

脊椎动物 NF- κ B1/2 基因亚家族分子进化和功能分歧分析

甘翔, 王继文*, 刘贺贺, 李亮, 孙文强, 王郁石, 张涛, 陈达, 黄惠兰

(四川农业大学, 畜禽遗传资源发掘与创新利用四川省重点实验室, 成都 611130)

摘要: 旨在分析 NF- κ B1/2 功能分化原因, 定义其各个结构域在其进化历程中的地位, 本研究对 NF- κ B1/2 基因进行序列比对, 进行了系统发育树的构建与选择压力预测等多种生物信息学技术分析。结果表明, 脊椎动物的 NF- κ B1/2 在进化历程中没有出现交织, 其中 RHD 与 IPT 表现出高度的保守性; ANK 在其直系与旁系同源结构中均表现出较大差异; DD 在直系同源结构中表现出高度的保守性, 在旁系同源结构中则出现了较大的差异。结果提示, 在脊椎动物中, NF- κ B1/2 是一对高度保守的基因, ANK 序列的变异是造成 NF- κ B1/2 在脊椎动物中功能分歧的重要原因, DD 则很可能对 NF- κ B1/2 基因的分歧有着重要的作用。

关键词: 脊椎动物; NF- κ B1/2; 序列比对; 系统发育树; 选择压力

中图分类号: Q953

文献标志码: A

文章编号: 0366-6964(2015)11-1987-07

Molecular Evolution and Functional Differences Analysis of NF- κ B1/2 Gene Subfamily of Vertebrate

GAN Xiang, WANG Ji-wen*, LIU He-he, LI Liang, SUN Wen-qiang, WANG Yu-shi, ZHANG Tao, CHEN Da, HUANG Hui-lan

(Farm Animal Genetic Resources Exploration and Innovation Key Laboratory of Sichuan Province, Sichuan Agricultural University, Chengdu 611130, China)

Abstract: The objective of this study was to analyze the reasons of different functions of NF- κ B1/2 and to identify the evolutionary status of its domains. The sequence information of NF- κ B1/2 was analyzed by following bioinformatic methods, including multiple sequence alignment phylogenetic tree construction and selective pressure prediction. The results showed that there was no intersection among vertebrate NF- κ B1/2 during evolution, and RHD and IPT were highly conserved; ANK had diverged among both orthologous and paralogous structures; DD showed a high conservation with orthologous structures but a large difference with paralogous structures. Our results suggested that NF- κ B1/2 were highly conservative among vertebrates, and the sequence variation of ANK might be an important factor causing the different functions of NF- κ B1/2 in vertebrate, and DD might be one of the reasons for the divergence of NF- κ B1/2 genes.

Key words: vertebrate; NF- κ B1/2; sequence alignment; phylogenetic tree; selection pressure

NF- κ B1 与 NF- κ B2 共同隶属于 NF- κ B 基因家族, 分别编码 p50 与 p52 的前体蛋白 p105 和 p100。

p105 与 p100 是核转录因子, 在二聚体化后能与免疫球蛋白轻链 (Immunoglobulin light chain,

收稿日期: 2014-12-29

基金项目: 国家水禽产业技术体系项目 (CARS-43-6)

作者简介: 甘翔 (1991-), 男, 广西桂林人, 硕士生, 主要从事动物遗传育种与繁殖研究, E-mail: ganxiangdk@163.com

* 通信作者: 王继文, 博士, 教授, 主要从事动物遗传育种与繁殖研究, E-mail: wjw2886166@163.com

MYLC) 基因增强子 cDNA 序列 (5'-GG-GACTTTC-3') 特异性结合并调控其转录, 从而启动和调节众多免疫和炎症反应^[1-2]。NF- κ B1/2 拥有 Rel homology domain (RHD)、IPT domain (IPT)、ankyrin repeats (ANK) 和 Death domain (DD) 4 个结构域^[3]。RHD 由大约 200 个氨基酸残基组成, 拥有核定位信号序列, 是与 DNA 结合的关键区域, 也是与抑制剂作用的重要区域; ANK 可作为 NF- κ B 家族的抑制剂, 通过与 RHD 结合来使其滞留在细胞质并保持静息态, ANK 蛋白的磷酸化、泛素化、降解释放 RHD 的过程也就是 NF- κ B 蛋白的激活过程^[4]; 而 IPT 和 DD 则与 NF- κ B 蛋白二聚体的稳定性与活性有着密切的联系^[5]。

研究表明, NF- κ B 基因的原始结构起源于后生动物以前, 在进化过程中分化为 NF- κ B1 和 NF- κ B2^[6]。利用模式脊椎动物的 NF- κ B1/2 氨基酸序列构建进化树, 发现它们分成两支, 且完全没有交叉的迹象, 说明两者在进化上出现了明显的分歧, 暗示它们在结构上存在一定差异。另外, 在脊椎动物中, NF- κ B1/2 蛋白的功能也表现出了一定的差异。有文献报道^[7-8], 不同的 NF- κ B1/2 蛋白二聚体对于相应的 κ B 位点会显示出不同的亲和力与激活能力。不仅如此, 两者的 ANK 抑制结构对于不同的二聚体也可能显示出不同的抑制强度。还有报告^[7-9]称, 许多刺激信号如 TNF- α 、IL-1、LPS 等对 NF- κ B1 和 NF- κ B2 氨基酸序列的刺激能力与刺激途径也有所差异。在小鼠中, 敲除 NF- κ B1 会导致 B 细胞的多种功能缺陷, 而敲除 NF- κ B2 则会引起淋巴结构改变, 影响上皮发育。另外, 两者的敲除反应都会一定程度上导致胃黏膜增生, 但 NF- κ B2 对胃黏膜的影响程度要大于 NF- κ B1^[4]。当前的研究已经基本定义了 NF- κ B1/2 基因的起源, 也明确了 NF- κ B 家族在脊椎动物中的功能差异, 但对造成这些差异的原因还没有明确的定义。不过可以推断, NF- κ B1/2 的 RHD、IPT、ANK 与 DD 在 NF- κ B1 与 NF- κ B2 基因亚家族分子进化特征的形成和分子功能分歧的产生方面扮演了重要的角色。

本研究选择模式脊椎动物的 NF- κ B1/2 基因与蛋白质序列, 针对 NF- κ B1/2 的 RHD、IPT、ANK 和 DD 4 个结构域进行生物信息学分析, 旨在揭示引起 NF- κ B1/2 功能分化的部分原因, 并阐述 NF- κ B1/2 在脊椎动物中进化、变迁的基本趋势。

1 材料与方法

1.1 序列数据的收集与下载

序列比对与核苷酸替代率分析所用的 NF- κ B1/2 氨基酸、核苷酸序列均下载自 NCBI 数据库 (<http://www.ncbi.nlm.nih.gov/>)。序列比对使用了包括人类 (*Homo sapiens*)、大鼠 (*Rattus norvegicus*)、小鼠 (*Mus musculus*)、原鸡 (*Gallus gallus*)、非洲爪蟾 (*Xenopus (Silurana) tropicalis*)、美国短吻鳄 (*Alligator mississippiensis*)、二色热带鱼 (*Stegastes partitus*) 7 种具有代表性的模式动物的 NF- κ B1/2 的氨基酸序列。而核苷酸替代率分析则选用了鱼类、两栖类、爬行类、鸟类和海陆哺乳类共计 20 余种脊椎动物的 NF- κ B1/2 核苷酸序列^[10]。

1.2 序列比对与系统发育树的构建

使用 ClustalW 将下载序列对齐, 使用 DAMBE 等软件进行格式处理, 得到的结果输入 MEGA 6.0 软件, 使用邻接法 (Neighbor-Joining, NJ), 经过自展法 (Bootstrap) 1 000 次重复验证, 以保证树拓扑结构的可靠性, 分别构建 NF- κ B1/2 氨基酸全序列及其各个结构域的氨基酸序列系统发育树^[11-12]。

1.3 NF- κ B1/2 蛋白的结构域预测

使用 Pfam 在线数据库 (<http://pfam.sanger.ac.uk/>) 对 NF- κ B1/2 蛋白进行结构预测。

1.4 进化速率计算检验与分析

使用 DAMBE 等软件处理序列文本格式, 使用 MEGA6.0、ClustalW 进行向导树的构建^[13]。将处理结果文件输入 paml 软件包, codeml 工具^[14-15], 分别选取模型 A: 所有的分支有一个 ω , 模型 B: 所有的分支有相同的 ω 且等于 1, 模型 H: 每一支都有自己的 ω , 选择最优模型从而得出各个结构域的整体 ω 值。每个位点 ω 的计算也使用 codeml 工具, NS-site 参数为 0、1、2、7、8, 在生成结果中优先选用 BEB 模型, 从而得到 NF- κ B1/2 的 RHD、ANK 和 DD 中各个位点的 ω 值^[16-17]。

2 结果

2.1 基于 NF- κ B1/2 与结构域 RHD、IPT、ANK 和 DD 氨基酸序列的系统发育树

如图 1a 所示, 进化树中 NF- κ B1/2 分为两大支, 进化趋势与物种进化趋势一致。由鱼类至两栖类之后, 系统发育树再分为两支, 其中一支为爬行类与鸟类, 另一支为哺乳类。树中 NF- κ B1/2 两个成

员在整个脊椎动物进化历程中并没有出现任何的交互,这表明 NF- κ B 基因家族早在脊椎动物与非脊椎动物分歧时,就已经分化成为 NF- κ B1 和 NF- κ B2, 然后开始独立进化。

脊椎动物 NF- κ B1/2 两个成员都包含 RHD、ANK、IPT 和 DD 4 个结构域(图 1b),由它们的氨基酸序列所构建的系统发育树(图 2)与基于全序列构建的发育树基本一致。

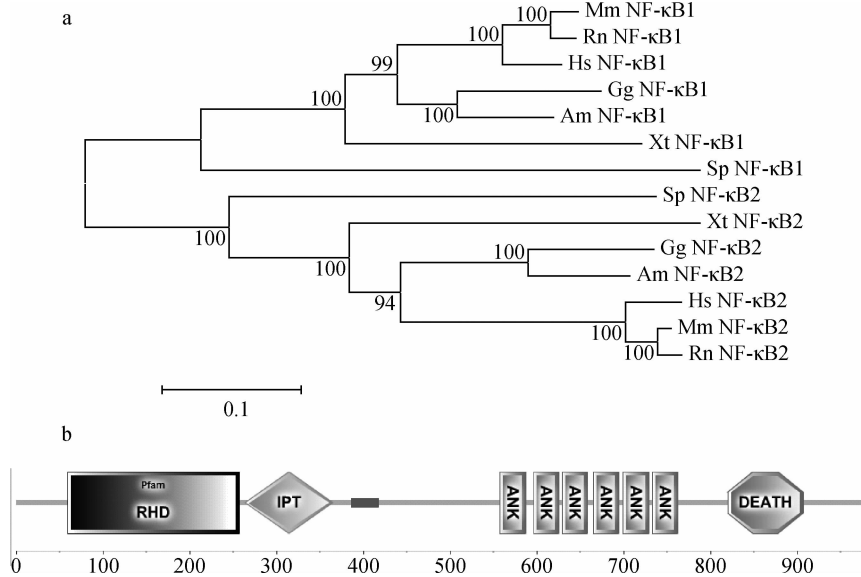


图 1a 为 NF- κ B1/2 系统发育树, 分支上的数字为 bootstrap 1 000 次的支持率, 代表分歧度; 图 1b 为 NF- κ B1/2 的保守结构域

Figure 1a represents the molecular phylogeny tree of NF- κ B1/2, the numerals on branches indicate the difference degrees; Figure 1b represents the conserved domains of NF- κ B1/2

图 1 NF- κ B1/2 氨基酸序列的系统发育树与保守结构域

Fig. 1 Phylogenetic tree of NF- κ B1/2 and their conserved domains

2.2 基于 NF- κ B1/2 结构域 (RND、IPT、ANK 和 DD) 的序列分析

图 3 中, NF- κ B1/2 之间在 RHD 上存在较多保守区域, 但在序列的 60~66、108~117、157~164 位氨基酸区段(图中方框标出)也发现了它们各自的特征序列。在 IPT 区域上, 两个基因在相应物种间都表现出较高的同源性。而在 ANK 区域上, 无论在直系同源还是旁系同源结构中, 该区域都出现了较大差异。另外, 相对于其他 3 个结构域, DD 区域的旁系同源结构之间存在很大的差异, 而直系同源结构则表现出较高的同源性。

2.3 NF- κ B1/2 各保守结构域选择压力比较

Codeml 中各模型比对结果显示, NF- κ B1 的 ANK、RHD 和 DD, NF- κ B2 的 ANK 与 RHD 都更适用于 H 模型, 检测结果为极显著; 而 NF- κ B1/2 的 IPT, NF- κ B2 的 DD 更适用于 A 模型, 检测结果也极显著。

分别比较了 NF- κ B1/2 的 RHD、ANK、IPT 与 DD 区域的选择压力, 结果(表 1)表明, 在鱼类及两栖动物上, NF- κ B1/2 的 RHD, NF- κ B1 的 ANK 显

示出较高的 ω 值; 在爬行动物与鸟类上, NF- κ B1/2 的 ANK 与 NF- κ B1 的 DD ω 值较大; 而在哺乳动物中, 大多数结构域的 ω 值均低于 0.1。

而对 NF- κ B1/2 氨基酸各位点选择压力(图 4)分析表明, RHD、IPT、ANK 与 DD 序列的 ω 值均小于 1.2, 且在所分析脊椎动物中, 4 个结构域均没有出现处于显著水平的正选择位点。NF- κ B1/2 的 RHD 区域大多数位点 ω 值在 0.1 以下。相对于 RHD、IPT 与 DD 区域, ANK 区域则拥有较高的进化速率; NF- κ B1 的 IPT 绝大多数位点 ω 值较低, 而 NF- κ B2 的 IPT 所有位点的 ω 值均在 0.01 以下; NF- κ B1/2 的 DD ω 曲线存在较大差异, NF- κ B1 的 DD 各个位点的 ω 值显示出较大的波动。

3 讨论

NF- κ B1/2 属于 NF- κ B 基因家族中的亚家族, 作为转录因子, 能够诱导机体多种正向调控基因的转录, 在动物机体免疫反应、骨骼肌发育中发挥重要作用^[1]。研究表明, NF- κ B 基因的原始结构起源于后生动物以前, 在进化过程中分化为 NF- κ B1 和

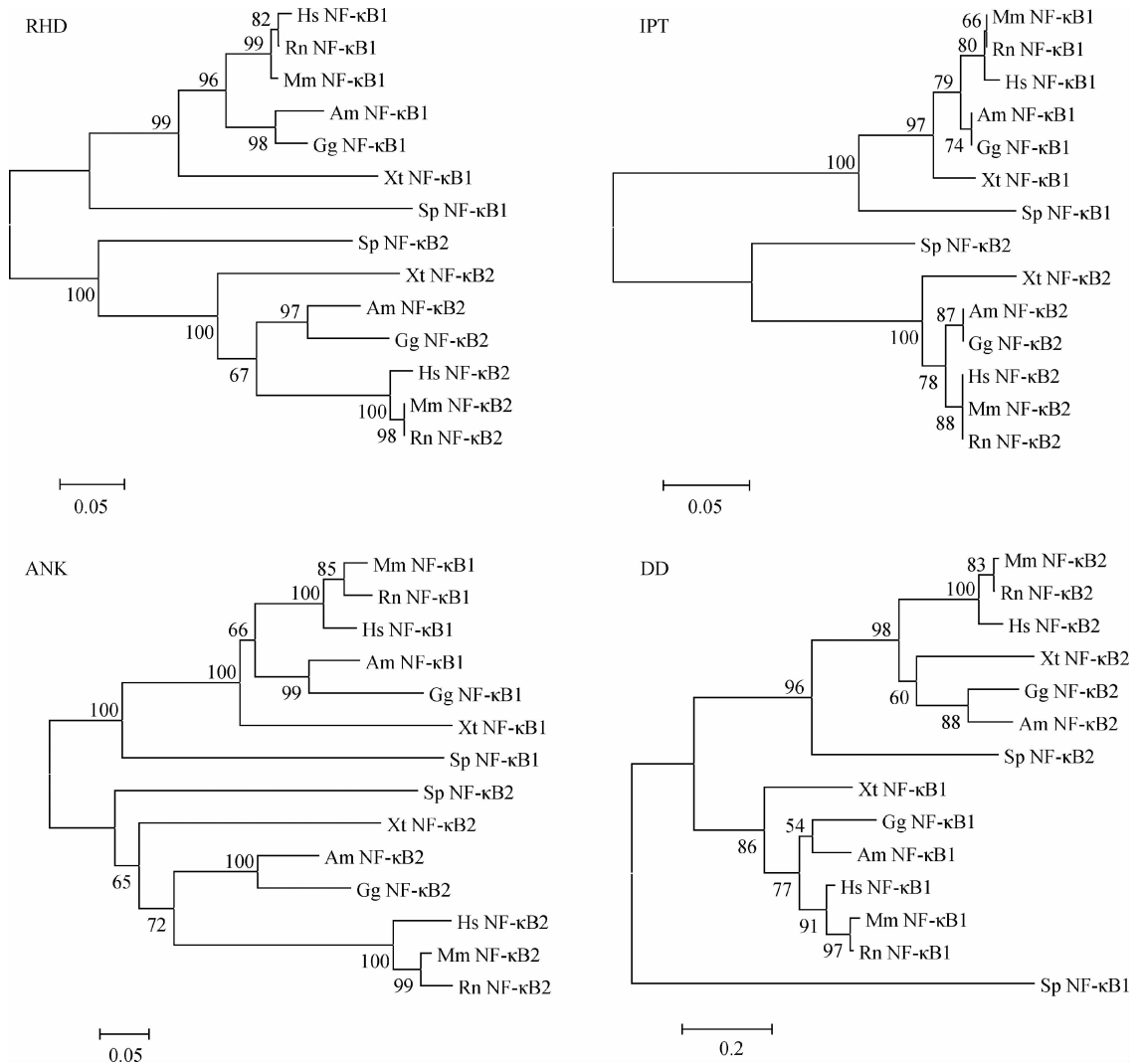


图 2 NF-κB1/2 4 个结构域 RHD、IPT、ANK 与 DD 氨基酸序列的系统发育树

Fig. 2 The phylogenetic trees of RHD, IPT, ANK and DD of NF-κB1/2 based on their amino acid sequences

表 1 NF-κB1/2 基因各结构域在脊椎动物各个群类中的 ω 均值

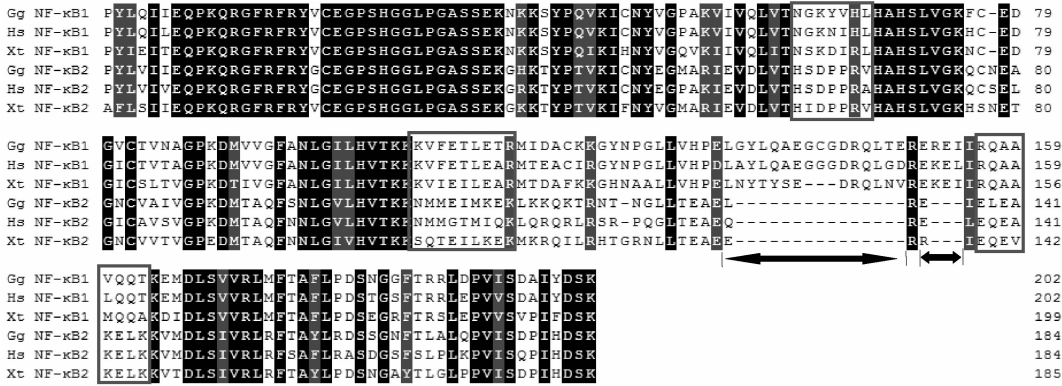
Table 1 The average value of ω of domains of NF-κB1/2 genes in different vertebrate groups

物种 Species	ANK	RHD	IPT	DD
NF-κB1 基因 NF-κB1 gene				
鱼类及两栖动物 Fish and amphibians	0.210 03	0.122 90	0.014 10	0.163 40
爬行动物及鸟类 Reptiles and birds	0.285 90	0.049 58	0.014 10	0.334 31
哺乳动物 Mammals	0.071 65	0.022 21	0.014 10	0.105 00
NF-κB2 基因 NF-κB2 gene				
鱼类及两栖动物 Fish and amphibians	0.095 75	0.151 32	0.007 70	0.068 99
爬行动物及鸟类 Reptiles and birds	0.165 95	0.011 40	0.007 70	0.068 99
哺乳动物 Mammals	0.055 79	0.004 23	0.007 70	0.068 99

NF-κB2^[6]。通过构建 NF-κB1/2 的进化树,发现 NF-κB 基因家族早在脊椎动物与非脊椎动物发生分歧前,就已经分化成为 NF-κB1 和 NF-κB2,然后开始独立进化,这与其他报道^[18]得出的结论一致。

RHD、IPT、ANK 与 DD 是 NF-κB1/2 亚家族成员重要的典型结构域,早在两侧对称动物时期 4 个结构域就已经形成^[8],且这样的序列结构一直持续到了哺乳动物^[8,18]。可以推断,RHD、IPT、ANK

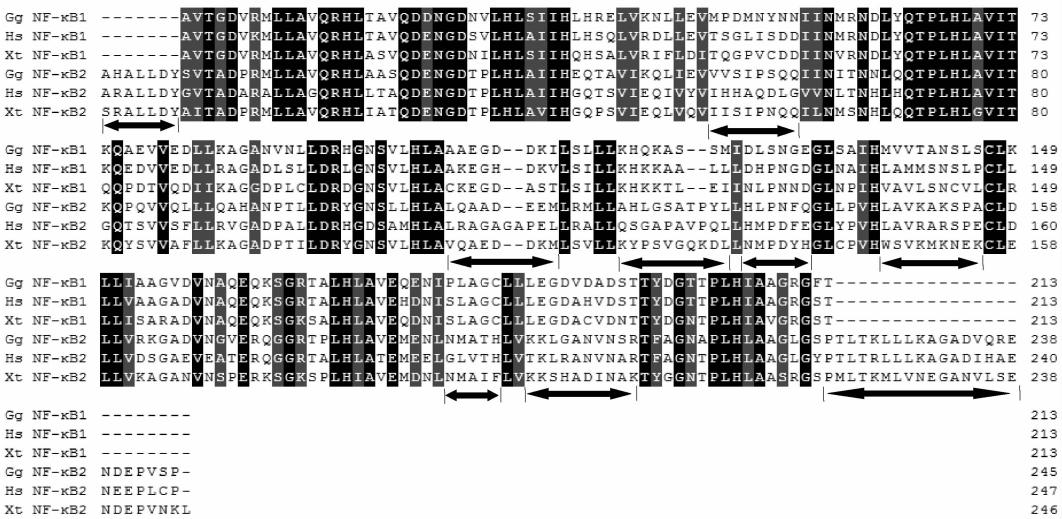
RHD



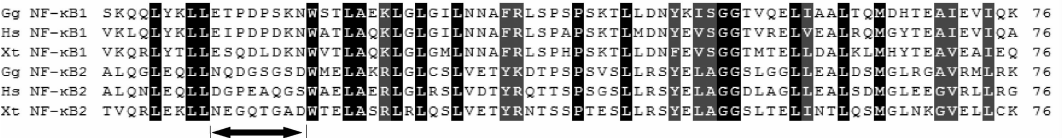
IPT



ANK



DD



NF-κB1/2 氨基酸序列间存在差异的区段用黑色箭头注明,完全相同的区段用黑色表示,同源性极高的区段以灰色表示,而 NF-κB1/2 各自的特征序列区段则用方框标出

The black arrow represents the different amino acids of NF-κB1/2, these identical sequences are marked as black and the highly homologous sequences are marked as grey. The special signature sequences of NF-κB1/2 are shown in the box

图 3 NF-κB1/2 的 RHD、IPT、ANK 和 DD 保守结构域区域对比

Fig. 3 Alignment of RHD, IPT, ANK and DD conserved domains of NF-κB1/2

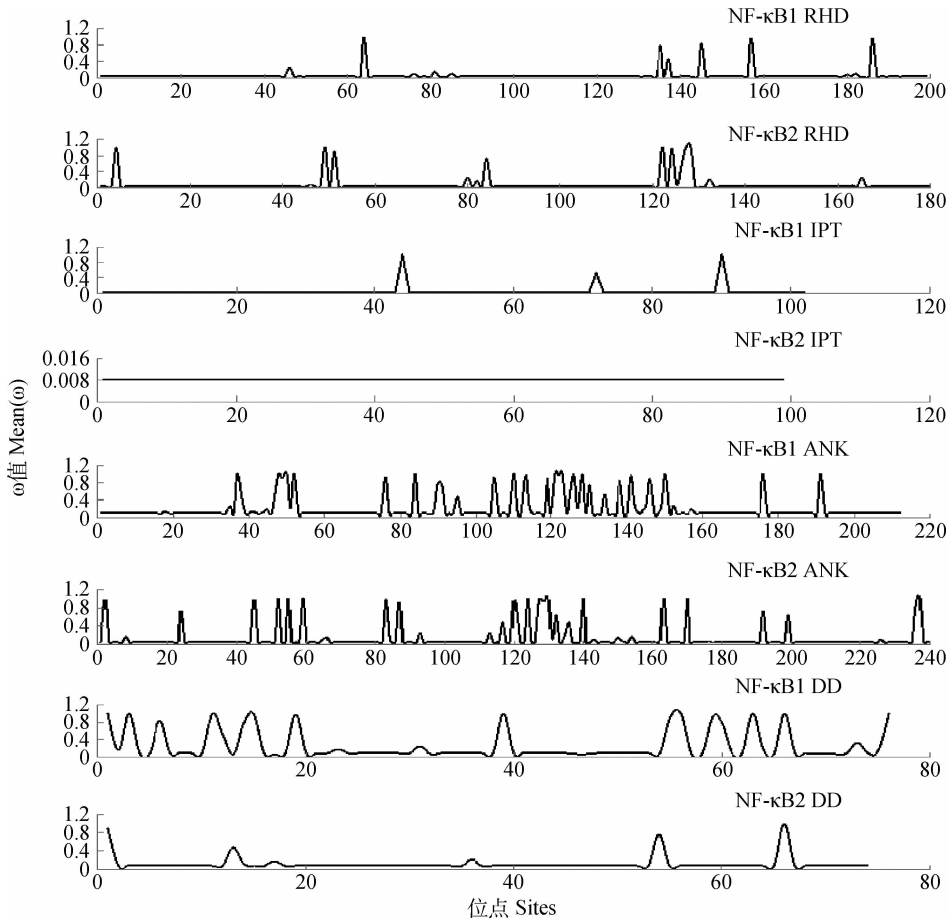


图 4 NF- κ B1/2 RHD、IPT、ANK 与 DD 结构域各位点选择压力

Fig. 4 The selection pressure of RHD, IPT, ANK and DD conserved domains of NF- κ B1/2

与 DD 在 NF- κ B1 与 NF- κ B2 基因亚家族分子进化特征的形成和分子功能分歧的产生方面扮演了重要的角色。本研究构建了各个结构域氨基酸系统发育树,发现由它们的氨基酸序列构建的系统发育树与基于全序列构建的发育树形态基本相同,说明 NF- κ B1 和 NF- κ B2 分歧后, NF- κ B1 和 NF- κ B2 各个结构域的进化趋势与其整体序列基本一致。

选择压力分析能使人们对 NF- κ B1/2 在进化中所受到的环境阻力与其序列位点的变异情况有更深入的认识。因此,为了更进一步探索 NF- κ B1/2 在脊椎动物阶段的进化历程,笔者又分别分析了 NF- κ B1/2 各个结构域在不同动物类群中(鱼类、鸟类、哺乳动物)的选择压力,发现不同种类动物 NF- κ B1/2 保守结构域 ω 值存在差异,说明 NF- κ B1/2 各个结构域在不同的进化阶段存在着明显不同的进化历程。而在对 NF- κ B1/2 各个结构域序列位点的分析中,发现 NF- κ B1/2 整体序列相对保守,不存在显著的正选择位点。不过 NF- κ B2 的 DD 所承受的

选择压力却明显大于 NF- κ B1 的 DD,这很可能是由于 NF- κ B2 的 DD 所含有的水解抑制区(Processing inhibitory domain, PID)在生理过程中的重要作用所导致^[19]。

另外,对脊椎动物 NF- κ B1/2 的 RHD、IPT、ANK 和 DD 与直系、旁系同源结构的同源性分析显示, NF- κ B1/2 的序列发生了各自独立的进化,且这样的独立进化在 RHD 中表现的尤为显著,暗示 NF- κ B1/2 在功能上的分工越来越明确。而 IPT 表现的高保守性与它在生理过程中的重要功能相吻合^[5,20]。ANK 相对于其他 3 个结构域在脊椎动物的 NF- κ B1/2 间拥有显著的差异,推测它可能是导致 NF- κ B1/2 在脊椎动物阶段序列、功能差异形成的重要因素。而 DD 仅在 NF- κ B1/2 之间存在较大的差异,且这样的差异在鱼类中就已经表现出来了,这不仅说明了它在脊椎动物阶段中的保守性,也暗示了 NF- κ B1/2 的 DD 分歧的年代较为久远,甚至可能与 NF- κ B1/2 基因的分歧有一定关联,但是因

为 NF- κ B1/2 的分化早于脊椎动物而晚于两侧对称动物,期间相关的动物序列素材较少,因此人们无法进行确切的论证。

综上,在脊椎动物的各个类群中,NF- κ B1/2 承受着不同的选择压力,它们的 4 个结构域也表现出了不同的变异程度与特点。NF- κ B1/2 在脊椎动物中的进化与功能分歧主要由 ANK 的序列变异造成,而 NF- κ B1/2 在脊椎动物之前的分歧很可能与 DD 有着密切的联系。

参考文献 (References):

- [1] VALLABHAPURAPU S, KARIN M. Regulation and function of NF- κ B transcription factors in the immune system[J]. *Annu Rev Immunol*, 2009, 27(1): 693-733.
- [2] IWANASZKO M, BRASIER A R, KIMMEL M. The dependence of expression of NF- κ B-dependent genes: statistics and evolutionary conservation of control sequences in the promoter and in the 3'UTR[J]. *BMC Genomics*, 2012, 13: 182.
- [3] GUTIERREZ H, DAVIES A M. Regulation of neural process growth, elaboration and structural plasticity by NF- κ B[J]. *Cell*, 2011, 34(6): 316-325.
- [4] BAEUERLE P A, BALTIMORE D. NF- κ B: ten years after[J]. *Cell*, 1996, 87(1): 13-20.
- [5] BAKKAR N, GUTTRIDGE D C. NF- κ B signaling: a tale of two pathways in skeletal myogenesis [J]. *Physiol Rev*, 2010, 90(2): 495-511.
- [6] GAUTHIER M, DEGNAN B M. The transcription factor NF- κ B in the demosponge *Amphimedon queenslandica*: insights on the evolutionary origin of the Rel homology domain[J]. *Dev Genes Evol*, 2008, 218(1): 23-32.
- [7] SIMEONIDIS S, STAUBER D, CHEN G, et al. Mechanisms by which I κ B proteins control NF- κ B activity[J]. *Proc Natl Acad Sci U S A*, 1999, 96(1): 49-54.
- [8] SAVINOVA O V, HOFFMANN A, GHOSH G. The Nfkb1 and Nfkb2 proteins p105 and p100 function as the core of high-molecular-weight heterogeneous complexes[J]. *Mol Cell*, 2009, 34(5): 591-602.
- [9] WHITESIDE S T, ISRAEL A. I κ B proteins: structure, function and regulation[J]. *Semin Cancer Biol*, 1997, 8(2): 75-82.
- [10] HUANG Y H, TEMPERLEY N D, REN L M, et al. Molecular evolution of the vertebrate TLR1 gene family—a complex history of gene duplication, gene conversion, positive selection and co-evolution [J]. *BMC Evol Biol*, 2011, 11(1): 149.
- [11] LU M X, HUA J, CUI Y D, et al. Five small heat shock protein genes from *Chilo suppressalis*: characteristics of gene, genomic organization, structural analysis, and transcription profiles [J]. *Cell Stress Chaperones*, 2014, 19(1): 91-104.
- [12] QIAN Y H, WU H Y, JI X Y, et al. Mitochondrial genome of the stonefly *kamimuria wangi* (Plecoptera: Perlidae) and phylogenetic position of plecoptera based on mitogenomes [J]. *PLoS ONE*, 2014, 9(1): e86328.
- [13] 徐怀亮. 7 个灵长类物种 PRSS12 基因编码区序列测定与进化分析 [J]. 东北林业大学学报, 2010, 38(4): 85-88.
- XU H L. Sequencing and phylogenetic analysis of PRSS12 coding region of seven primate [J]. *Journal of Northeast Forestry University*, 2010, 38(4): 85-88. (in Chinese)
- [14] PICONE B, CHRISTOFFELS A. Molecular evolution of key receptor genes in primates and non-human primates [J]. *J Genet Genomics*, 2014, 2(4): 57-61.
- [15] ARENAS M, POSADA D. Simulation of genome-wide evolution under heterogeneous substitution models and complex multispecies coalescent histories [J]. *Mol Biol Evol*, 2014, 31(5): 1295-1301.
- [16] ZHAO H, ZHOU Y, PINTO C M, et al. Evolution of the sweet taste receptor gene *Tas1r2* in bats [J]. *Mol Biol Evol*, 2010, 27(11): 2642-2650.
- [17] YU L, JIN W, ZHANG X, et al. Evidence for positive selection on the leptin gene in Cetacea and Pinnipedia [J]. *PLoS ONE*, 2011, 6(10): e26579.
- [18] HUGUET C, CREPIEUX P, LAUDET V. Rel/NF- κ B transcription factors and I κ B inhibitors: evolution from a unique common ancestor [J]. *Oncogene*, 1997, 15(24): 2965-2974.
- [19] 李 贤, 刑 达, 陈小佳. 核因子 KB 抑制蛋白家族的研究进展 [J]. 中国病理生理杂志, 2006, 22(12): 2480-2484.
- LI X, XING D, CHEN X J. Advances of nuclear factor KB of inhibit protein family [J]. *Chinese Journal of Pathophysiology*, 2006, 22(12): 2480-2484. (in Chinese)
- [20] CRAMER P, MÜLLER C W. A firm hand on NF κ B: structures of the I κ B α -NF κ B complex [J]. *Structure*, 1999, 7(1): R1-R6.