植物学通报 *Chinese Bulletin of Botany* 2008, **25** (4): 401–406, www.chinbullbotany.com

·综述·

植物进化中的正选择作用

高远[1, 2], 田李[1, 2], 秦松[1*]

[1] 中国科学院海洋研究所, 青岛 266071; [2] 中国科学院研究生院, 北京 100049

摘要 正选择是指将因含有有利突变而提高个体适合度的等位基因固定下来的选择作用, 研究正选择对理解生物进化过程具有重要意义。本文回顾了近年来在植物基因中发现的正选择作用, 分别对陆生植物和藻类中经历正选择作用的基因进行了总结, 其中在陆生植物中发现的正选择位点主要集中在与生殖相关及与抗逆相关的基因上, 这为以后对植物中正选择作用的研究提供了线索。

关键词 适应性进化, 植物, 正选择, 生殖相关基因, 抗逆相关基因

高远, 田李, 秦松 (2008). **植物进化中的正选择作用**. **植物学通报 25**, 401–406.

达尔文在《物种起源》中指出自然选择是促进物种进化最主要的动力, 关于此观点的争论持续了一个多世纪。中性理论(Neutral Theory)指出: 在进化过程中, 绝大多数核苷酸替代都是中性或近中性的突变随机固定, 而不是正选择的结果(Kimura,1968)。King 和 Jukes (1969)对中性理论进行了证实和补充, 他们从另一个角度指出: 许多蛋白质多态性在选择上表现为中性或近中性, 并在突变引入与随机消亡之间保持平衡。但随后许多生物学家对此提出了异议。Nei(1987)提出: 尽管大多数的氨基酸替代都是通过随机固定产生的, 但的确有部分替代是适应性的。这种适应性进化理论可以看作是对中性理论的进一步修正。对鳄鱼血红蛋白(Perutz et al., 1981)等的多项研究均发现了正选择位点的存在, 从而支持了适应性进化(adaptive evolution)的理论。

所谓正选择(positive selection), 又称定向选择(directional selection)或达尔文选择(Darwinian selection), 是指将含有发生了有利突变并因此提高个体适合度的等位基因固定下来的选择作用(Shimizu and Purugganan, 2005)。对于一个物种而言, 曾经历过的正选择是进化历史中的重大事件, 因为这意味着该物种对环境中与其生存密切相关的某一因素的极大适应, 也暗示着在该进化时间点某些蛋白新功能的产生。对正选择作用的重要位点的研究既有助于对基因功能区的预测, 又为解决某些生物技术和生物医学难题提供了崭新的思路。目前, 正选择作用的检测也开始为越来越多的生物学家所重视, 并成为研究各种分子生物学问题的重要方法。

目前, 在陆生植物和藻类进化中都发现了正选择作用, 其中对陆生植物的研究中发现的正选择作用主要集中在与生殖及与抗逆相关的基因。

1 陆生植物

1.1 与生殖相关基因

与生殖相关的基因直接决定着生物的各种生殖相关性状, 也决定着生物在繁殖方式和各环节上对所处环境的适应性。在动物中, 已经在许多生殖蛋白中发现适应性进化, 原因可能是生殖蛋白中氨基酸替代更易于被固定(Lee et al., 1995; Metz and Palumbi, 1996; Bielawski and Yang, 2001; Swanson et al., 2001, 2003)。目前, 植物中关于生殖蛋白经历正选择的事例也有报道。

其中已发现的正选择作用的位点之一存在于拟南芥

收稿日期: 2007-07-24; **接受日期**: 2008-02-15
基金项目: 国家自然科学基金(No. 30670165)和中国科学院知识创新重要方向项目(No. KSCX2-YW-G-002)
* **通讯作者**。E-mail: sqin@ms.qdio.ac.cn

https://www.sciengine.com/doi/10.3724/CBB-t27-274

编码 oleopollenin 的基因上。在拟南芥的花粉表面，oleopollenin 是花粉壁的主要组成部分之一。Schein 等(2004)通过对拟南芥和其相近物种中 7 个花粉特异的 oleopollenin 的基因序列的系统进化分析和正选择检测发现，来源于拟南芥的oleosin-like结构域处于快速进化中，而且处于正选择作用之下，受到的正选择作用可能有助于作为配体的 oleosin-like 结构域与雌蕊受体之间的物种特异性识别。而这种特异性识别被认为是导致生殖隔离和物种形成的最重要机制之一。

开花植物中开花时间也是与生殖有关的重要性状，它直接限制着纬度等植物生存的地域条件。*FRIGIDA* (*FRI*)基因是拟南芥中调节开花时间的重要基因。它与开花位点 C(*FLC*)协同作用，使植物开花时间延迟。*FLC* 是参与开花起始的关键基因， 受到 *FRI* 基因的调节(Sheldon et al.,1999)。Corre 等(2002)对来自西欧的 25 个野生生态型拟南芥的*FRI*序列研究发现，在种内有过量的非同义替代位点，指出在西欧拟南芥中 *FRI* 基因多态性是由一系列隔离群体的正选择作用造成的。植物 MADS-box 基因编码在发育多方面起重要作用的转录调节因子，该基因经过复制形成了很大的基因家族。Martínez-Castilla 和 Martínez-Castiua(2003)对拟南芥基因组中 104 个 MADS 序列研究发现，不同的选择机制影响了系统树中不同分支的进化，并发现影响开花时间基因的分支受到了正选择作用。

异丁子香酚 -O- 转甲基酶((iso)eugenol-o-methyltran-sferase, IEMT)是*Clarkia breweri*中参与产生花的挥发性化合物甲基丁子香酚和甲基异丁子香酚的酶。IEMT 可能是由咖啡酸 -O- 甲基转移酶经基因复制和氨基酸歧化进化而来，并由此获得了其新功能(Wang et al.,1997)。Barkman(2003)通过最大似然法获得了正选择导致 IEMT 自产生以来适应性进化的统计学证据。虽然具体压力尚不清楚，但推测该基因的适应性进化可能增强了植物对花粉传播者的吸引和对食草动物的抵抗效果。

此外，*LEGCYC* 基因是控制花对称性的调控基因。在羽扇豆属(*Lupinus*)中，*Lupinus densiflorus* 有着更小的近轴花瓣和更大的侧面花瓣，与该属的其它种在比例上有所不同。Ree 等(2004)在对该物种的 *LEGCYC* 基因研究时指出，系统树上 *L. densiflorus* 分支上一个旁系同源*LEGCYC1B*的部分密码子位点经测定显示经历过正选择，并认为这种正选择作用与 *L. densiflorus* 在进化中的形态学变化有关。

1.2 与抗逆相关基因

当植物体遇到对自己生存极端不利的环境或受到病虫害侵袭时，这类基因能够保护植物体免受伤害，它们直接决定着植物体的抵抗能力以及是否能够生存并完成生活周期，对植物体有十分重要的作用。

病原抗性基因广泛存在于动物和植物中，主要参与宿主对真菌和细菌等病原微生物的识别和抑制作用，在这些基因的许多位点上经常发现正选择作用。植物中的一类病原抗性基因又常称为植物抗性基因(plant R-genes)。该基因往往在染色体上以多拷贝形式串联排布，并具有多态性。R 基因被认为参与一种宿主和病原在进化上的军备竞赛(evolutionary arms race)。一种典型的军备竞赛是: R 基因的一个有利突变使宿主能够识别以前所不能识别的抗原决定簇(称为Avr因子)，因此被选择压力所固定并获得了高基因频率(Clay and Kover, 1996)。典型的植物 R 基因一般都含有一个亮氨酸重复(leucine-rich repeat, LRR)区。它被认为具有病原识别功能，因此也决定着抗性的特异性。通过对拟南芥基因组中 49 个单 R 基因和 32 组含 2-12 个 R 基因进行聚类分析，发现 R 基因受到显著的正选择作用(Bergelson et al., 2001)。该检测结果符合上述军备竞赛理论。

植物的内 -β-1,3 葡聚糖酶(EGases) 也参与植物体的防御功能，它能够降解攻击性病原体的细胞壁多糖。大豆中该酶的同功酶 EGaseA 和 EGaseB 的区别在于它们对一种大豆病原体 *Phytophthora sojae*(大豆疫霉菌)产生的葡聚糖酶抑制蛋白(GIP1)的敏感性有差异。主要的同功酶 EGaseA 是 GIP1 的高亲和配基，其活性能被GIP1完全抑制，而EGaseB却不受GIP1影响(Ham et al.,1997)。Bishop 等(2004)对来自 20 个 *Glycine soja*(野大豆)、4 个 *Glycine* 属其它品种以及一些双子叶植物的局部序列进行了检测， 发现在 *Glycine* 属

https://www.sciengine.com/doi/10.3724/CBB-t27-274

EGaseA 上有很强的正选择迹象, 而在双子叶植物 EGases 和 Glycine EGaseB 中有较弱的正选择迹象。对此的解释是: 在宿主与病原进化上的军备竞赛中, EGaseA由于在引发进一步防御反应中的关键作用而受到了比 EGaseB 更强的来自病原对抗手段的选择作用。

拟南芥等植物的次级代谢混合物中的主要组分是硫代葡萄糖苷(glucosinolates), 该物质对植物抵抗食草昆虫有一定作用。*MaM* 基因聚簇编码的 MAM 控制着硫代葡萄糖苷合成反应的前一步, 是硫代葡萄糖苷代谢的关键酶。Benderoth 等(2006)对来自拟南芥和其近缘植物的包含 *MaM* 聚类的 BAC(bacterial artificial chromosomes)进行了测序和分析, 发现MAM聚类中的 MAM1 的进化受到正选择的作用, 因此具有更大的底物多样性; 并认为该基因的进化与植物 R 基因军备竞赛式的进化模式有很大的相似之处。

此外, 一年生草本植物 *Alopecurus myosuroides* (black grass)的乙酰辅酶A羧化酶(ACCase)是高效除草剂的作用靶位。虽然乙酰辅酶 A 羧化酶是控制脂肪酸合成的关键酶, 直接影响着植物体的生存, 序列分析也证明其相当保守, 但Christophe等(2004)通过对18个大穗看麦娘群体86个植株的研究, 发现其群体间的分化非常明显; 虽然目前以乙酰辅酶 A 羧化酶为靶位的除草剂的使用只有15代, 但乙酰辅酶A羧化酶的除草剂作用位点已经表现出正选择作用的迹象。

2 藻类

对于生活在水中的藻类来说, 其生存环境与陆生植物有很大不同, 进化的模式和环境因子之间的关系也不一样。这表现在光对其生存和发育的影响十分重大。而水中入射光的强度和光质又受到水深的直接影响, 因此作为光合植物, 藻类的进化就明显受到了这些因素的影响。

藻胆蛋白是一类广泛存在于蓝藻、红藻和隐藻中的重要的捕光色素蛋白,主要包括藻红蛋白(phycoerythrin, PE)、藻蓝蛋白(phycocyanin, PC)和别藻蓝蛋白(allophycocyanin, APC)3 种。所捕获的光能沿前两者组成的杆, 传递到 APC 复合体形成的轴, 最终传递给光反应中心, 实现光的捕获(Christophe et al., 2004)。Qin 等(2005)对来源于红藻和蓝藻的 11 条 PE 序列的正选择分析显示, 在 3 个分支上分别有 40.6%、13.3% 和 12.8% 的位点处于正选择作用下; 并认为低光适应型原绿球藻和海洋聚球藻对所处的大洋环境分别采取了 2种适应机制, 因而具有不同的进化模式。Zhao 和 Qin (2006a)对蓝藻基因组中参与藻胆体组装的各类基因分析指出, 在不同分支上的色基结合区和 N 端螺旋发卡区上的部分位点存在着正选择; 并发现每个区域内部的正选择位点都存在相关性, 提示它们受到了相同的正选择压力。此后, 通过对来自高光适应型和低光适应型的原绿球藻群体的藻红蛋白序列比较分析发现, 在 2 种群体中正选择位点在蛋白三级结构上的分布不同, 这意味着 2 种群体在不同环境下具有完全不同的选择模式(Zhao and Qin, 2006b)。

除藻胆蛋白外, 类胡萝卜素也是一类广泛分布在光合植物中的捕光色素, 其在光收集和对植物的光保护中起着重要作用。在某些藻类中, 八氢番茄红素去饱和酶(phytoene desaturase, PDS)是参与类胡萝卜素合成反应的重要酶类(Rabbani et al.,1998)。Liang 等(2006)通过研究高等植物、绿藻和蓝藻的八氢番茄红素去饱和酶基因序列发现, 在绿藻分支上, 7.9% 的位点处于正选择压力下, 绿藻的*PDS*基因显示出与其它种类不同的进化模式。

除此之外, 作为蓝藻细胞中的一种防御系统, 限制性修饰(restriction modification, RM)系统能够降解 DNA 以防止外源 DNA 的入侵。RM 系统主要由具有 DNA 降解活性的限制性内切酶(RMase)和能够特异性甲基化宿主DNA以保护其不受降解的相应转甲基酶(MTase)构成(Bujnicki, 2001)。Zhao 等(2005)通过对 MTase 基因家族的分析发现10.2%的位点受到正选择作用, 并发现这些位点参与了 MTase 的 DNA- 蛋白质相互作用。

温度也是影响植物生存的重要非生物环境条件之一。某一地域的温度变化深刻地影响着植物体内蛋白的结构和功能活性, 因此可能是影响整个群落生物的生理和形态基础。Miller(2003)对因热耐受性不同而产生

https://www.sciengine.com/doi/10.3724/CBB-t27-274

分化的热泉蓝藻聚球藻的碳固定基因 *rbcL* 进行正选择分析发现，与多个蛋白活性位点有接触的第389位氨基酸处于正选择作用下。该位点的丙氨酸在高温耐受性的种类中被具有更高疏水性的异亮氨酸所替代，这种替代能明显提高该蛋白在高温下构象的稳定性。

通过以上对正选择作用位点的总结，可以发现经历正选择的所有基因有3个共同特点。(1)这些基因绝大多数与植物的生存和繁殖等密切相关。该基因的缺失往往会导致植物体在某一环境下死亡或不育。无论是与生殖相关的决定花粉识别的花粉壁蛋白基因和吸引传粉者的花香基因，还是与抗逆相关的决定植物对各类病原抗性的基因以及直接决定藻类光合作用效率的藻胆蛋白基因，均符合这个特点。对此可以这样理解：一旦环境发生改变，这些与个体或群体生存息息相关的基因容易受到更大的选择压力作用，因此发生有利突变的等位基因更易于被固定下来而出现正选择位点。(2)基因所决定的表现型往往容易受到环境因子的作用。在植物体中符合上述第1个特点的基因位点比比皆是，但并不是都发生了正选择作用。藻类的藻红蛋白位于整个光合系统的最外层，因此光质和光强的改变更容易形成选择压力影响其进化。更极端的例子是作为高效除草剂作用靶位的乙酰辅酶A羧化酶基因，在短短的15代后就表现出正选择作用的迹象。这可能是因为环境因子形成的选择压力易于传递到此类基因位点上，进而影响其进化。(3)在这些基因进化过程中环境因子往往发生过巨大变化。植物体与病原微生物之间的军备竞赛可以理解为彼此形成对方的环境因子，在作为选择压力作用于对方的进化后，又受到对方进化形成的新的选择压力的作用。造成该特点的原因可能是，剧烈变化的环境因子能形成新的选择压力，从而加速了基因中有利突变的固定。

今后在研究植物的正选择作用时，对符合上述特点的基因应予以关注。因为正选择作用是环境对基因的选择作用，要深刻理解正选择作用对植物进化的影响，就必须掌握基因及环境相关的多方面知识背景和各层次信息。其中应包括：(1)通过生物信息学工具所得到的关于正选择位点的信息；(2)通过三级结构定位得到的该位点对整个蛋白生物活性的影响；(3)突变导致的相应表型效应及其与环境的关系；(4)对进化过程中生存环境变化的一定认识。

近年来，利用比较基因组学和生物信息学方法检测生物进化过程中发生的正选择作用已成为分子进化的研究趋势，相关研究将会有助于正确理解该生物发生历史中经受的环境变化及基因和蛋白等大分子物质的结构功能进化。目前在与人类相关的哺乳动物和各种模式动物的研究中，都有不少正选择作用方面的发现。但相比之下，植物的进化历史更长，显示出很高的多态性，有更加广阔的研究前景。在植物正选择作用研究中的更多新发现将会大大有益于对生命进化的深刻理解。

参考文献

Barkman T (2003). Evidence for positive selection on the floral scent gene isoeugenol-O-methyltransferase. *Mol Biol Evol* **20**, 168-172.

Benderoth M, Textor S, Windsor AJ, Mitchell-Olds T, Gershenzon J, Kroymann J (2006). Positive selection driving diversification in plant secondary metabolism. *Proc Natl Acad Sci USA* **103**, 9118-9123.

Bergelson J, Kreitman M, Stahl EA, Tian D (2001). Evolutionary dynamics of plant R-genes. *Science* **292**, 2281-2285.

Bielawski J, Yang Z (2001). Positive and negative selection in the DAZ gene family. *Mol Biol Evol* **18**, 523-529.

Bishop J, Ripoll D, Bashir S (2004). Selection on *Glycine* β-1,3-endoglucanase genes differentially inhibited by a *Phytophthora* glucanase inhibitor protein. *Genetics* **169**, 1009-1019.

Bujnicki J (2001). Understanding the evolution of restriction-modification systems: clues from sequence and structure comparisons. *Acta Biochim Pol* **48**, 935-967.

Christophe D, Cécile S, Séverine M, Valérie LC (2004). Nucleotide variability at the acetyl coenzyme a carboxylase gene and the signature of herbicide selection in the grass weed *Alopecurus myosuroides*. *Mol Biol Evol* **21**, 884-892.

Clay K, Kover P (1996). The red queen hypothesis and plant/pathogen interactions. *Annu Rev Phytopathol* **34**, 29-50.

Corre V, Roux F, Reboud X (2002). DNA polymorphism at the *FRIGIDA* gene in *Arabidopsis thaliana*: extensive

https://www.sciengine.com/doi/10.3724/CBB-t27-274

nonsynonymous variation is consistent with local selection for flowering time. *Mol Biol Evol* **19**, 1261-1271.

Ham K, Wu S, Darvill A, Albersheim P (1997). Fungal pathogens secrete an inhibitor protein that distinguishes isoforms of plant pathogenesis-related endo-beta-1,3-glucanases. *Plant J* **11**, 169-179.

Kimura M (1968). Evolutionary rate at the molecular level. *Nature* **217**, 624-626.

King JL, Jukes TH (1969). Non-Darwinian evolution. *Science* **164**, 788-798.

Lee YH, Ota T, Vacquier V (1995). Positive selection is a general phenomenon in the evolution of abalone sperm lysine. *Mol Biol Evol* **12**, 231-238.

Liang CW, Zhao FQ, Meng CX, Tan CP, Qin S (2006). Molecular cloning, characterization and evolutionary analysis of phytoene desaturase (PDS) gene from *Haematococcus pluvialis*. *World J Microb Biot* **22**, 59-64.

Martínez-Castilla LP, Martínez-Castilla ER (2003). Adaptive evolution in the Arabidopsis MADS-box gene family inferred from its complete resolved phylogeny. *Proc Natl Acad Sci USA* **100**, 13407-13412.

Metz EC, Palumbi SR (1996). Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein bindin. *Mol Biol Evol* **13**, 397-406.

Miller SR (2003). Evidence for the adaptive evolution of the carbon fixation gene *rbcL* during diversification in temperature tolerance of a clade of hot spring cyanobacteria. *Mol Ecol* **12**, 1237-1246.

Nei M (1987). Molecular evolutionary genetics. New York: Columbia University Press. pp. 91-137.

Peruta MF, Bauer C, Gros G (1981). Allosteric regulation of crocodilian haemoglobin. *Nature* **291**, 682-684.

Qin S, Zhao FQ, Tseng CK (2005). Evidence for positive selection in phycoerythrin genes of red algae and cyanobacteria *Prochlorococcus* and *Synechococcus*. *Photosynthetica* **43**, 141-146.

Rabbani S, Beyer P, Lintig J, Hugueney P, Kleinig H (1998). Induced beta-carotene synthesis driven by triacylglycerol deposition in the unicellular alga *Dunaliella bardawi*. *Plant Physiol* **116**, 1239-1248.

Ree RH, Citerne HL, Lavin M, Cronk QCB (2004). Heterogeneous selection on LEGCYC paralogs in relation to flower morphology and the phylogeny of *Lupinus* (Leguminosae). *Mol Biol Evol* **21**, 321-331.

Schein M, Yang Z, Mitchell-Olds T, Schmid KJ (2004). Rapid evolution of a pollen-specific oleosin-like gene family from *Arabidopsis thaliana* and closely related species. *Mol Biol Evol* **21**, 659-669.

Sheldon CC, Burn JE, Perez PP (1999). The FLF MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. *Plant Cell* **11**, 445-458.

Shimizu K, Purugganan M (2005). Evolutionary and ecological genomics of Arabidopsis. *Plant Physiol* **138**, 578-584.

Swanson WJ, Clark AG, Waldrip-Dail HM, Mariana FW, Aquadro CF (2001). Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. *Proc Natl Acad Sci USA* **19**, 7375-7379.

Swanson W, Nielsen R, Yang Q (2003). Pervasive adaptive evolution in mammalian fertilization proteins. *Mol Biol Evol* **20**, 18-20.

Wang J, Dudareva N, Bhakta S, Raguso RA, Pichersky E (1997). Floral scent production in *Clarkia breweri* (Onagraceae). II. Localization and developmental modulation of the enzyme S-adenosyl-L-methionine: (iso)eugenol O-methyltransferase and phenylpropanoid emission. *Plant Physiol* **114**, 213-221.

Zhao FQ, Qin S (2006a). Evolutionary analysis of phycobiliproteins: implications for their structural and functional relationships. *J Mol Evol* **63**, 330-340.

Zhao FQ, Qin S (2006b). Comparative molecular population genetics of phycoerythrin locus in *Prochlorococcus*. *Genetica* **129**, 291-299.

Zhao FQ, Zhang XW, Liang CW, Wu J, Bao Q, Qin S (2005). Genome-wide analysis of restriction-modification system in unicellular and filamentous cyanobacteria. *Physiol Genomics* **24**, 181-190.

https://www.sciengine.com/doi/10.3724/CBB-t27-274

Positive Selection in Plant Evolution

Yuan Gao[1, 2], Li Tian[1, 2], Song Qin[1*]

[1]*Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China*
[2]*Graduate School of the Chinese Academy of Sciences, Beijing 100049, China*

Abstract Positive selection fixes an allele harboring an advantageous mutation that increases individual fitness. Research into positive selection helps us understand processes of evolution. This paper reviews positive selection found in plant genes in recent years. Positive selections discovered in genes of land plants and algae are summarized, respectively. The positive selection genes found in land plants are mainly related to reproduction and resistance. This review will help further research into the search for loci under positive selection.

Key words adaptive evolution, plant, positive selection, reproduction-related genes, resistance-related genes

Gao Y, Tian L, Qin S (2008). Positive selection in plant evolution. *Chin Bull Bot* **25**, 401–406.

* Author for correspondence. E-mail: sqin@ms.qdio.ac.cn

(**责任编辑**: **孙冬花**)

第三届全国微生物遗传学青年学者学术研讨会

为展示我国微生物学和微生物遗传学青年学者的最新研究成果, 加强学术交流, 推动微生物学和微生物遗传学的发展, 第七届全国微生物学暨第三届全国微生物遗传学青年学者学术研讨会将于2008年10月12日 - 10月14日在山东济南召开。本届会议由中国微生物学会基础微生物学专业委员会和中国遗传学会微生物遗传专业委员会主办, 山东大学微生物技术国家重点实验室和中国科学院微生物所微生物资源国家重点实验室联合承办。热忱欢迎全国各地的微生物学和微生物遗传学青年工作者、同行及研究生踊跃投稿并积极参会。

联系电话: 0531-88564288, 13583118964(**李越中**); 0531-88566275, 13791034198(**刘 红**)
传真: 0531-88564288

https://www.sciengine.com/doi/10.3724/CBB-t27-274