

裸子植物 *PHYP* 和被子植物 *PHYB* 光感受区适应性进化的研究

王 静, 李 万 昌

(河南师范大学 生命科学院, 河南 新乡 453007)

摘 要:光敏色素是一类红光/远红光受体,在植物从萌发到成熟的整个生长发育过程中都起重要的调节作用。光敏色素 N 端的光感受区是发色团结合的部位,高度保守。该区域的保守性是确保光敏色素光逆转和光信号转导等功能正常运行的先决条件之一。裸子植物的 *PHYP* 基因对应于被子植物的 *PHYB* 基因,都属于 B 类光敏色素基因。为了研究这 2 个基因的光感受区是否经历选择压力以及是否经历同样的选择压力,基于分支模型,位点模型以及分支-位点模型对裸子植物中 8 条,被子植物 8 条共 16 条序列进行适应性进化分析。结果表明:在该区域构建的系统发育树中,强烈的正选择压力趋向于发生在裸子植物的近期分化的属内种间的谱系中,而在相对距离较远的古老谱系中缺少这种压力,然而,通过目前的模型并未检测到具体的适应性进化的位点;被子植物的 *PHYB* 基因中检测出 6 个发生适应性进化的位点。

关键词:裸子植物;被子植物;*PHYP*; *PHYB*;适应性进化

中图分类号:Q 941⁺.2 **文献标识码:**A **文章编号:**1001-0009(2012)14-0123-04

同义替换速率(d_s)或非同义替换的速率(d_N)是一年中或一个世代内一个同义位点上发生同义替换的数目或一个非同义位点上发生非同义替换的数目。实际上,由于不知道 2 个序列的分歧时间,因此,习惯上总是考虑一对序列的平均每个同义位点上发生同义替换的数目或平均每个非同义位点上发生非同义替换的数目。 d_s 等于 d_N 时(即 $\omega = 1$),表明没受到自然选择; d_s 大于 d_N 时($\omega < 1$)表明受到负选择; d_s 小于 d_N 时($\omega > 1$)表明受到正选择^[1]。

光敏色素(Phy)是一类红光/远红光受体,在植物整个生长发育过程中都起重要的调节作用^[2-4]。Phy 分子是由大约 1 100 个氨基酸残基的脱辅基蛋白和 1 个线性的四吡咯环生色团通过共价键链接构成。其 N 端是光感受区,包括后色胆素裂解酶亚结构域(Bilin lyase domain, BLD)和光敏色素亚结构域(Phytochrome domain, PHY),GAF(cGMP-specific phosphodiesterase/adenylate cyclases/formate hydrogen lyase transcription)结构域包含于 BLD 结构域中,是发色团结合的部位,高

度保守^[5-6];C 端是光调节区,光调节区域含有 2 个 PAS(Per/Arnt/Sim)同源重复序列和 1 个组氨酸激酶类作用区(Histidine kinase related domain, HKRD)。研究表明,GAF 结构域某些位点发生突变,可能会严重影响脱辅基蛋白和发色团的结合^[7],从而影响光敏色素的生理功能。试验的前期研究发现,在裸子植物 *PHY* 进化过程中,GAF 结构域中存在 2 个正选择位点^[8],但是并未涉及 C 端的其它结构域。*PHYB* 存在于被子植物,和裸子植物中的 *PHYP* 同为 B 类光敏色素基因。Yang 等^[9]和 Mathews 等^[10]在对被子植物光敏色素 A-C/F 基因以及光敏色素 B-D/E 基因的分析中检测到了正选择。但是 Yang 等选用的基因序列有限,每个光敏色素类型只有 1 条基因序列。样本的大小可能会影响检测结果。因此,现在前期研究工作的基础上,采用 16 条光敏色素基因序列,检测了分别存在于裸子植物和被子植物中同为 B 类光敏色素的 *PHYP* 和 *PHYB* 基因的光感受区,以期了解这 2 个基因的光感受区是否经历选择压力以及经历同样的选择压力,从而深化对该区域功能的认识。

1 材料与方法

1.1 试验材料

试验采用了 8 条裸子植物的 *PHYP* 光感受区序列,8 条被子植物的 *PHYB* 光感受区序列进行适应性进化分析。其中被子植物的序列数据以及外类群风兜地

第一作者简介:王静(1978-),女,河南温县人,博士,讲师,现主要从事植物进化研究工作。E-mail:jingwangkuang@163.com.

基金项目:“十二五”农村领域国家科技计划课题资助项目(2011BAD17B04);河南师范大学博士科研启动基金资助项目(525440)。

收稿日期:2012-03-29

钱 (*Marchantia paleacea* Bertol. var. *diptera* (Mont.) Hatt.) 的序列数据取自 GenBank(序列名称及登陆号见

表 1); 裸子植物序列由测序所得, 植物材料均采自中国科学院武汉植物园。

表 1 植物材料名称以及 GeneBank 登录号

Table 1 Names of the plant materials and the accession numbers of sequences

物种名称 Species	Genebank 登录号 Accession No.	物种名称 Species	Genebank 登录号 Accession No.
南方红豆杉 <i>Taxus wallichiana</i> var. <i>mairei</i> (Lemée et H. Léveillé) L. K. Fu et Nan Li	FJ393233	水稻 <i>Oryza sativa Japonica</i> L.	AB183525
白豆杉 <i>Pseudotsuga chienii</i> (W. C. Cheng) W. C. Cheng	FJ393234	高粱 <i>Sorghum bicolor</i> (L.) Moench	AF182394
三尖杉 <i>Cephalotaxus fortunei</i> Hook.	FJ393235	马铃薯 <i>Solanum tuberosum</i> L.	Y14572
柳杉 <i>Cryptomeria fortunei</i> Hooibrenk ex Otto et Dietr.	FJ393239	烟草 <i>Nicotiana tabacum</i> L.	L10114
日本柳杉 <i>Cryptomeria japonica</i> D. Don	FJ393238	番茄 <i>Lycopersicon esculentum</i> P. Miller	AJ002281
池杉 <i>Taxodium ascendens</i> Brongn.	FJ393240	杨 <i>Populus tremula</i> L.	AM072337
穗花杉 <i>Amentotaxus argotaenia</i> (Hance) Pilg.	FJ393237	番茄 <i>Lycopersicon esculentum</i> P. Miller	AF122901
巴山榧树 <i>Torreya fargesii</i> Franch.	FJ393236	拟南芥 <i>Arabidopsis thaliana</i> (L.) Heynh.	EU352789

1.2 总 DNA 提取、PCR、克隆和测序

用于 DNA 提取的材料为新鲜的叶片, 方法采用改进的 CTAB 法^[11]。扩增 *PHYB* 光感受区序列采用的引物是根据王静得到的南方红豆杉 *PHYB* 序列(数据未发表)和欧洲赤松 *PHYB* 的 cDNA 序列(X96738)设计, 引物为 2 对(表 2), 第 1 对扩增 BLD 结构域的序列, 第 2 对扩增了 *PHY-PAS1* 结构域的序列。测序完成后, 光感受区域序列由这 2 个部分通过 Editseq 软件拼接而成。每 25 μ L PCR 反应体系包括: 模板 50 ng, 引物各 8 pmol, *Taq* DNA 聚合酶 2 U(Takara, 大连)和由 Takara 提供的 10 倍 PCR 缓冲液 2.5 μ L 及 2.5 mmol/L dNTPs 0.5 μ L。聚合酶链式反应 (PCR) 在 Biometra PCR 扩增仪上进行 (Thermobio, 德国)。PCR 反应程序为: 94 $^{\circ}$ C 5 min; 94 $^{\circ}$ C 30 s, 55 $^{\circ}$ C 60 s, 72 $^{\circ}$ C 90 s, 35 个循环; 72 $^{\circ}$ C 10 min。

PCR 产物在 1.0 % 琼脂糖凝胶、1 \times TAE 电泳液中电泳 35 min (电压为 120 V), 切下目的片段, 用 QIAEX II (Qiagen, 德国) 柱式 DNA 胶回收试剂盒回收 PCR 产物。回收所得产物与 PMD19-T 载体在 4 $^{\circ}$ C 下连接 16 h, 然后转化导入由大肠杆菌 DH-5 α 所制的感受态细胞。感受态细胞在含有 X-Gal、IPTG 和氨苄青霉素的 LB 琼脂平板培养基上培养 12 h, 通过蓝白斑筛选获得阳性克隆。至少有 3 个阳性克隆送至英俊公司, 用 ABI377 型自动测序仪测序, 测序引物为 M13+/M13-。为了保证测序的准确性, 对所得序列的正、反链进行测定并加以校准。

表 2 扩增 BLD 和 *PHY-PAS1* 结构域的引物

Table 2 The primers of the amplification of *PHY-PAS1* and BLD domain

引物名称 Primer	引物序列 Sequence
BLD F1	5'-CTCTGACTTGGAGCCTTAC-3'
R1	5'-ATACTGGGAGACTGGGTTA-3'
<i>PHY-PAS1</i> F2	5'-GAAATGGCGTGTAGCAGCAG-3'
R2	5'-CATACGTACCTGTTCTGCATAAAAAG-3'

1.3 数据分析

根据文献[12], 对 *PHYB* 和 *PHYA* 的分析是以地钱类风兜地钱为外类群。

序列经 Clustal X 1.81^[13] 软件比对, 个别位点进行适当人工校正。采用最大简约法 (Maximum Parsimony, MP) 构建分子系统树。空位 (Gap) 被编码为缺失 (Missing)。MP 分析使用 PAUP^[14] 软件进行, 采用如下选项: 树二组重新连接 (Tree Bisection-reconnection, TBR)、启发式搜索 (Heuristic search)、多重性选择 (MULTREES option)、ACCTRAN 优化和 100 次随机附加的重复。利用自展分析 (Bootstrap, 1 000 次重复) 检验各分支的置信度。

在该研究中, 采用 PAML4^[15] 分析方法中的 3 个模型: 分支模型^[16]、位点模型^[17-18] 以及分支-位点模型^[19-20] 进行分析。采用系统发育分析得到的 MP 树文件, 以及该分析中的密码子矩阵。目的分支在图 1 中以字母标出, ω 是背景比率。整个分析采用 PAML4 软件包中的 Codeml 程序完成。

2 结果与分析

2.1 构建的系统发育树

图 1 是基于裸子植物 *PHYB* 和被子植物 *PHYA* 光感受区序列构建的 MP 树, 这 2 个类型的光敏色素分别聚在不同的 2 个分支。在 *PHYB* 分支中, 红豆杉+白豆杉与香榧+穗花杉形成姊妹群 (bootstrap=83), 然后和三尖杉聚在一起 (bootstrap=95), 池杉和柳杉+日本柳杉形成一个分支且得到有力支持 (bootstrap=100), 该分支同红豆杉科和三尖杉科聚在一起 (bootstrap=100), 共同构成欧洲赤松的姊妹群 (bootstrap=100)。在 *PHYA* 分支中, 单子叶植物水稻 (*Oryza sativa Japonica*) + 高粱 (*Sorghum bicolor*) 形成一个分支, 双子叶植物马铃薯 (*Solanum tuberosum*)、烟草 (*Nicotiana tabacum*)、番茄 (*Lycopersicon esculentum*, tomato)、杨 (*Populus tremula*)、拟南芥 (*Arabidopsis thaliana*) 形成另外一个分支, 支持率很高 (bootstrap=100)。

2.2 选择压力的分析

由表 3 可知, 裸子植物 *PHYB* 和被子植物 *PHYA* 光感受区不同模型下的似然值以及参数估计。单比率模

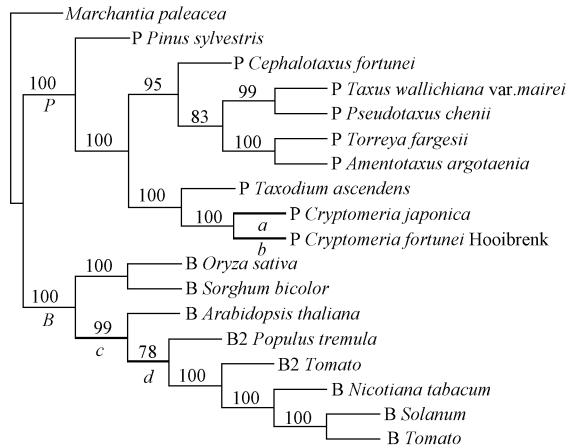


图1 基于裸子植物 PHYP 和被子植物 PHYB 光感受区构建的 MP 树, 凤兜地钱 (*Marchantia paleacea* var. *diptera*) 用作外类群

注:步长为 2 267,一致性指数 $CI=0.60$,保留指数为 $RI=0.70$;节点上的数字代表自展支持率(1 000 次重复);字母表示目的分支;加粗表示 $\omega>1$ 的分支。

Fig.1 MP tree based on photosensory domain of PHYP of gymnosperm and PHYB of angiosperm. *Marchantia paleacea* var. *diptera* was used as the outgroup

Note:2 267 steps in length,consistency index $CI=0.60$,retention index $RI=0.70$;numbers at node are bootstrap values for MP (1000 replicates);the alphabet indicate the branches of interesting;the branches with $\omega>1$ are drawn in thick lines.

表3 裸子植物 PHYP 和被子植物 PHYB 基因光感受区在不同模型下的参数估计值以及对数似然值

Table 3 Parameter estimates and Log-Likelihood values for the photosensory domain of the PHYP of gymnosperm and PHYB of angiosperm under different models

模型 Model	参数个数 P	ts/tv ts/tv	似然值 l	参数估计值 Estimates of parameters	正选择位点 Positively selected sites
分支模型					
单比率模型 M0;(one-ratio)	35	2.19	-10 535.68	$\omega_0=0.051$	None
三比率模型;(Three-ratio)	37	2.18	-10 529.55	$\omega_0=0.051$ $\omega_P=0.009$, $\omega_B=0.276$	None
自由比率模型 (Free-ratio)	67	2.26	-10 471.25	$\omega_{h'}=\infty$ $\omega_{b'}=29.37$, $\omega_{c'}=27.18$ $\omega_{d'}=259.52$	None
位点模型					
M1a;中性(M1a;neutral)	36	2.33	-10 470.32	$p_0=0.951$ ($p_1=0.049$)	Not allowed
M2a;选择(M2a;selection)	38	2.33	-10 470.32	$p_0=0.951$ $p_1=0.049$ ($P_2=0.0000$), $\omega_2=31.859$	None
M3;离散(M3;discrete)	39	2.21	-10 343.95	$p_0=0.644$, $p_1=0.281$ ($p_2=0.075$) $\omega_0=0.009$, $\omega_1=0.102$, $\omega_2=0.317$	None
M7;(beta)	36	2.21	-10 345.68	$p=0.358$ $q=5.404$	Not allowed
M8;(beta & ω)	38	2.21	-10 345.68	$p_0=1.000$ ($p_1=0.000$), $p=0.358$ $q=5.404$, $\omega=57.656$	None
分支-位点模型					
模型 A(model A)					
P*	38	2.33	-10 470.32	$p_0=0.951$, $p_1=0.049$ ($p_1+p_2=0.000$), $\omega=1.000$	None
B*	38	2.31	-10 453.35	$p_0=0.875$, $p_1=0.043$ ($p_1+p_2=0.082$), $\omega=1.964$	$P>0.99$;320R
					$P>0.95$;22G 93C 244S 261G 265W
					$P>0.50$;168R 186A 266L 285E 398I 416N

注:P,模型中参数的数目; ω_0 是指所有分支上的 d_N/d_S , ω_P , ω_B 分别是分支 P、B 的 d_N/d_S 比率; P^* , B^* ,表示在用分支位点模型检测正选择位点时,分支 P、B 是前景分支。

Note:P,number of parameters in the model; ω_0 is the d_N/d_S ratio for all branches,and ω_P , ω_B are the d_N/d_S ratios for branches P and B, respectively; P^* , B^* , indicate the branches P and B are the foreground branches for detecting positive selection under branch-site model.

3 结论与讨论

裸子植物的 PHYP 和被子植物中的 PHYB 同为 B 类光敏色素^[21],试验对这 2 个类型 PHY 基因的光感受区研究发现,被子植物 PHYB 的分支中在 $P>95\%$ 的水平上有 6 个正选择位点的出现(表 3),该研究结果和

型包括 35 个参数,在该模型下,似然值 $l_0=-10\ 535.68$, $\omega_0=0.051$ 。三比率模型假定 3 个不同的 ω ,即 ω_P 、 ω_B 以及 ω_0 ,分别表示 P、B 分支以及背景分支上的 ω 值(图 1),这 2 个模型相互比较时,自由度是 2, $2\Delta l=2(-10\ 529.55-(-10\ 535.68))=12.26$, χ^2 测验显示 $P<0.01$ 。该结果表明,这 2 个分支的确有不同的 ω 。自由比率模型假定每个分支都有自己的 ω 。在该模型中,大部分分支的 ω 小于 1,所以它们处于加强的负选择下(图 1)。但是分支 a、b、c、d 可能处于减轻的负选择甚至正选择下,因为它们的 ω 大于 1(表 3)。不过,分支 a 的 ω 无穷大是源于缺乏同义替换。在该分析中,自由模型下的似然值 $l_1=-10\ 471.25$ (表 3),该模型共有 67 个参数,和单比率模型比较,2 倍的似然值差异为 $2\Delta l=128.86$ 。依 χ^2 分布,自由度 $df=32$ 检验显著性,结果表明,该模型显著优于单比率模型。位点模型假定位点间的选择压力在变化而分支间的压力不变化,而分支-位点模型认为选择压力在位点间和分支间都改变。运用这 2 个模型可以检测那些受到正选择的位点。位点模型从分支 P 和 B 上都没有检测到受到正选择的位点,但是分支-位点模型从分支 B 上检测到了正选择位点($P>95\%$:320R、22G、93C、244S、261G、265W),而在分支 P 上没有检测到(表 3)。

Yang^[9] 在对被子植物光敏色素 B-D/E 的研究中检测到了正选择位点的结果保持一致。这说明,样本的大小对研究结果影响不大。在该区域构建的系统发育树中,强烈的正选择压力趋向于发生在裸子植物近期分化的属内种间的谱系中,而在相对距离较远的古老谱系中缺少

这种压力(图 1),并且,尽管在少数最近的 *PHYB* 分支(Recent lineages)中检测到有正选择的发生,但是没有检测到正选择位点,该结果与试验前期研究裸子植物 *PHYB* 的 PHY-PAS1 结构域适应性进化的结果类似^[22]。裸子植物 *PHYB* 和被子植物 *PHYB* 基因的这种差异可能是由光敏色素的进化速率不同引起的。据 Schneider-Poetsch 等^[21]报道,裸子植物光敏色素的进化速率要低于被子植物光敏色素。因此,裸子植物和低等植物的光敏色素进化距离要小于被子植物的光敏色素。所以,即使裸子植物 *PHYB* 的光感受区发生过正选择(图 1,分支 a,b,c,d),但是可能目前的模型还不够灵敏,不能准确鉴定出该区域中的正选择位点。

参考文献

- [1] Nei M, Kumar S. Molecular evolution and phylogenetics[M]. New York: Oxford University Press Inc, 2000.
- [2] 徐平珍,刘涛,杨莹,等.脱落酸在植物花发育过程中的作用[J]. 云南植物研究, 2007, 29(2): 215-222.
- [3] 徐燕,薛立,屈明.植物抗寒性的生理生态学机制研究进展[J]. 林业科学, 2007, 43(4): 88-93.
- [4] 王静,王艇.高等植物光敏色素的分子结构、生理功能和进化特征[J]. 植物学通报, 2007, 24(5): 649-658.
- [5] Rockwell N C, Lagarias J C. The structure of phytochrome. A picture is worth a thousand spectra[J]. Plant Cell, 2006a, 18: 4-14.
- [6] Rockwell N C, Su Y S, Lagarias J C. Phytochrome structure and signaling mechanisms[J]. Annual Review of Plant Biology, 2006b, 57: 837-858.
- [7] Schneider-Poetsch H, Braun B, Marx S, et al. Phytochromes and bacterial sensor proteins are related by structural and functional homologies[J]. FEBS LETTERS, 1991, 281: 245-249.
- [8] Wang J, Yan B, Chen G, et al. Adaptive evolution in the GAF domain of phytochromes in gymnosperms[J]. Biochemical Genetics, 2010, 48(3): 236-247.
- [9] Yang Z, Nielsen R. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages[J]. Molecular Biology and Evolution, 2002, 19: 908-917.
- [10] Mathews S, Burleigh J G, Donoghue M J. Adaptive evolution in the photosensory domain of phytochrome A in early angiosperms[J]. Molecular Biology and Evolution, 2003, 20: 1087-1097.
- [11] 苏应娟,王艇,杨维东,等.罗汉松属植物 DNA 的提取和 RAPD 分析[J]. 中山大学学报(自然科学版), 1998, 37(71): 13-18.
- [12] Schmidt M, Schneider-Poetsch HAW. The evolution of gymnosperms redrawn by phytochrome genes: The Gnetales appear at the base of the gymnosperms[J]. Journal of molecular evolution, 2002, 54: 715-724.
- [13] Thompson J D, Gibson T J, Plewniak F, et al. The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools[J]. Nucleic Acids Research, 1997, 25: 4876-4882.
- [14] Swofford, Paup D L. Phylogenetic analysis using parsimony (and other methods)[M]. Sinauer Associates: Sunderland MA, 2002.
- [15] Yang Z. PAML 4: a program package for phylogenetic analysis by maximum likelihood[J]. Molecular Biology and Evolution, 2007, 24 (8): 1586-1591.
- [16] Yang Z. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution[J]. Molecular Biology and Evolution, 1998, 15: 568-573.
- [17] Nielsen R, Yang Z. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene[J]. Genetics, 1998, 148: 929-936.
- [18] Yang Z. Maximum likelihood estimation on large phylogenies and analysis of adaptive evolution in human influenza virus A[J]. Journal of Molecular Evolution, 2000, 51(5): 423-432.
- [19] Yang Z, Wong W S W, Nielsen R. Bayes empirical Bayes inference of amino acid sites under positive selection[J]. Molecular Biology and Evolution, 2005, 22: 1107-1118.
- [20] Zhang J, Nielsen R, Yang Z. Evaluation of an Improved Branch-Site Likelihood Method for Detecting Positive Selection at the Molecular Level[J]. Molecular Biology and Evolution, 2005, 22: 2472-2479.
- [21] Schneider-Poetsch H, Kolukisaoglu Ü, Clapham D H, et al. Non-angiosperm phytochromes and the evolution of vascular plants[J]. Plant Physiology, 1998, 102: 612-622.
- [22] 王静,王艇,苏应娟,等.裸子植物中光敏色素 PHY-PAS1 结构域的适应性进化[J]. 植物学报, 2009, 44(5): 608-618.

Adaptive Evolution in the Photosensory Module of *PHYB* Gene in Gymnosperms and *PHYB* Gene in Angiosperm

WANG Jing, LI Wan-chang

(Life Science College of Henan Normal University, Xinxing, Henan 453007)

Abstract: Phytochromes, which are the best-characterized plant photosensors, receive and transmit signals generated by red and far-red light. They can regulate a wide range of developmental and physiological responses to changes in light intensity and spectral quality during plant development. The N-terminal of phytochrome is photosensory module, which is covalently attached to the bilin chromophore. The conservation of the photosensory module is one of the functional prerequisites for phytochrome photoconversion and signal transduction. *PHYB* is the gymnosperm orthologue of *PHYB*. Both of the genes belong to B-type gene. In order to identify lineages and sites under positive selection, the molecular evolution of 8 gymnosperm and 8 angiosperm photosensory module sequences were assayed under the branch model, the site model and the branch-site model established by Yang et al in this study. The results showed that strong positive selective pressure tends to occur in the recent *PHYB* lineages of congeneric species but absent in old lineages consisting of distantly related species; 6 positively selected sites have been identified in *PHYB* photosensory module of angiosperms.

Key words: gymnosperm; angiosperm; *PHYB*; *PHYB*; adaptive evolution