

毛果杨全基因组 ANK 基因的鉴定及表达分析

龙甜甜, 张红梅, 王 婕, 项 艳*

(安徽农业大学林学与园林学院, 合肥 230036)

摘 要: 锚蛋白重复序列 (ankyrin repeats) 是普遍存在于生物体中的介导蛋白与蛋白相互作用的一种结构域。具有这种结构域的蛋白参与转录的启动、细胞周期的调控、细胞骨架系统的形成及离子的运输和信号转导。利用生物信息学方法, 以杨树最新基因及蛋白序列为基础, 通过对毛果杨全基因组中 ANK 基因的鉴定, 获得较为完整的 ANK 家族信息。同时, 对该家族成员的分类、进化、定位、结构和表达进行全面分析, 为下一步研究 ANK 基因的功能提供重要的生物信息依据。

关键词: 杨树; ANK; 表达; 生物信息学

中图分类号: S792.11

文献标识码: A

文章编号: 1672-352X (2013)02-0210-09

Genome-wide identification and expression analysis of ANK gene family in *Populus trichocarpa*

LONG Tian-Tian, ZHANG Hong-mei, WANG Jie, XIANG Yan

(School of Forestry and Landscape Agriculture, Anhui Agricultural University, Hefei 230036)

Abstract: The ankyrin repeat (ANK) is one of the most common structure motifs mediating protein to protein interaction. The repeat found in proteins participated in transcriptional initiation, cell cycle regulation, cytoskeletal, ion transportation and signal transduction. Based on the latest genome protein sequences of poplar, ANK genes and comparatively complete gene family information were obtained by bioinformatics methods. Subsequently, classification, phylogenetic analysis, chromosomal localization, gene structure analysis and gene expression pattern analysis were also investigated. These results provide an important clue for functional study of ANK genes.

Key words: poplar; ANK; gene expression; bioinformatics

1987年, Breeden 和 Nasmyth 从酵母中发现了 ANK(ankyrin repeat) 模体^[1]。随后 Bennett 在细胞骨架锚蛋白(ankyrin) 中发现了含有 24 个此重复序列模体的结构^[2], 并将其命名为锚蛋白重复序列模体 (ANK repeats motif)。锚蛋白重复序列(ANK)是生物体中广泛利用的一种序列模体。典型的 ANK 重复序列一般含有 33 个氨基酸残基, ANK 模体在 ANK 结构域中折叠成 $\beta_2\alpha_2$ 结构, 在空间上则形成 L 型结构^[3-4]。多数 ANK 蛋白家族一般含有 4~6 个 ANK 重复序列, 数目不等的 ANK 重复序列串联起来, 依靠氢键和疏水作用, 组成紧密、稳定的蛋白结构域。由于该结构域介导蛋白与蛋白的相互作用^[5-6], 可以与多种配体结合, 因此可以实现纷繁复杂的生物学

功能^[7]。

含有 ANK 重复序列的蛋白质分布十分广泛, 从病毒、原核生物、真核生物到人体; 从细胞质到细胞外介质都能找到^[8]。这类蛋白质数量众多, 形成了一个大的家族, 其中主要包括 CDK (cyclin-dependent kinase)抑制剂, 转录调控因子, 细胞骨架形成体(cytoskeleton organisers), 发育调控子, 膜蛋白和毒素等^[9-10]。

目前, 对 ANK 基因家族的系统性研究主要集中在模式植物拟南芥和水稻中^[11-12], 而对其他物种特别是木本植物中 ANK 基因的研究还停留在起步阶段。对木本植物基因组中 ANK 基因进行鉴定和分析, 将有助于完善该家族的功能及进化信息。作者

收稿日期: 2012-10-31

基金项目: 国家科技支撑计划子课题 (2009BADA6B006-1) 资助。

作者简介: 龙甜甜, 女, 硕士研究生。E-mail: ltt89295@163.com

* 通信作者: 项 艳, 女, 博士, 教授。E-mail: xiangyan@aahu.edu.cn

利用生物信息学方法, 以毛果杨最新基因及蛋白序列为基础, 通过对毛果杨全基因组中 ANK 基因的鉴定, 获得较为完整的毛果杨 ANK 家族信息。同时, 对该家族成员的分类、进化、定位、结构和表达进行全面分析, 为下一步 ANK 基因家族的功能研究奠定生物信息学基础。

1 材料与方法

1.1 ANK 基因的鉴定

首先, 在 NCBI 在线服务器中以“ANK”为关键词搜索拟南芥标准 ANK 基因^[13], 结合 Pfam 在线分析软件(<http://pfam.janelia.org/search/sequence>)^[14], 获取 ANK 基因的 ANK 保守结构域序列的隐马氏模型(HMM)^[15]。同时借助 DNATOOLS 软件对下载的毛果杨全基因组氨基酸序列建立本地数据库, 接着将获得的 HMM 序列作为查询对象与已建立的本地毛果杨全基因组氨基酸序列数据库进行 TBLASTN 搜寻同源序列, 搜索参数设定为 10^{-3} , 以期获得最大数量的候选蛋白序列。将得到的候选序列导入 MEGA v4.1^[16]软件中, 使用工具 Clustal W (v1.83)^[17]进行多重序列比对, 去除冗余序列。最后, 为了保证筛选到的 ANK 蛋白序列的准确性, 将去除重复序列后剩下的蛋白序列导入 Pfam 在线工具进行蛋白结构分析验证, 去除 ANK 结构域不完整的蛋白序列。

将每个 ANK 家族的氨基酸序列导入 ExPasy (http://au.expasy.org/tools/pi_tool.html) 站点, 使用 Compute pI/MW 工具, 获得该家族所有蛋白的分子量、长度以及等电点等属性。

1.2 ANK 基因染色体物理位置定位及命名

目前国内外尚未建立基因在杨树染色体上定位的信息库。本研究通过 NCBI 数据库提供的染色体定位工具进行 BLAST 比对, 以期获得较为准确的候选基因的起始位点。所有的毛果杨 ANK 类型基因使用这种方法确认起始位点, 使用 MapInspect 软件 (<http://www.plantbreeding.wur.nl/uk/software-mapinspect.html>)^[13], 并结合已公布的三倍体杨 [*P. tomentosa* × *P. bolleana*] × *P. tomentosa*] ANK 基因命名特性, 将所有预测得到的毛果杨 ANK 基因在杨树染色体上的物理位置绘制成图, 基因的名称大体上按照在染色体上的不同位置进行命名。

1.3 ANK 结构域分析和分类

运用 SMART(<http://smart.embl-heidelberg.de/>) 及 Pfam 数据库, 分析每个 ANK 基因的结构域组成。将杨树每个 ANK 基因的蛋白序列输入 SMART 数据

库^[18], 可以得到每个基因的结构域分布图。然后参考 Pfam 数据库的分析结果, 将毛果杨的 75 个基因分为 10 个亚类。

1.4 ANK 基因系统进化树的构建

运用 Clustal W(v1.83)^[17]工具, 对毛果杨所有的 ANK 基因进行多重序列比对。然后将比对结果用 MEGA v4.0^[16]软件打开, 运用邻近法(Neighbor-Joining method) (bootstrap = 1 000)进行进化树的构建, 为分析 ANK 基因家族的进化关系提供理论依据。

1.5 ANK 基因家族 motif 分析

使用 GENESTUDIO 软件, 将所有 ANK 蛋白序列拼接到一个 TXT 文本文件中。然后, 登陆 MEME^[19]在线网站, 递交拼接结果, 操作参数设定如下: (1) 最大可查询基序数 50; (2) 最小可查询基序数 6; (3) 最大基序蛋白长度 50。该网站会在未来一天内以邮件的形式反馈查询 MEME 保守基序的研究结果。

1.6 ANK 基因的内含子和外显子分析

首先, 在杨树基因组数据库中以 ANK 蛋白的登录号为关键词进行搜索, 获得所有 ANK 基因及其完整的编码序列(coding sequence, CDS)。然后, 将 ANK 基因及 CDS 序列分别在线递交到 GSDS 网站系统中, 递交完信息后即可获得杨树所有 ANK 基因结构示意图。

1.7 ANK 基因表达谱聚类分析

杨树微阵列数据来源于 PopGeneIE 网站 (<ftp://aspnas.fysbot.umu.se/>), 这些数据都是经过归一化处理后的 Ratio 值, 可以直接用于下一步分析。利用 Genesis^[20]工具对获得的数据进行标准化和分层化处理, 完成聚类分析电子图谱的绘制。

2 结果与分析

2.1 ANK 基因家族的鉴定及命名

利用 NCBI 在线网站并结合 Pfam 在线蛋白数据库, 获得拟南芥 ANK 标准类型基因(PF03106)的 HMM 蛋白序列。将获得的 HMM 序列作为查询对象与已建立的本地杨树全基因组氨基酸序列数据库进行 TBLASTN 搜寻同源序列, 初步筛选获得 80 个 ANK 基因。利用 MEGA v4.1 软件进行多重序列比对, 同时结合 Pfam 在线工具对候选蛋白序列进行保守结构域分析, 删除重复序列和结构域残缺的基因, 最终确定了 75 个 ANK 类型基因。将获得的杨树 ANK 类型基因的编码蛋白序列导入 ExPasy 站点, 获得所有 ANK 蛋白的分子量、氨基酸数量以及

等电点等理化属性。

通过 NCBI 综合数据库提供的染色体定位工具进行 BLAST 比对, 获得所有杨树 ANK 基因的起始位点。使用 Mapinspect 基因定位软件, 初步获得杨树 ANK 基因在染色体的物理定位信息。为了统一杨树基因命名法, ANK 类型基因命名规则如下: 参考三倍体杨[(*P. tomentosa* × *P. bolleana*) × *P. tomentosa*]

中对其基因的命名方式, 即该类基因统称为“ANK”, 并在其前缀加上代表物种的特异性字母, 如: 杨树中的 ANK 类型基因, 命名统一为 *PtANK*; 登录 NCBI 网站, 获得已公布的 38 条三倍体杨 ANK 基因。再根据基因定位在染色体位置的不同, 获得 *PtANK* 的后缀数字, 即组成了完整的命名系统 (表 1)。

表 1 杨树中 ANK 蛋白数目及其属性
Table 1 A catalog of ANK proteins in the poplar

基因名称 Gene name	序列号 Sequenced ID	氨基酸数量 Number of amino acids	分子量/kDa Molecular weight	等电点/pI Isoelectric point	染色体位置 Chromosome
<i>PtANK1</i>	POPTR_0001s13070.1	711	78 816.5	7.12	1
<i>PtANK2</i>	POPTR_0001s18810.1	523	58 486.8	9.25	1
<i>PtANK3</i>	POPTR_0001s24570.1	442	47 953.6	6.63	1
<i>PtANK4</i>	POPTR_0001s06680.1	1 621	178 365.2	5.89	1
<i>PtANK5</i>	POPTR_0001s28170.2	1 604	176 228.6	5.93	1
<i>PtANK6</i>	POPTR_0001s28170.1	1 616	177 717.5	5.94	1
<i>PtANK7</i>	POPTR_0001s36320.1	665	73 303.8	9.35	1
<i>PtANK8</i>	POPTR_0001s34120.1	753	81 260.3	6.19	1
<i>PtANK9</i>	POPTR_0001s30240.1	320	35 484.8	5.12	1
<i>PtANK10</i>	POPTR_0005s10790.1	443	50 494.1	9.00	1
<i>PtANK11</i>	POPTR_0857s00210.1	259	28 692.9	4.80	1
<i>PtANK12</i>	POPTR_0002s22820.1	478	51 321.5	6.69	2
<i>PtANK13</i>	POPTR_0001s01270.1	250	26 722.7	6.08	3
<i>PtANK14</i>	POPTR_0003s10240.1	249	26 636.5	6.27	3
<i>PtANK15</i>	POPTR_0001s01190.1	176	19 665.2	7.79	3
<i>PtANK16</i>	POPTR_0003s10350.1	176	19 626.1	7.84	3
<i>PtANK17</i>	POPTR_0003s16170.1	830	92 455.9	6.65	3
<i>PtANK18</i>	POPTR_0004s06320.1	450	51 044.6	6.94	4
<i>PtANK19</i>	POPTR_0004s21970.3	366	39 138.2	4.47	4
<i>PtANK20</i>	POPTR_0004s21980.1	365	39 057.1	4.46	4
<i>PtANK21</i>	POPTR_0005s07030.1	526	57 329.2	7.80	5
<i>PtANK22</i>	POPTR_0005s07040.1	541	59 166.3	8.70	5
<i>PtANK23</i>	POPTR_0005s24070.1	337	37 295.7	8.87	5
<i>PtANK24</i>	POPTR_0005s24680.1	261	29 050.2	6.52	5
<i>PtANK25</i>	POPTR_0018s06510.1	877	98 960.2	6.84	6
<i>PtANK26</i>	POPTR_0006s19070.1	507	54 949.7	7.92	6
<i>PtANK27</i>	POPTR_0006s23100.1	585	64 155.8	9.54	6
<i>PtANK28</i>	POPTR_0006s23110.1	538	59 231.8	9.24	6
<i>PtANK29</i>	POPTR_0006s24010.1	491	55 606.7	9.28	6
<i>PtANK30</i>	POPTR_0006s24020.1	471	52 575.5	8.77	6
<i>PtANK31</i>	POPTR_0006s26600.1	706	78 955.3	5.91	6
<i>PtANK32</i>	POPTR_0007s00400.1	663	72 659.9	9.58	7
<i>PtANK33</i>	POPTR_0007s00720.1	505	57 033.5	9.16	7
<i>PtANK34</i>	POPTR_0007s04760.1	504	54 979.5	7.81	7
<i>PtANK35</i>	POPTR_0008s04830.1	588	64 272.4	9.08	8
<i>PtANK36</i>	POPTR_0010s08160.1	455	49 500.7	5.62	8
<i>PtANK37</i>	POPTR_0008s17920.1	656	72 725.4	6.62	8

续表 1 Continued table 1

<i>PtANK38</i>	POPTR_0009s07500.1	153	16 322.1	4.54	9
<i>PtANK39</i>	POPTR_0001s28280.2	153	16 337.2	4.53	9
<i>PtANK40</i>	POPTR_0001s28280.1	153	16 406.3	4.53	9
<i>PtANK41</i>	POPTR_0010s06740.1	760	82 877.3	6.77	10
<i>PtANK42</i>	POPTR_0010s06580.1	641	70 373.1	6.97	10
<i>PtANK43</i>	POPTR_0010s16870.1	244	26 015.4	4.89	10
<i>PtANK44</i>	POPTR_0010s19260.1	480	52 151.3	5.93	10
<i>PtANK45</i>	POPTR_0010s21950.1	587	64 196.3	9.26	10
<i>PtANK46</i>	POPTR_0011s15450.1	753	82 189.2	6.11	11
<i>PtANK47</i>	POPTR_0017s02430.1	841	96 568.7	6.43	12
<i>PtANK48</i>	POPTR_0012s04000.1	820	93 848.1	6.46	12
<i>PtANK49</i>	POPTR_0004s08170.1	842	96 464.7	6.48	12
<i>PtANK50</i>	POPTR_0012s02480.1	245	26 844.9	4.82	12
<i>PtANK51</i>	POPTR_0012s13410.1	511	55 237.9	6.19	12
<i>PtANK52</i>	POPTR_0013s03960.2	430	46 233.2	5.09	13
<i>PtANK53</i>	POPTR_0013s03960.1	456	48 950.7	4.98	13
<i>PtANK54</i>	POPTR_0005s05650.1	474	51 222.5	4.93	13
<i>PtANK55</i>	POPTR_0013s05860.1	429	46 444.3	7.82	13
<i>PtANK56</i>	POPTR_0015s01250.1	353	38 243.4	4.60	15
<i>PtANK57</i>	POPTR_0014s05000.1	459	50 597.6	6.53	15
<i>PtANK58</i>	POPTR_0012s11530.1	448	50 046.6	6.34	15
<i>PtANK59</i>	POPTR_0012s11540.1	169	19 506.7	9.48	15
<i>PtANK60</i>	POPTR_0015s13360.1	511	55 138.0	6.16	15
<i>PtANK61</i>	POPTR_0016s04010.1	443	48 839.9	6.37	16
<i>PtANK62</i>	POPTR_0016s08150.1	542	59 489.8	9.10	16
<i>PtANK63</i>	POPTR_0016s10820.1	197	21 220.2	4.26	16
<i>PtANK64</i>	POPTR_0006s12330.1	201	21 400.5	4.17	16
<i>PtANK65</i>	POPTR_0017s03960.1	623	68 732.4	9.59	17
<i>PtANK66</i>	POPTR_0018s00390.1	509	56 926.6	8.55	18
<i>PtANK67</i>	POPTR_0018s01360.1	714	80 209.6	6.28	18
<i>PtANK68</i>	POPTR_0018s05940.1	521	57 941.2	7.31	18
<i>PtANK69</i>	POPTR_0018s05990.1	488	54 762.4	8.06	18
<i>PtANK70</i>	POPTR_0018s06000.1	479	53 607.3	8.50	18
<i>PtANK71</i>	POPTR_0018s06020.1	641	71 593.5	6.41	18
<i>PtANK72</i>	POPTR_0130s00210.1	478	53 401.1	6.59	18
<i>PtANK73</i>	POPTR_0018s08770.1	477	53 379.1	6.40	18
<i>PtANK74</i>	POPTR_0006s16110.1	478	53 410.2	6.34	18
<i>PtANK75</i>	POPTR_0018s10750.1	509	55 293.1	8.02	18

2.2 ANK 基因家族的染色体定位

为了研究 ANK 基因家族在杨树染色体上的定位, 本研究通过 NCBI 数据库提供的染色体定位工具进行 BLAST 比对, 获得较为准确的 ANK 基因的起始位点信息。利用染色体定位绘图软件 Mapinspect, 获得毛果杨 ANK 基因家族染色体定位图。

图 1 展示了 ANK 类型基因在杨树 18 条染色体上的分布, 可以明显地看出该类型基因在染色体上

分布不太均匀。75 个 ANK 基因广泛的分布于杨树的 17 条染色体上, 除 14 号染色体没有分布。ANK 基因在染色体上具体分布情况如下: 1 号染色体含有最多的 ANK 基因, 共有 11 个 ANK 成员, 其次为 18 号染色体, 有 10 个基因分布; 7 个基因定位在 6 号染色体上, 3、10、12 和 15 号染色体都含有 5 个 ANK 基因; 5、13 和 16 号染色体都含有 4 个 ANK 基因; 7、8 和 9 号染色体都含有 3 个 ANK 基因; 2、11 和 17 号染色体都只含有 1 个 ANK 基因。

2.3 ANK 基因家族的分类

运用 SMART 软件(<http://smart.embl-heidelberg.de/>) 和 Pfam 数据库, 分析了每个 ANK 基因结构域 (如图 2), 将杨树 ANK 基因家族中 75 个基因分为 10 个亚类。表格 2 展示了 28 个 ANK 基因仅含有锚蛋白重复序列, 命名为 ANK-M 亚类。ANK-TM (20

个) 亚类除含有锚蛋白重复序列之外还含有锚定膜蛋白, ANK-TPR, ANK-AAA 和 ANK-RT 亚类都含有 3 个 ANK 基因, ANK-TC 和 ANK-RF 亚类含有 6 个 ANK 基因, ANK-PA、ANK-ST、ANK-O 含有 2 个 ANK 基因。

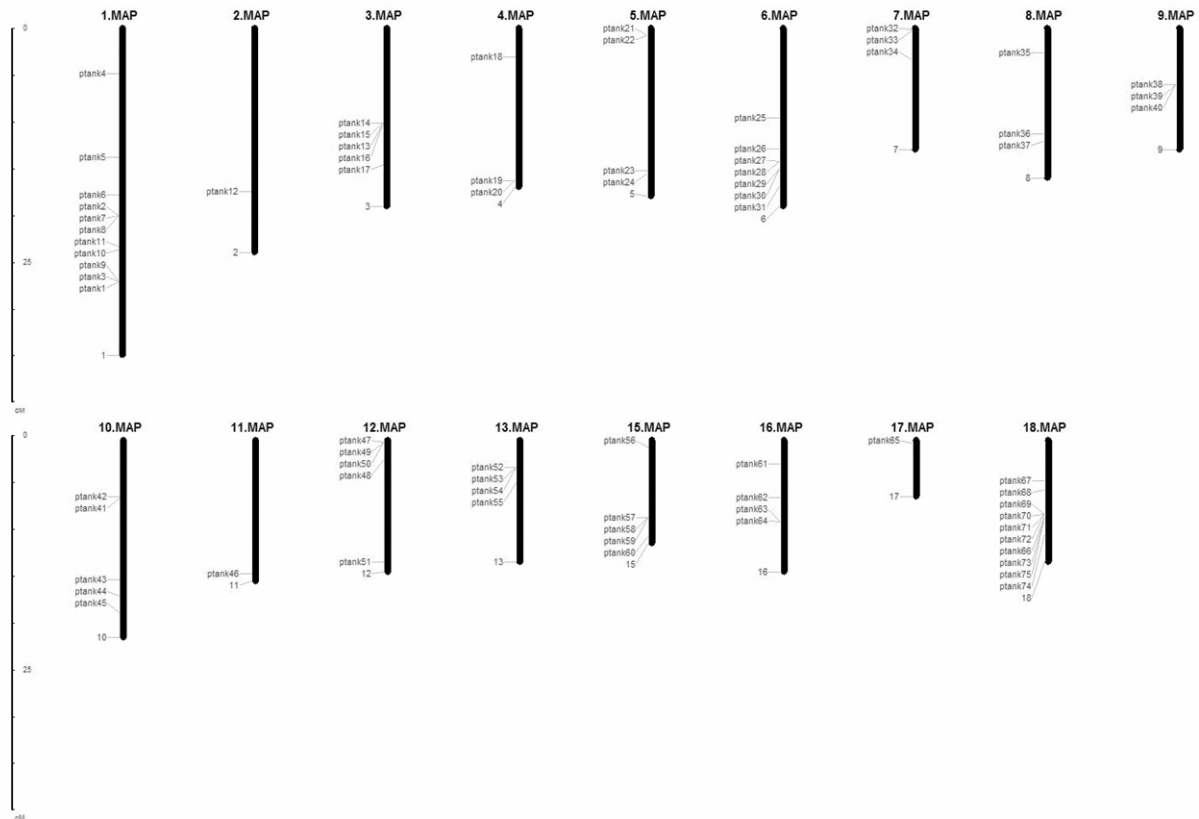


图 1 杨树 ANK 基因的染色体定位

Figure 1 Physical locations of the poplar ANK genes in poplar chromosomes

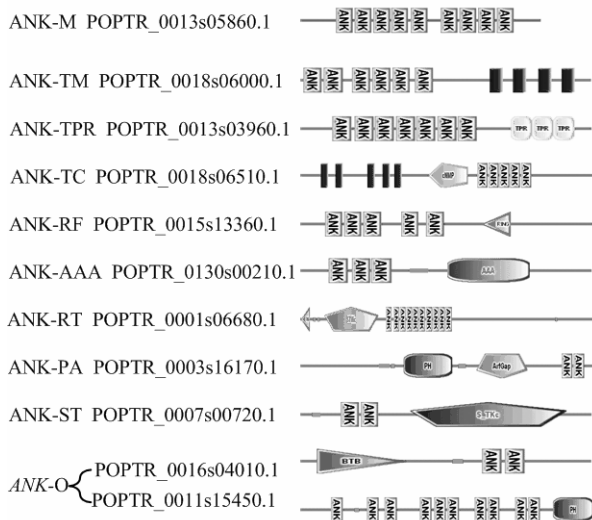


图 2 每个亚家族的代表基因及其结构域特点

Figure 2 Representative *PtANK* and its domains of each subfamily

表 2 杨树 ANK 基因亚家族基因数目分布

Table 2 The quantitative distribution of ANK gene subfamily in poplar

亚类 Subfamily	信息详述 Detail information	总数 Total
ANK-M	Proteins with only ankytin repeat	28
ANK-TM	Ankyrin-transmembrane proteins	20
ANK-TPR	Protein with teratrico peptide repeats	3
ANK-TC	Proteins and CNMP domain	6
ANK-RF	Ring finger domain	6
ANK-AAA	ATPase family associated with various cellular activities	3
ANK-RT	Ring finger domain and Tyrosine Kinase, catalytic domain	3
ANK-PA	Pleckstrin homology domain ArfGap domain	2
ANK-ST	Tyrosine Kinase, catalytic domain	2
ANK-O	Proteins with other domains	2
Total		75

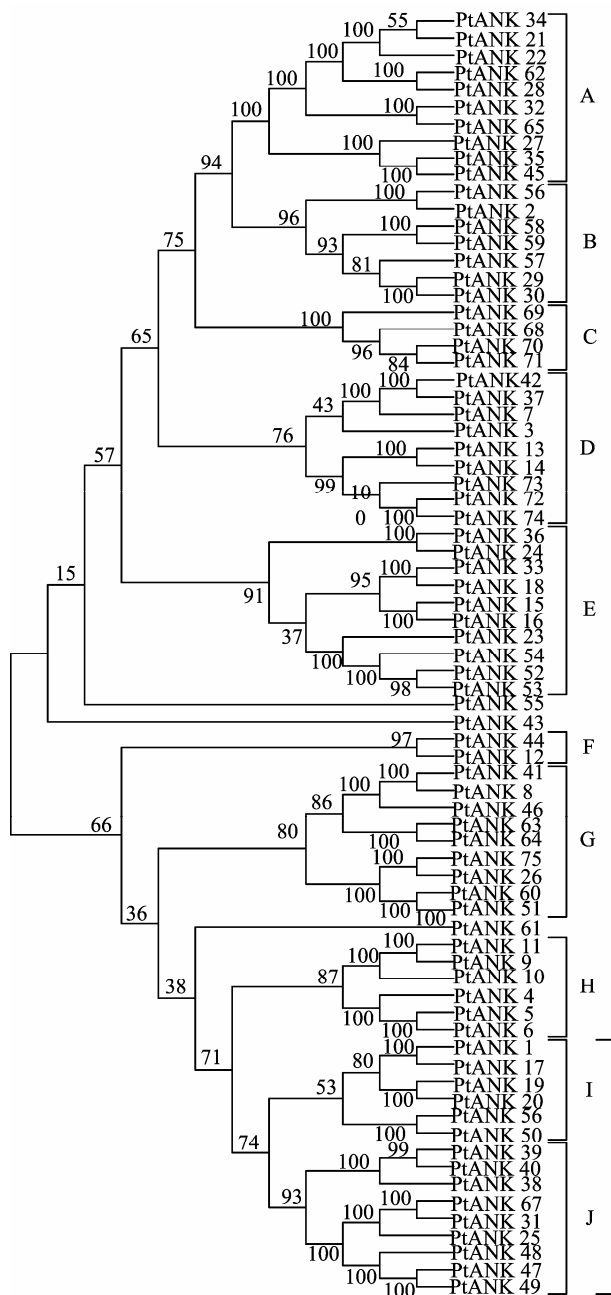


图 3 ANK 基因家族进化分析

Figure 3 Phylogenetic analysis of ANK gene family

2.4 ANK 基因家族保守结构域进化分析

使用 Clustal X(v1.83)软件, 对 75 个杨树 ANK 保守结构域进行多重序列比对, 在其运算结果的基础上, 用 MEGA v4.1 构建 *PtANK* 保守结构域的进化树 (图 3)。参照 ANK 保守结构域多重比对结果对其进行分组, 总共分为 10 组, 分别为 A-J 组, 从图中可以看出除了 A 组只含有 ANK-TM 结构域外, 同一组的 ANK 基因含有不同的结构域。如 B 组含有 ANK-AAA、ANK-RT 和 ANK-M 3 种结构域, F 组含有 ANK-0、ANK-M、ANK-AAA 和 ANK-RF 4 种结构域, 可以看出这些基因的功能是多样化的。

ANK-ST 和 ANK-TPR 结构域都分布在 E 组相邻的基因中, ANK-M 结构域分布于除了 A 组之外的每一组中相邻的基因中, 可以推断在进化树中位置相近的基因, 含有相同或相近的结构域。

2.5 ANK 基因家族保守 motif 分析

为了更加全面的了解杨树 ANK 蛋白保守区域的结构, 本研究利用 MEME 软件对其保守 motif 进行分析, 结果如图 4 所示。不同颜色的方块代表不同的 motif, 一共有 49 个预测的 motif。可以直观的看出在 *PtANK* 中存在多个保守基序, 不同类型的 ANK 含有 motif 数目和种类存在较大差异。通过 SMART 软件分析, 可以看出除了基序 1、2、3、8、9、11 和 17 外, 其他都是未知功能的基序。所有的已知功能的基序都被解释为锚蛋白重复序列。通过对 motif 进行分析有助于了解杨树 ANK 家族的功能。

2.6 ANK 基因内含子和外显子分析

利用 GSDS 在线网站对 ANK 基因家族所有基因进行内含子和外显子分析 (图 5)。每条 *PtANK* 的序列被内含子割裂成多段, 内含子数目不等, 从 0~15 个不等。属于同一亚类的基因其基因结构较为相似, 例如, *PtANK7* 和 *PtANK8* 同属于 ANK-RT 亚组, 由 2 个 UTR、16 个外显子和 15 个内含子构成; *PtANK18* 和 *PtANK33* 同属于 ANK-ST 亚组, 则由 2 个 UTR、12 个外显子和 11 个内含子构成。

2.7 ANK 基因家族表达谱聚类分析

所有杨树 ANK 基因中, 共有 9 个基因的表达谱未能从 PopGeneIE 数据库中获得, 其他 66 个基因在不同组织中的 (根、成熟的叶片、年幼的叶片、木质部、雄花、雌花) 表达谱如图 6 所示。图中不同颜色的小方块代表基因表达谱数据, 一个小块代表一个数据, 红色和绿色反映基因在同等条件下转录水平是上调还是下调, 而颜色的深浅表示数据绝对值的大小。根据基因的表达情况, 可以人为的将其划分为 5 个基因聚类(I、II、III、IV、V), 相同的聚类反映了具有相同表达特征的基因。从图中可以直观的看出, 几乎所有的基因都呈现出了组织表达特异性。

3 讨论

基因的重复、重组和序列的变异被看作为生物进化的三大基本来源, 因此 ANK 基因的结构域和进化分析, 对探索 ANK 基因家族的起源和进化具有重要的意义。本文从 ANK 基因的结构域和进化分析中, 可以看出 ANK 基因的功能是多样化的, 而且发现在进化树中位置相近的基因, 含有相同或相近结构域。

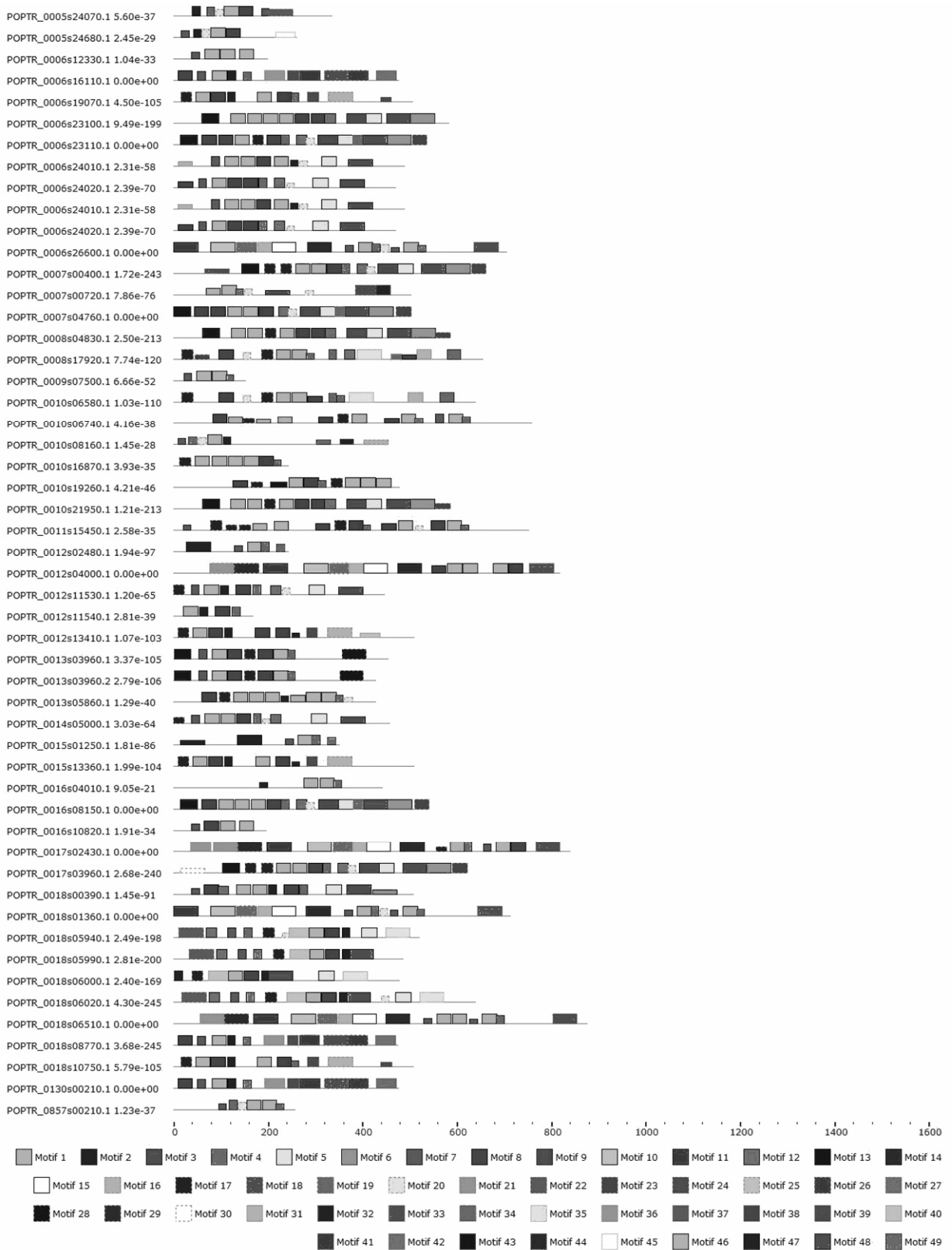


图 4 ANK 家族的保守基序分析

Figure 4 Analysis of conserved motifs in *PtANK* family

本研究利用 GSDS 在线网站对 ANK 基因家族所有基因进行内含子和外显子分析,发现 *PtANK12* 和 *PtANK55* 是杨树 ANK 家族中不含有内含子区域

的基因。这 2 个基因编码的蛋白中 ANK 结构域有部分缺失,产生的原因可能是基因长期进化过程中部分结构域选择性丢失或注释拼接过程中发生错误。

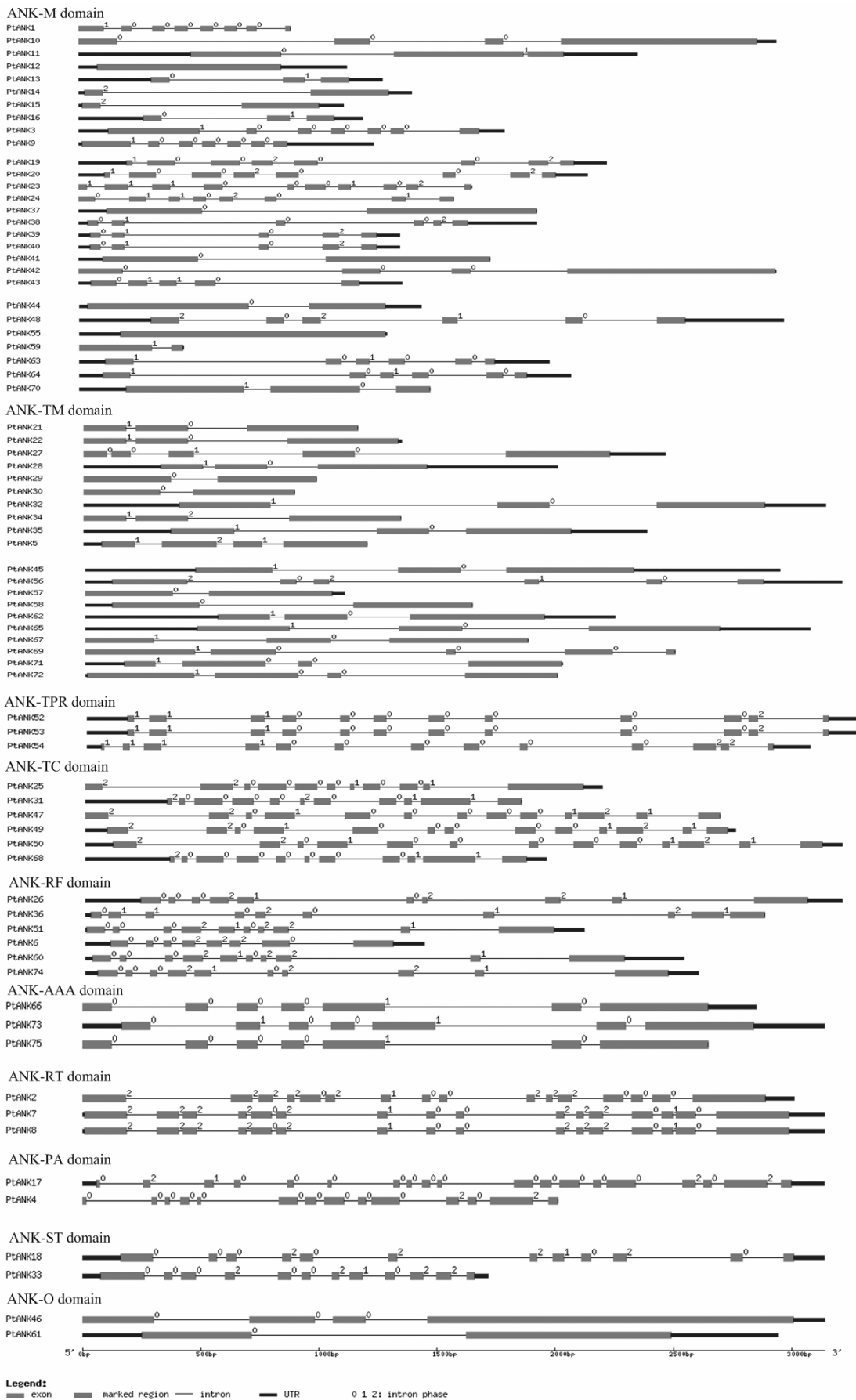


图 5 ANK 基因内含子和外显子分析

Figure 5 The intron-exon structures of *PtANK* genes

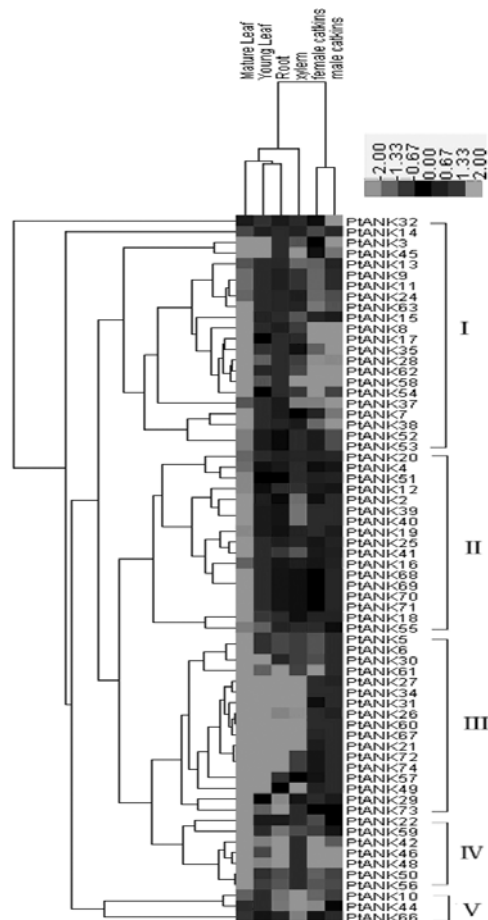


图 6 杨树 ANK 基因表达谱分析

Figure 6 Expression pattern analysis of *PtANK*s

通过对 66 个杨树 ANK 基因在不同组织中的表达谱中看出, 近似一半的基因 (35 个) 在杨树雄花中转录丰度增高。特别值得注意的是, 这些增高表达的基因几乎全部分布于 II 号和 III 号聚类基因中, 而在其他组织中转录水平较低, 暗示着 II 号和 III 号聚类基因可能在雄花的发育生长过程中发挥着重要的生物学功能。在杨树未成熟叶片中, 接近 1/3 的 ANK 基因 (28 个) 转录水平呈上调趋势, 而在成熟叶片中, 所有用于分析的 ANK 基因转录水平都呈下调趋势, 例如: *PtANK16*、*PtANK -25*、*PtANK -68*、*PtANK -69* 等基因在未成熟叶片中转录丰度接近 1.4, 而在成熟叶片中转录丰度则邻近 -2.0, 预示着该类基因在叶片生长过程中可能扮演着重要角色。通过对木质部与其他组织中表达量比较, 发现有 12 个基因 (*PtANK4*、*PtANK -22*、*PtANK -32*、*PtANK -42*、*PtANK -46*、*PtANK -48*、*PtANK -50*、*PtANK -51*、*PtANK -52*、*PtANK -53*、*PtANK -56*、*PtANK -73*) 在杨树木质部形成过程中表达水平高, 暗示着这些基因可能与木质素的形成相关, 这为下一步培育优质杨树新品种提供了参考。

参考文献:

- [1] Breeden L, Nasmyth K. Similarity between cell-cycle genes of budding yeast and fission yeast and the Notch gene of *Drosophila*[J]. *Nature*, 1987, 329: 651-654.
- [2] Lux S E, John K M, Bennett V. Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins[J]. *Nature*, 1990, 344(6261): 36-42.
- [3] Foord R, Taylor I A, Sedgwick S G, et al. X-ray structural analysis of the yeast cell cycle regulator Swi6 reveals variations of the ankyrin fold and has implications for Swi6 function[J]. *Nat Struct Biol*, 1999, 6(2): 157-165.
- [4] Jacobs M D, Harrison S C. Structure of an I κ B α /NF- κ B complex[J]. *Cell*, 1998, 95 (6): 749-758.
- [5] Michaely P, Bennett V. The ANK repeat: a ubiquitous motif involved in macromolecular recognition[J]. *Trends Cell Biol*, 1992, 2(5): 127-129.
- [6] Bork P. Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally[J]. *Proteins*, 1993, 17(4): 363-374.
- [7] Al-Khodor S, Price C T, Kalia A, et al. Functional diversity of ankyrin repeats in microbial proteins[J]. *Trends Microbiol*, 2010, 18(3): 132-139.
- [8] Mosavi L K, Cammett T J, Desrosiers D C, et al. The ankyrin repeat as molecular architecture for protein recognition[J]. *Protein Sci*, 2004, 13: 1435-1448.
- [9] Sedgwick S G, Smerdon S J. The ankyrin repeat: a diversity of interactions on a common structural framework[J]. *Trends Biochem Sci*, 1999, 24(8): 311-316.
- [10] Cunha S R, Moher P J. Ankyrin protein networks in membrane formation and stabilization[J]. *J Cell Mol Med*, 2009, 13: 4362-4376.
- [11] Becerra C, Jahrman T, Puigdomenech P, et al. Ankyrin repeat-containing proteins in *Arabidopsis*-transmembrane protein[J]. *Gene*, 2004, 340: 111-121.
- [12] Huang J, Zhao X, Yu H, et al. The ankyrin repeat gene family in rice: genome-wide identification, classification and expression profiling[J]. *Plant Mol Biol*, 2009, 71(3): 207-226.
- [13] Huang J, Zhao X, Yu H, et al. The ankyrin repeat gene family in rice: genome-wide identification, classification and expression profiling[J]. *Plant Mol Biol* 2009, 71(3): 207-226.
- [14] Finn R D, Mistry J, Schuster-Böckler B, et al. Pfam: clans, web tools and services[J]. *Nucleic acids research*, 2006, 34(suppl 1): D247-D251.
- [15] Eddy S R. Profile hidden markov models[J]. *Bioinformatics*, 1998, 14: 755-763.
- [16] Tamura K, Dudley J, Nei M, et al. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0[J]. *Mol Biol Evol*, 2007, 24: 1596-1599.
- [17] Thompson J D, Higgins D G, Gibson T J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice[J]. *Nucleic Acids Res*, 1994, 22(22): 4673-4680.
- [18] Schultz J, Milpetz F, Bork P, et al. SMART, a simple modular architecture research tool: identification of signaling domains[J]. *Proc Natl Acad Sci U S A* 1998, 95(11): 5857-5864.
- [19] Bailey T L, Williams N, Misleh C, et al. MEME: discovering and analyzing DNA and protein sequence motifs[J]. *Nucleic Acids Res*, 2006, 34(Web Server issue): W369-W373.
- [20] Sturn A, Quackenbush J, Trajanoski Z. Genesis: cluster analysis of microarray data[J]. *Bioinformatics*, 2002, 18(1): 207-208.