domestication of the *RAG* transposon. Dev Comp Immunol, 2018, 81: 105–115

22 Wilson A, Held W, MacDonald H R. Two waves of recombinase gene expression in developing thymocytes. J Exp Med, 1994, 179: 1355–1360

23 Grawunder U, Leu T M J, Schatz D G, et al. Down-regulation of *RAG1* and *RAG2* gene expression in PreB cells after functional immunoglobulin heavy chain rearrangement. Immunity, 1995, 3: 601–608

24 Hsu L Y, Lauring J, Liang H E, et al. A conserved transcriptional enhancer regulates *RAG* gene expression in developing B cells. Immunity, 2003, 19: 105–117

25 Wei X C, Kishi H, Jin Z X, et al. Characterization of chromatin structure and enhancer elements for murine recombination activating gene-2. J

https://engine.scichina.com/doi/10.1360/SSV-2024-0149

Immunol, 2002, 169: 873–881
26 Kuo T C, Schlissel M S. Mechanisms controlling expression of the *RAG* locus during lymphocyte development. Curr Opin Immunol, 2009, 21: 173–178
27 Lin W C, Desiderio S. Regulation of V(D)J recombination activator protein RAG-2 by phosphorylation. Science, 1993, 260: 953–959
28 Jiang H, Chang F C, Ross A E, et al. Ubiquitylation of RAG-2 by Skp2-SCF links destruction of the V(D)J recombinase to the cell cycle. Mol Cell, 2005, 18: 699–709
29 Gan T, Wang Y, Liu Y, et al. RAG2 abolishes RAG1 aggregation to facilitate V(D)J recombination. Cell Rep, 2021, 37: 109824
30 Banerjee J K, Schatz D G. Synapsis alters RAG-mediated nicking at *Tcrb* recombination signal sequences: implications for the "beyond 12/23" rule. Mol Cell Biol, 2014, 34: 2566–2580
31 Yu K, Lieber M R. Mechanistic basis for coding end sequence effects in the initiation of V(D)J recombination. Mol Cell Biol, 1999, 19: 8094–8102
32 Gerstein R M, Lieber M R. Coding end sequence can markedly affect the initiation of V(D)J recombination. Genes Dev, 1993, 7: 1459–1469
33 Zhang Y, Zhang X, Dai H Q, et al. The role of chromatin loop extrusion in antibody diversification. Nat Rev Immunol, 2022, 22: 550–566
34 Willett C E, Cherry J J, Steiner L A. Characterization and expression of the recombination activating genes (*rag1* and *rag2*) of zebrafish. Immunogenetics, 1997, 45: 394–404
35 Greenhalgh P, Olesen C E, Steiner L A. Characterization and expression of recombination activating genes (*RAG-1* and *RAG-2*) in *Xenopus laevis*. J Immunol, 1993, 151: 3100–3110
36 Zhong H, Li Z, Lin S, et al. Initiation of V(D)J recombination in zebrafish (*Danio rerio*) ovaries. Mol Immunol, 2007, 44: 1784–1792
37 Agrawal A, Eastman Q M, Schatz D G. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. Nature, 1998, 394: 744–751
38 Dreyfus D, Gelfand E W. Comparative analysis of invertebrate Tc6 sequences that resemble the vertebrate V(D)J recombination signal sequences (RSS). Mol Immunol, 1999, 36: 481–488
39 Kapitonov V V, Jurka J. RAG1 core and V(D)J recombination signal sequences were derived from *Transib* transposons. PLoS Biol, 2005, 3: e181
40 Hencken C G, Li X, Craig N L. Functional characterization of an active Rag-like transposase. Nat Struct Mol Biol, 2012, 19: 834–836
41 Fugmann S D, Messier C, Novack L A, et al. An ancient evolutionary origin of the *Rag1/2* gene locus. Proc Natl Acad Sci USA, 2006, 103: 3728–3733
42 Zhang Y, Xu K, Deng A, et al. An amphioxus RAG1-like DNA fragment encodes a functional central domain of vertebrate core RAG1. Proc Natl Acad Sci USA, 2014, 111: 397–402
43 Huang S, Tao X, Yuan S, et al. Discovery of an active RAG transposon illuminates the origins of V(D)J recombination. Cell, 2016, 166: 102–114
44 Morales Poole J R, Huang S F, Xu A, et al. The RAG transposon is active through the deuterostome evolution and domesticated in jawed vertebrates. Immunogenetics, 2017, 69: 391–400
45 Martin E C, Vicari C, Tsakou-Ngouafo L, et al. Identification of RAG-like transposons in protostomes suggests their ancient bilaterian origin. Mobile DNA, 2020, 11: 17
46 Tao X, Huang Z, Chen F, et al. The RAG key to vertebrate adaptive immunity descended directly from a bacterial ancestor. Natl Sci Rev, 2022, 9: nwac073
47 Chen S, Li X. Molecular characterization of the first intact *Transib* transposon from *Helicoverpa zea*. Gene, 2008, 408: 51–63
48 Liu C, Yang Y, Schatz D G. Structures of a RAG-like transposase during cut-and-paste transposition. Nature, 2019, 575: 540–544
49 Feschotte C, Pritham E J. DNA transposons and the evolution of eukaryotic genomes. Annu Rev Genet, 2007, 41: 331–368
50 Sinzelle L, Izsvák Z, Ivics Z. Molecular domestication of transposable elements: from detrimental parasites to useful host genes. Cell Mol Life Sci, 2009, 66: 1073–1093
51 Liu S, Yuan S, Gao X, et al. Functional regulation of an ancestral RAG transposon *ProtoRAG* by a trans-acting factor YY1 in lancelet. Nat Commun, 2020, 11: 4515
52 Tao X, Yuan S, Chen F, et al. Functional requirement of terminal inverted repeats for efficient *ProtoRAG* activity reveals the early evolution of V(D)J recombination. Natl Sci Rev, 2020, 7: 403–417
53 Zhang Y, Cheng T C, Huang G, et al. Transposon molecular domestication and the evolution of the RAG recombinase. Nature, 2019, 569: 79–84
54 Zhang Y, Corbett E, Wu S, et al. Structural basis for the activation and suppression of transposition during evolution of the RAG recombinase.

https://engine.scichina.com/doi/10.1360/SSV-2024-0149

EMBO J, 2020, 39: e105857
55 Martin E C, Le Targa L, Tsakou-Ngouafo L, et al. Insights into RAG evolution from the identification of “missing link” family A RAGL transposons. Mol Biol Evol, 2023, 40: msad232
56 Finstad J, Good R A. The evolution of the immune response. J Exp Med, 1964, 120: 1151–1168
57 Das S, Boehm T, Holland S J, et al. Evolution of two distinct variable lymphocyte receptors in lampreys: VLRD and VLRE. Cell Rep, 2023, 42: 112933
58 Guo P, Hirano M, Herrin B R, et al. Dual nature of the adaptive immune system in lampreys. Nature, 2009, 459: 796–801
59 Rogozin I B, Iyer L M, Liang L, et al. Evolution and diversification of lamprey antigen receptors: evidence for involvement of an AID-APOBEC family cytosine deaminase. Nat Immunol, 2007, 8: 647–656
60 Kishishita N, Matsuno T, Takahashi Y, et al. Regulation of antigen-receptor gene assembly in hagfish. EMBO Rep, 2010, 11: 126–132
61 Nagawa F, Kishishita N, Shimizu K, et al. Antigen-receptor genes of the agnathan lamprey are assembled by a process involving copy choice. Nat Immunol, 2007, 8: 206–213
62 Morimoto R, O’Meara C P, Holland S J, et al. Cytidine deaminase 2 is required for *VLRB* antibody gene assembly in lampreys. Sci Immunol, 2020, 5: eaba0925
63 Holland S J, Berghuis L M, King J J, et al. Expansions, diversification, and interindividual copy number variations of AID/APOBEC family cytidine deaminase genes in lampreys. Proc Natl Acad Sci USA, 2018, 115: E3211–E3220
64 Chen Y, Luo L, Deng L, et al. New insights into the lineage-specific expansion and functional diversification of lamprey AID/APOBEC family. Front Immunol, 2022, 13: 822616
65 Schmucker D, Clemens J C, Shu H, et al. *Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. Cell, 2000, 101: 671–684
66 Watson F L, Puttmann-Holgado R, Thomas F, et al. Extensive diversity of Ig-superfamily proteins in the immune system of insects. Science, 2005, 309: 1874–1878
67 Chen B E, Kondo M, Garnier A, et al. The molecular diversity of Dscam is functionally required for neuronal wiring specificity in *Drosophila*. Cell, 2006, 125: 607–620
68 Dong Y, Taylor H E, Dimopoulos G. AgDscam, a hypervariable immunoglobulin domain-containing receptor of the *Anopheles gambiae* innate immune system. PLoS Biol, 2006, 4: e229
69 Li X J, Yang L, Li D, et al. Pathogen-specific binding soluble down syndrome cell adhesion molecule (Dscam) regulates phagocytosis via membrane-bound Dscam in crab. Front Immunol, 2018, 9: 801
70 Zhang S M, Adema C M, Kepler T B, et al. Diversification of Ig superfamily genes in an invertebrate. Science, 2004, 305: 251–254
71 Terwilliger D P, Buckley K M, Brockton V, et al. Distinctive expression patterns of 185/333 genes in the purple sea urchin, Strongylocentrotus purpuratus: an unexpectedly diverse family of transcripts in response to LPS, β-1,3-glucan, and dsRNA. BMC Mol Biol, 2007, 8: 16
72 Buckley K M, Smith L C. Extraordinary diversity among members of the large gene family, 185/333, from the purple sea urchin, *Strongylocentrotus purpuratus*. BMC Mol Biol, 2007, 8: 68
73 Dheilly N M, Nair S V, Smith L C, et al. Highly variable immune-response proteins (185/333) from the sea urchin, *Strongylocentrotus purpuratus*: proteomic analysis identifies diversity within and between individuals. J Immunol, 2009, 182: 2203–2212
74 Dishaw L J, Leigh B, Cannon J P, et al. Gut immunity in a protochordate involves a secreted immunoglobulin-type mediator binding host chitin and bacteria. Nat Commun, 2016, 7: 10617
75 Dishaw L J, Giacomelli S, Melillo D, et al. A role for variable region-containing chitin-binding proteins (VCBPs) in host gut-bacteria interactions. Proc Natl Acad Sci USA, 2011, 108: 16747–16752
76 Cannon J P, Haire R N, Litman G W. Identification of diversified genes that contain immunoglobulin-like variable regions in a protochordate. Nat Immunol, 2002, 3: 1200–1207
77 Lemaitre B, Nicolas E, Michaut L, et al. The dorsoventral regulatory gene cassette *spätzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults. Cell, 1996, 86: 973–983
78 Poltorak A, He X, Smirnova I, et al. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in *Tlr4* gene. Science, 1998, 282: 2085–2088
79 Yoneyama M, Kikuchi M, Natsukawa T, et al. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate

https://engine.scichina.com/doi/10.1360/SSV-2024-0149

antiviral responses. Nat Immunol, 2004, 5: 730–737
80 Martinon F, Burns K, Tschopp J. The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-beta. Mol Cell, 2002, 10: 417–426
81 Sun L, Wu J, Du F, et al. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. Science, 2013, 339: 786–791
82 Takaoka A, Wang Z C, Choi M K, et al. DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. Nature, 2007, 448: 501–505
83 Huang S, Yuan S, Guo L, et al. Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. Genome Res, 2008, 18: 1112–1126
84 Hibino T, Loza-Coll M, Messier C, et al. The immune gene repertoire encoded in the purple sea urchin genome. Dev Biol, 2006, 300: 349–365
85 Rast J P, Smith L C, Loza-Coll M, et al. Genomic insights into the immune system of the sea urchin. Science, 2006, 314: 952–956
86 Yuan S, Huang S, Zhang W, et al. An amphioxus TLR with dynamic embryonic expression pattern responses to pathogens and activates NF-κB pathway via MyD88. Mol Immunol, 2009, 46: 2348–2356
87 Peng J, Tao X, Li R, et al. Novel Toll/IL-1 receptor homologous region adaptors act as negative regulators in amphioxus TLR signaling. J Immunol, 2015, 195: 3110–3118
88 Yuan S, Liu H, Gu M, et al. Characterization of the extrinsic apoptotic pathway in the basal chordate amphioxus. Sci Signal, 2010, 3: ra66
89 Slavik K M, Morehouse B R, Ragucci A E, et al. cGAS-like receptors sense RNA and control 3′2′-cGAMP signalling in *Drosophila*. Nature, 2021, 597: 109–113
90 Holleufer A, Winther K G, Gad H H, et al. Two cGAS-like receptors induce antiviral immunity in *Drosophila*. Nature, 2021, 597: 114–118
91 Li Y, Slavik K M, Toyoda H C, et al. cGLRs are a diverse family of pattern recognition receptors in innate immunity. Cell, 2023, 186: 3261–3276.e20
92 Cai H, Li L, Slavik K M, et al. The virus-induced cyclic dinucleotide 2′3′-c-di-GMP mediates STING-dependent antiviral immunity in *Drosophila*. Immunity, 2023, 56: 1991–2005.e9
93 Boehm T, Morimoto R, Trancoso I, et al. Genetic conflicts and the origin of self/nonself-discrimination in the vertebrate immune system. Trends Immunol, 2023, 44: 372–383
94 Ghaffari S H, Lobb C J. Structure and genomic organization of a second cluster of immunoglobulin heavy chain gene segments in the channel catfish . J Immunol, 1999, 162: 1519–1529
95 Lee S S, Fitch D, Flajnik M F, et al. Rearrangement of immunoglobulin genes in shark germ cells. J Exp Med, 2000, 191: 1637–1648
96 Rumfelt L L, Avila D, Diaz M, et al. A shark antibody heavy chain encoded by a nonsomatically rearranged VDJ is preferentially expressed in early development and is convergent with mammalian IgG. Proc Natl Acad Sci USA, 2001, 98: 1775–1780
97 Ding S, Wu X, Li G, et al. Efficient transposition of the *piggyBac* (PB) transposon in mammalian cells and mice. Cell, 2005, 122: 473–483
98 Yusa K. *PiggyBac* transposon. Microbiol Spectr, 2015, 3: MDNA3-2014

https://engine.scichina.com/doi/10.1360/SSV-2024-0149

Origin and evolution of antigen receptor gene diversity

YUAN ShaoChun, HUANG ZiWen, CAI ZhenXi, HUANG ShengFeng & XU AnLong

Guangdong Province Key Laboratory of Pharmaceutical Functional Genes, MOE Engineering Research Center of South China Sea Marine Biotechnology, Center of Evolutionary and Synthetic Biology, School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China

The immune system can be divided into innate immunity and adaptive immunity. Since key organs and molecules of adaptive immunity, such as the thymus and the major histocompatibility complex, are present only in jawed vertebrates, they are considered an evolutionary innovation in vertebrates. Their origin has attracted substantial attention in the field of evolutionary biology and immunology. This review focuses on recent advancements made in the mechanism and evolution of antigen receptor gene diversity, specifically in the transposon origin of V(D)J rearrangement and host domestication mechanism of the *RAGL* transposons. This study also summarizes the research progress on *VLR* and other immune receptors with diversity in distinct evolutionary stages. This study delves into the evolutionary trajectory of antigen receptor gene diversity and provides new insights and technologies in the healthcare field.

antigen receptor gene, RAG, V(D)J recombination, evolution

doi: 10.1360/SSV-2024-0149

https://engine.scichina.com/doi/10.1360/SSV-2024-0149