

小麦 *TaCAT* 新基因克隆及分子生物学和生化特性分析

余茂云, 雷昊, 殷桂香, 冯晨, 别晓敏, 杜丽璞, 徐惠君, 叶兴国

中国农业科学院作物科学研究所; 国家农作物基因资源与基因改良重大科学工程; 农业部作物遗传育种重点开放实验室, 北京 100081

摘要 小麦组织培养再生潜力与抗氧化胁迫能力具有相关性。为解析小麦 *TaCATs* 基因家族的分子生物学及生化特性并为后续试验验证奠定理论基础, 利用 e-PCR 方法进行小麦过氧化氢酶(*CAT*)新基因克隆, 结合 *in silico* 技术对小麦 *TaCATs* 基因家族的生化特性进行分析和预测。氨基酸序列同源性比对结果表明, 克隆的小麦 *TaCATs* 基因家族新成员与水稻的 *CatA* 和玉米的 *Cat3* 具有较高的相似性, 分别达 89% 及 81.1%, 命名为 *TaCAT3*, 基因组 DNA 和 cDNA 长度分别为 1986 和 1482bp, 编码 494 个氨基酸的蛋白。亚细胞定位结果表明, *TaCAT3* 可能定位在线粒体中, 且所有的功能活性位点在小麦 *TaCATs* 家族中具有一致的保守性。系统发生树构建结果表明, 小麦 *TaCATs* 能形成 3 个独立的分支。在蛋白质编码序列同源比对的基础上, 利用 SwissModel 的 Swiss-PdbViewer 3.7 软件包对 *TaCATs* 高级结构进行同源模拟, 发现所构建的模型能很好地反映 *TaCATs* 的高级结构。

关键词 小麦; *TaCAT*; *in silico*; 同源建模

中图分类号 Q785

文献标识码 A

doi 10.3981/j.issn.1000-7857.2011.12.002

Clone of *TaCAT* Genes in Wheat (*Triticum aestivum* L.) and Its Biochemical and Molecular Biological Characterization

SHE Maoyun, LEI Hao, YIN Guixiang, FENG Chen, BIE Xiaomin, DU Lipu, XU Huijun, YE Xingguo

National Key Facility of Crop Gene Resources and Genetic Improvement; Key Laboratory of Crop Genetic and Breeding, Ministry of Agriculture; Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China

Abstract There is a close relationship between *in vitro* regeneration potential of wheat (*Triticum aestivum* L.) tissues and their resistant ability to oxidative stress. In order to dissect and provide theoretical evidences for subsequent experimental molecular functional characterization of *TaCATs*, e-PCR method and *in silico* techniques are employed for the isolation of the new member of wheat *TaCAT* family and the prediction on biochemical characters of *TaCATs*, respectively. *TaCAT3*, as a new member in wheat *CAT* gene family, is encoded a deduced 494-amino acid protein with the 1986bp and 1482bp of genomic DNA and cDNA, respectively, it shares 89% homology with *CatA* (*Oryza sativa* L.) and 81.1% identity with *Cat3* (*Zea mays* L.), and is possibly targeted in the mitochondria according to the subcellular localization prediction. Furthermore, all the functional sites in *TaCATs* are highly conserved. Phylogenetic tree construction shows that the wheat *TaCATs* are able to form three separate branches with one isoform each. Swiss-PdbViewer 3.7 software package is employed for homology modeling of *TaCATs* after an appropriate template is acquired in Swiss Model online. The reconstructed model could well reflect the *TaCAT* senior structure.

Keywords wheat; *TaCAT*; *in silico*; homology modeling

收稿日期: 2011-02-23; 修回日期: 2011-03-21

基金项目: 国家自然科学基金项目(30971776); 《科技导报》博士生创新研究计划(KJDB200902-14)

作者简介: 余茂云, 博士研究生, 研究方向为小麦生物技术育种, 电子信箱: ahxiaoshe@126.com; 叶兴国(通信作者), 研究员, 研究方向为小麦生物技术育种, 电子信箱: yexg@mail.caas.net.cn

0 引言

过氧化氢酶(catalase,缩写为CAT;EC:1.11.1.6)是一种清除生物体内多余H₂O₂并将其还原为分子氧和水的氧化还原酶^[1-2]。在大多数单子叶植物中主要存在3种CAT基因,编码不同亚型的同工酶^[3-4]。在水稻(*Oryza sativa* L.)和玉米(*Zea mays* L.)皆已发现3种CAT编码基因,分别是*CatA*^[5]、*CatB*^[6]、*CatC*(GenBank登录号:DQ118681)和*Cat1*^[7]、*Cat2*^[8]、*Cat3*^[9]。其中,水稻*CatA*、*CatB*和*CatC*分别与玉米*Cat3*、*Cat1*及*Cat2*互为直向同源关系(orthologous genes)^[9-10],其表达具有时序性和器官特异性^[6,8,11],且*CatA*及*Cat3*编码蛋白为单子叶植物所特有^[3]。此外,蓖麻(*Ricinus communis* L.)^[12]、马铃薯(*Solanum tuberosum* L.)^[13]、大麦(*Hordeum vulgare* L.)^[14]和拟南芥(*Arabidopsis thaliana* L.)^[15]等也有CAT同工酶编码基因的报道。目前,在小麦(*Triticum aestivum* L.)中已见报道2个CAT同工酶编码基因^[2,16]。

H₂O₂是好氧生物有氧代谢过程中的一个重要产物,同时也是植物在遭受病原菌胁迫及逆境胁迫时产生的活性氧(Active Oxygen Species, AOS)的一种重要组成形式^[6,17]。在植物体内,过多的H₂O₂积累会导致抗氧化能力失衡,从而诱发植物发生氧化胁迫,进而启动植物体内抗氧化胁迫机制,降低因氧化胁迫造成的伤害^[6,18]。植物体内抗氧化能力减弱及氧化还原动态平衡机制破坏会直接影响组织培养过程已分化细胞的再分化,进而降低植物组织培养再生能力^[19-21]。

植物组织培养的低再生能力严重影响利用基因工程技术进行的遗传改良研究。根据基因的遗传定位,利用分子生物学手段鉴定和分离控制植株再生的主效基因已受到广泛关注。相继在菊苣(*Cichorium intybus* L.)、矮牵牛(*Petunia hybrida* L.)、泡桐(*Paulownia kawakamii* L.)、番茄(*Lycopersicon peruvianum* L.)、胡萝卜(*Daucus carota* L.)、拟南芥、水稻、玉米及大麦中分离到*AGP*^[22]、*Sho*^[23]、*PkSFI*^[24]、*Rg-1*^[25]、*SERK*^[26]、*ESR1*^[27]、*NiR*^[28]、*ZmLEC1*^[29]及*Shd1*和*KN1*等有关控制再生的候选基因,并利用转基因技术提高了这些物种组织培养的再生能力。然而,作为世界上第一大粮食作物的小麦,其再生主效基因克隆尚未见报道。据此,本课题组根据小麦再生相关基因的差别杂交结果^[31-32],以2,4-D处理与不处理的普通小麦材料CB037的幼胚为材料,采用

同源克隆及半定量RT-PCR技术成功分离到一个与小麦幼胚再生具有显著正相关的候选基因*TaCAT1*(待发表)。在此基础上,本研究利用同源克隆技术从小麦中分离到第3种CAT同工酶编码基因——*TaCAT3*,扩大了控制小麦幼胚再生能力的基因资源,同时利用*in silico*技术对该家族基因的生化 and 分子生物学特性进行预测和分析,为后续该基因的生物分子生物学鉴定提供理论依据。

1 材料与方法

1.1 植物材料及其DNA提取

普通小麦(*T. aestivum* L.)基因型CB037由中国农业科学院作物科学研究所陈孝研究员选育,本课题组保存。利用SDS法从小麦叶片中提取基因组DNA,用于小麦CAT家族基因序列扩增。

1.2 RNA提取与cDNA合成

取CB037小麦种子进行恒温培养,待种子萌动后种于蛭石中,3周后取叶片,液氮速冻并于-80℃保存。按照TRNzol Kit试剂盒(北京天根生物公司)描述的方法提取小麦叶片总RNA,依据PrimeScript 1st Strand cDNA Synthesis Kit(大连TaKaRa公司)操作说明合成cDNA第一链。

1.3 小麦TaCAT3基因序列获得

根据水稻*CatA*和玉米*Cat3*氨基酸序列,利用美国国立生物技术信息中心(NCBI)的tblastn工具搜索相关的小麦EST,用CodonCode 2.0.4(美国CodonCode公司)进行序列拼接。根据拼接结果,利用DNAMAN 6.0软件在预测的5'和3'UTR区域设计扩增cDNA和gDNA全长序列的引物,并命名为*TaCAT3*,引物序列如表1。PCR反应体系总体积为20μL,包括ddH₂O 5.8μL,2×GC Buffer 10μL(大连TaKaRa公司),引物(10μmol/L)各1μL,dNTP(各2.5mmol/L)2μL,LA Taq酶(5U/μL,大连TaKaRa公司)0.2μL,模板gDNA或cDNA(50ng/μL)1μL。PCR扩增程序为94℃变性4min,35个循环包括94℃变性50s,62℃退火40s,72℃延伸2min,最后72℃延伸10min。1%普通琼脂糖凝胶电泳检测并回收目的片段。利用NCBI的ORF Finder和在线软件FGENESH⁺进行ORF和内含子预测。同时根据已发表的小麦CAT家族另外两个成员(*TaCAT1*、*TaCAT2*)的序列,采用同样的策略扩增出cDNA和

表1 小麦TaCAT扩增引物

Table 1 Primer sets for TaCAT amplification of wheat

基因	GenBank 登录号	正向引物(5'→3')	反向引物(5'→3')	片段大小/bp
<i>TaCAT1</i>	GU984379	ATCGTTCCTCCGTCGCCAT	CGCACTCCATGGACTGATCT	1516
<i>TaCAT2</i>	X94352	CCCGGCCATGGATCCCTACA	CITGTCCATCTTACATGTTCCG	1496
<i>TaCAT3</i>	HQ860268	AGGATGGATCCCTGCAAGTT	ATCGATCCATCCATCACATGCT	1502

gDNA 序列。

1.4 TaCAT 氨基酸组成生物学特性分析

利用 ExPASy 在线蛋白质理化性质分析工具对 TaCAT 的

氨基酸组成及其生化特性进行分析和预测,同时利用 PSORT (<http://www.psорт.org/>) 在线分析软件对 TaCAT 进行亚细胞定位预测。采用 Deléage 等^[36]方法对 3 种 TaCAT 进行二级结构

预测 (http://npsa-pbil.ibcp.fr/cgi-bin/seqpred_dpm.pl)。

1.5 同源比对和聚类分析

根据小麦 *TaCAT* 的氨基酸序列, 运用 blastp 在 NCBI 中进行同源比对, 通过在线 clustalW2 对搜索到的相关物种的 *CAT* 进行聚类分析后, 利用 Mega 4.0 软件包中的 Neighbor-Joining 构建系统发育树, 参数设置为 1000 bootstraps。

1.6 小麦 *TaCAT* 同源建模和模型评价

根据 Bordoli 等^[3]介绍的方法对小麦 *TaCATs* 高级结构进行预测。根据 ExPASy 的中 SwissModel 工具得到 *TaCATs* 建模模板, 采用 SwissModel 的 Swiss-PdbViewer 3.7 软件包进行同源模型构建, 同时对 *TaCATs* 保守的功能位点的高级结构进行模拟, 并利用 SwissModel 在线模型评估工具 Procheck 和 Whatcheck 进行评估。

2 结果与分析

2.1 小麦 *TaCAT3* 基因克隆和结构分析

研究表明, 小麦 *TaCAT1* 和 *TaCAT2* 分别与水稻 *CatC*、玉米 *Cat2* 和水稻 *CatB*、玉米 *Cat1* 具有较高的相似性, 推断其可能来源于共同的进化分支^[34]。因此本研究以水稻的 *CatA* 及玉米的 *Cat3* 为种子序列, 利用 NCBI 的 blast 工具, 成功得到 6 条具有重叠的小麦 EST 片段, GenBank 登录号分别为 CK209694、CK209786、CJ711217、DR740318、CJ930266 和 CJ958975。片段组装后得到一条长为 1737bp 的序列 (图 1), ORF Finder 预测结果表明其包含完整的 ORF。据此, 利用 ORF 两侧的 UTR 区域设计引物, 采用常规 PCR 技术在普通小麦材料 CB037 基因组 DNA 和 cDNA 中分别克隆一条特异片段, 大小分别为 2086bp 和 1502bp (图 2)。同时, 根据已发表

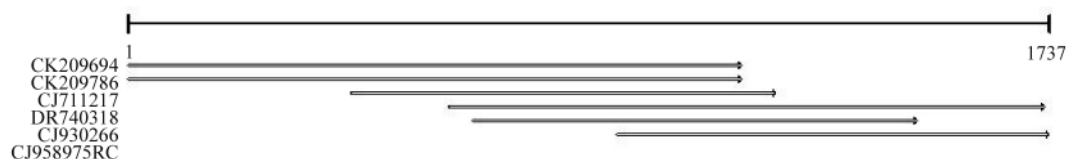


图 1 小麦 *TaCAT3* 序列拼接结果

Fig. 1 Sequence assembly of *TaCAT3* in wheat

的小麦 *TaCAT1* 和 *TaCAT2* 序列设计引物, 成功分离到两者的 gDNA 和 cDNA 序列 (结果未出示)。

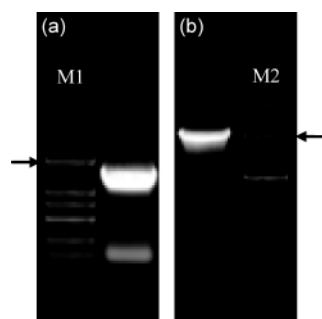


图 2 小麦 *TaCAT3* 序列扩增

Fig. 2 Isolation of *TaCAT3* in wheat

注: (a): cDNA 克隆; (b): gDNA 克隆; M1: DL2000; M2: DL5000; 箭头所示为 2000bp 的片段。

Notes: (a): cDNA; (b): gDNA; M1: DL2000; M2: DL5000; 2000bp fragment was shown by arrows.

对小麦 3 个不同 *TaCAT* 基因的 gDNA 和 cDNA 全长序列进行比对分析, 发现 3 个基因分别具有 6、8、4 个外显子和 5、7、3 个内含子 (图 3), 其中第一个及最后一个外显子在长度上具有极高的保守性。*TaCAT2* 是迄今为止所报道的具有最多外显子数的源自高等植物的 *CAT* 基因^[3]。尽管 3 个 *TaCAT* 基因外显子及内含子数目存在较大差异, 但 cDNA 长度变化较小, 分别为 1476、1476 和 1482bp (按起始密码子到

终止密码子计算, 下同)。3 个 *TaCAT* 基因的 cDNA 全长序列及其预测的编码产物比对结果表明, *TaCAT1* 与 *TaCAT2* 同源性较高, 分别达到 73.1% 和 80.7%, 而二者与 *TaCAT3* 同源性较低, 分别为 80.1% 和 72.8%, 67.3% 和 72.4% (表 2)。在蛋白水平上, *TaCAT3* 与水稻 *CatA*、*CatB*、*CatC* 和玉米 *Cat1*、*Cat2*、*Cat3* 分别达到 89.0%、72.1%、75.4%、72.5%、72.4%、81.1% 相似性, 表明 *TaCAT3* 与水稻 *CatA* 及玉米 *Cat3* 可能来源于一个进化分支, 这与预期的结果相符。测序结果表明, 小麦 3 个 *TaCAT* 基因的 gDNA 长度分别为 3013、3358 和 1986bp (计算同上)。由图 3 可知, 3 个 *TaCAT* 基因的 gDNA 序列长度差异可能是由于内含子在进化上高度活跃所致。此外, *TaCAT1* 及 *TaCAT2* 的终止密码子皆为 TAA, 而 *TaCAT3* 则为 TGA。

表 2 小麦中 3 个 *TaCAT* 基因同源性比较

Table 2 List of the homology for the three members of *TaCAT* in wheat

基因	<i>TaCAT1</i> /%	<i>TaCAT2</i> /%	<i>TaCAT3</i> /%
<i>TaCAT1</i> (%)	100(100)	—	—
<i>TaCAT2</i> (%)	73.1(80.7)	100(100)	—
<i>TaCAT3</i> (%)	80.1(72.8)	67.3(72.4)	100(100)

注: 括号中为氨基酸序列比对结果。

Note: Alignment results of amino acids are shown in brackets.

2.2 小麦 *TaCAT* 编码产物的理化性质和二级结构分析

小麦 3 种 *TaCAT* 基因分别编码 492、492 和 494 个氨基酸, 利用 ExPASy 在线蛋白质理化性质分析工具对其各种理

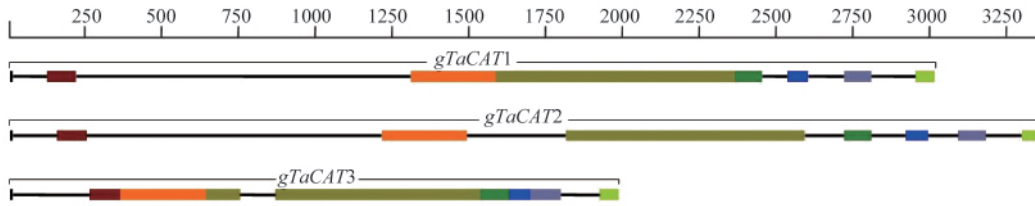


图3 小麦 TaCAT 基因结构

Fig. 3 Exon-intron structures of catalase genes in wheat

注:带颜色的粗线条为外显子,其中相同颜色表示相似的外显子,细线为内含子,标尺为碱基数目。

Notes: Exons are indicated by colored boxes, with the same color representing corresponding regions among genes. The thick lines between exons represent introns. The number of base pairs is corresponding to the uppermost scale.

化性质进行分析,结果如表3。尽管3种TaCAT的相似性存在一定差异(表2),但其氨基酸组成却极为相似,尤以极性与非极性氨基酸含量较高,皆在50%左右,而极小分子的氨基酸、脂肪族氨基酸及带电性R基团氨基酸含量相近,以芳香族氨基酸含量最低,约占16%。

功能位点分析表明,所有的功能活性位点在3种TaCAT中存在组成及定位上的高度保守性,如底物H₂O₂结合位点(Asp350, Arg353)、活性位点(His65, Ser104, Asn138)以及结合亚铁血红素(Val63, Thr105, Phe143, Pro326, Arg344, Phe346, Tyr348)位点等(图4)。此外,这些位点与所有已报道的其他植物CAT相应位点具有高度的一致性(数据未出示),表明这些位点具有进化上的保守性,同时也暗示它们对CAT的功能具有决定作用。利用ExPASy中的亚细胞定位预测软件进行TaCAT定位预测,结果表明,3种TaCAT皆具有过氧化物酶

的定位信号(图4中的SSS/TSG、SRL/NRL和表4中PTS1栏),但其定位在叶绿体中的可能性均很低,仅为0.159、0.076和0.162;此外,尚未发现信号肽序列存在,表明3种TaCAT可能并非分泌蛋白。线粒体定位预测表明,3种TaCAT可能定位在线粒体中,但TaCAT3具有更高的线粒体定位可能性,定位阈值达0.92076359。本研究中的TaCAT3是根据水稻CatA和玉米Cat3同源性克隆得到,而后者已被广泛认为与线粒体定位具有很高的相关性^[9],进一步说明该定位方法的可靠。

基于Deléage等^[10]报道的蛋白质二级结构预测结果,4种基本的二级结构单元(α-螺旋、β-折叠、β-转角、无规则卷曲)在TaCAT中皆存在,其中以无规则卷曲结构居多,分别占59.96%、57.11%和40.08%(表5),TaCAT1与TaCAT2在二级结构组成上极为相似,而与TaCAT3存在一定差异,表现为在

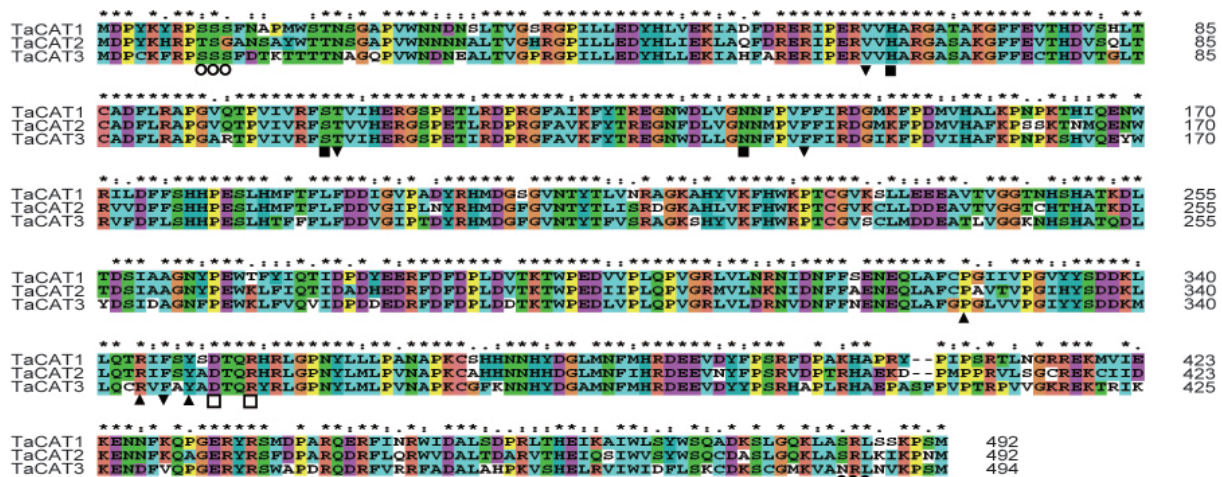


图4 小麦3种TaCAT序列比对结果

Fig. 4 Sequence alignment of three TaCATs in wheat

注:根据clustalW2默认的色彩标注氨基酸残基。星号表示完全一致,冒号表示高度保守,点号表示部分保守;空心、实心圆和方框分别表示过氧化物酶体或乙醛酸循环体定位信号、H₂O₂结合位点和CAT活性位点;近、远端亚铁血红素结合位点分别上、下三角符号表示。

Notes: All the amino acid residues are colored in clustalW2 default settings. Consensus, highly conserved, and semi-conserved residues are shown as asterisks, colons, and dots, respectively. Empty and full rounds and boxes represent targeting sites of peroxisome or glyoxysome, H₂O₂ binding sites and active sites, respectively. Proximal and distal heme-binding legends are shown by upper and lower triangles, respectively.

表 3 3 种小麦 TaCAT 理化特性分析
Table 3 Physiological analysis of three TaCATs in wheat

CAT 亚型	氨基酸类型	残基	数目	比例/%	等电点	分子量/Da
TaCAT1	Tiny	A+C+G+S+T	116	23.577	7.0189	56807.83
	Small	A+B+C+D+G+N+P+S+T+V	246	50.000		
	Aliphatic	A+I+L+V	115	23.374		
	Aromatic	F+H+W+Y	81	16.463		
	Non-polar	A+C+F+G+I+L+M+P+V+W+Y	253	51.423		
	Polar	D+E+H+K+N+Q+R+S+T+Z	239	48.577		
	Charged	B+D+E+H+K+R+Z	143	29.065		
	Basic	H+K+R	80	16.260		
	Acidic	B+D+E+Z	63	12.805		
TaCAT2	Tiny	A+C+G+S+T	119	24.187	7.0686	56479.71
	Small	A+B+C+D+G+N+P+S+T+V	252	51.220		
	Aliphatic	A+I+L+V	121	24.593		
	Aromatic	F+H+W+Y	78	15.854		
	Non-polar	A+C+F+G+I+L+M+P+V+W+Y	254	51.626		
	Polar	D+E+H+K+N+Q+R+S+T+Z	238	48.374		
	Charged	B+D+E+H+K+R+Z	144	29.268		
	Basic	H+K+R	82	16.667		
	Acidic	B+D+E+Z	62	12.602		
TaCAT3	Tiny	A+C+G+S+T	115	23.279	7.0853	56883.14
	Small	A+B+C+D+G+N+P+S+T+V	257	52.024		
	Aliphatic	A+I+L+V	116	23.482		
	Aromatic	F+H+W+Y	83	16.802		
	Non-polar	A+C+F+G+I+L+M+P+V+W+Y	264	53.441		
	Polar	D+E+H+K+N+Q+R+S+T+Z	230	46.559		
	Charged	B+D+E+H+K+R+Z	151	30.567		
	Basic	H+K+R	84	17.004		
	Acidic	B+D+E+Z	67	13.563		

注: A, Ala; B, Asx; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; U, Sec (selenocysteine); V, Val; W, Trp; X, Xaa; Y, Tyr; Z, Glx.

Notes: A, Ala; B, Asx; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; U, Sec (selenocysteine); V, Val; W, Trp; X, Xaa; Y, Tyr; Z, Glx.

表 4 小麦中 3 种 TaCAT 定位信号预测
Table 4 Prediction on subcellular location of three TaCATs in wheat

CAT 亚型	大小/aa	cTP ^[37]	SLP-Local ^[38]	SignalP ^[39]	PTS1 ^[40]
TaCAT1	492	0.159	0.45342794	—	7.754
TaCAT2	492	0.076	0.62997647	—	8.683
TaCAT3	494	0.162	0.92076359	—	9.206

注: —, 未发现。

Note: —, not found.

TaCAT3 中 α -螺旋、 β -折叠和 β -转角的比例有所上升, 分别达 31.78%、19.64% 和 8.50%, 而无规则卷曲结构降至 40.08%, 其二级结构如图 5 所示。

表 5 小麦 TaCAT 二级结构统计表

Table 5 Statistical summary of secondary structure of TaCAT in wheat

CAT 亚型	α -螺旋/%	β -折叠/%	β -转角/%	无规则卷曲/%
TaCAT1	24.59(121)	11.59(57)	3.86(19)	59.96(295)
TaCAT2	27.85(137)	11.79(58)	3.25(16)	57.11(281)
TaCAT3	31.78(157)	19.64(97)	8.50(42)	40.08(198)

注: 括号中为参与二级结构形成的氨基酸残基数。

Note: Amino acid residues composing the secondary structures are shown in brackets.

2.3 小麦 TaCAT 高级结构同源建模

蛋白质高级结构与其功能具有极大相关性, 因此, 正确

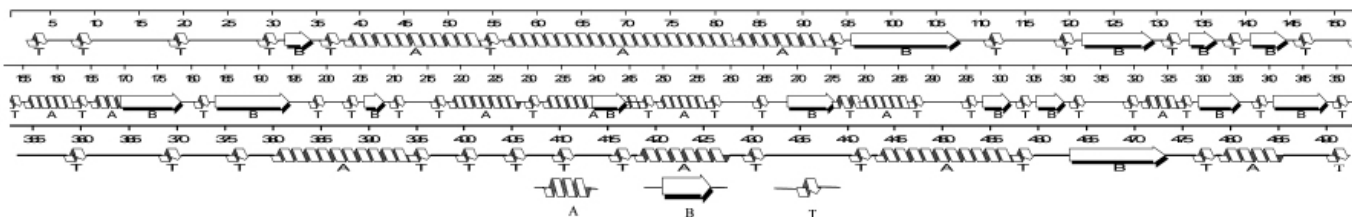


图5 小麦 TaCAT3 二级结构示意图

Fig. 5 Schematic of TaCAT3 secondary structure of wheat

注:A, α -螺旋;B, β -折叠;T, β -转角。

Notes: A, α -helix; B, β -sheet; T, β -turn.

预测蛋白质高级结构对功能分析具有重要意义,尤其对于高级结构未知的蛋白^[3,41]。小麦 TaCAT 蛋白质的高级结构尚未测定。本研究利用 ExPASy 的 spdbv 离线软件包 (Swiss Pdbviewer 3.7) 对小麦 TaCAT 进行同源建模。通过在线 Swiss Model 寻找一致性最高且结构已知的 CAT 作为同源建模的模板,结果见表 6。通过 spdbv 软件的 Magic Fit 工具进行 TaCAT 模型构建 (图 6, 此处仅显示 TaCAT3 的同源建模结果)。采用 Procheck 和 Whatcheck 对建模结果进行分析,根据计算得出 Ramachandran plot 对所建模型进行评价 (不包括甘氨酸及脯氨酸^[42])。同源建模及评估结果表明,预测模型包含 13 个 α -螺旋、11 个 β -折叠和大量无规则卷曲结构组成,其中有 97.4% 的氨基酸残基位于由 ψ 和 ϕ 二面角构成的 Ramachandran plots 中黄色区域 (即许可的 ψ 和 ϕ 二面角区,图 7), Procheck 评估结果的 G-factor 为 -1.09, 表明本模型预

表 6 小麦 TaCAT 同源建模模板

Table 6 Templates for TaCAT homology modeling of wheat

CAT 亚型	模型模板来源	PDB 登录号	一致性	E 值
TaCAT1	<i>Exiguobacterium oxidotolerans</i>	2j2mA	50%	e-132
TaCAT2	<i>Pseudomonas syringae</i>	1m7sA	49%	e-129
TaCAT3	<i>Pseudomonas syringae</i>	1m7sA	47%	e-125

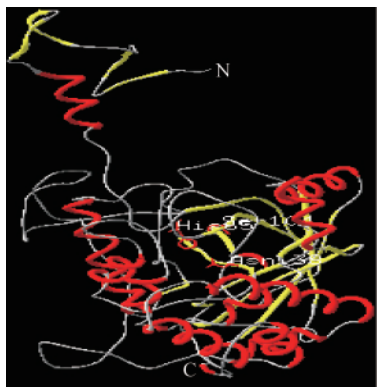


图6 小麦 TaCAT3 的同源建模

Fig. 6 Homology modeling of TaCAT3

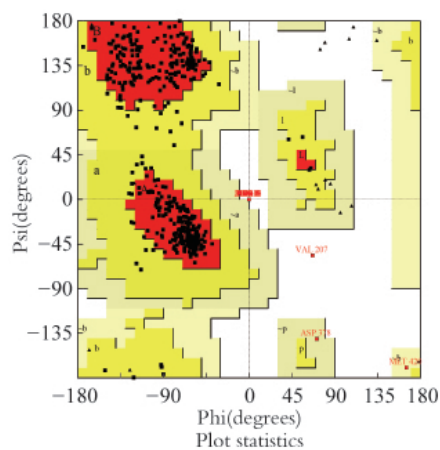


图7 小麦 TaCAT3 三维模型的 Ramachandran 图

Fig. 7 Ramachandran plots of TaCAT3 3-D model

注:3 个活性位点残基以线框模式显示,模型以 3D 形式表示,红、黄、灰色分别代表 α -螺旋 β -折叠、无规则卷曲。

Notes: Active sites are shown in wire-frame model. The advanced structure of TaCAT3 is in 3-D model. Red, yellow, and grey represent α -helix, β -sheet, and random coil, respectively.

测的可信度很高,由 *Pseudomonas syringae* 来源的 CAT 晶体结构能很好地反映出小麦 TaCAT3 的高级结构特征。同时对两种来源的 CAT 活性位点残基的空间构象进行重叠图构建,发现 TaCAT3 活性位点的 3 个残基 His65、Ser104 和 Asn138 能够与 1m7sA 的活性位点 His56、Ser95 和 Asn129 完全重合,表明活性位点、毗邻残基的种类和空间构象对 CAT 功能发挥具有重要作用,也进一步说明该模型预测可行 (图 8)。

2.4 小麦 TaCAT 基因进化树构建

为进一步研究小麦 TaCAT 与其他种 CAT 间的进化关系,利用 MEGA 4.0 软件对 blastp 得到的几种高等植物的 CAT 基因进行系统发育树构建。图 9 显示 4 个明显的分支,其中 TaCAT1 与水稻 CatC、玉米 Cat2 和高羊茅草 (*Festuca arundinacea* L.)Cat 形成一个分支, TaCAT2 与水稻 CatB、玉米 Cat1、大麦 Cat1、高粱 (*Sorghum bicolor* L.)Cat 形成另一个分支, TaCAT3 与水稻 CatA、玉米 Cat3、大麦 Cat2 及黑麦 (*Secale cereale* L.)Cat 形成第 3 分支,而其他双子叶植物来源的 CAT

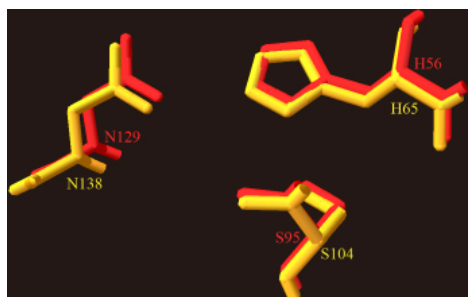


图 8 过氧化氢酶活性位点重叠图

Fig. 8 Iterative map of active sites of TaCAT3

注: 红色和黄色分别为 *Pseudomonas syringae* CAT 和小麦 TaCAT3 的活性位点。残基名称和位置已标注。

Notes: Red color and yellow color denote the active sites of catalase from *Pseudomonas syringae* and wheat with the names and sites beside the structures, respectively.

构成第 4 个分支,由此可见,双子叶植物体较单子叶植物可能具有更多 CAT 亚型。然而,同种来源的 CAT 未必包括在同一个亚群中,如皱叶烟草(*Nicotiana plumbaginifolia* L.)的 Cat3 与 Cat1 并不在同一亚群,而与陆地棉(*Gossypium hirsutum* L.) Cat1 形成一群,尽管置信度仅为 42%,同样的现象也出现在陆地棉 Cat2(图 9),因此推测,在皱叶烟草的 Cat1 与 Cat3 之间以及陆地棉 Cat1 与 Cat2 之间的 CAT 功能差异可能较大。

3 讨论

小麦组织培养再生基因分离及鉴定是从分子角度解析小麦组织培养再生机理的关键步骤,有利于采用生物技术手段提高小麦组织培养的再生能力。Nishimura 等^[20]认为植物的再生能力差异与遗传背景有很大关系,故本研究选择再生性能较好的普通小麦材料 CB037 进行 *CAT* 基因克隆以及生化和分子特性预测和分析。

Iwamoto 等^[9]对水稻 *CAT* 的进化研究认为,不同的 *CAT* 同工酶亚型出现可能是在单子叶植物从双子叶植物进化分离出来后由于内含子的丢失或插入所致。本研究中所克隆的 3 种 *TaCAT* 也存在这种差异,表现为内含子数目不一致(图 3),推测小麦 *CAT* 与水稻 *CAT* 可能存在同样的内含子进化机制。von Ossowski 等^[43]研究发现,在 *CAT* 进化上单子叶植物与双子叶植物不能形成 2 个独立分支,认为可能由于植物 *CAT* 进化缓慢所致。然而,本研究发现 *CAT* 基因能很好地形成单、双子叶植物 2 个独立的进化分支,其置信度达到 84%(图 9),其机制有待进一步研究。

同源建模技术是用于预测蛋白高级结构的一项重要技术,通过对未知蛋白结构的同源模拟便于利用突变模拟途径研究蛋白质功能^[33],揭示蛋白质间或酶与底物间的互作机制。同时结合多种模型评估手段,为后续蛋白质功能实证研究奠定基础。

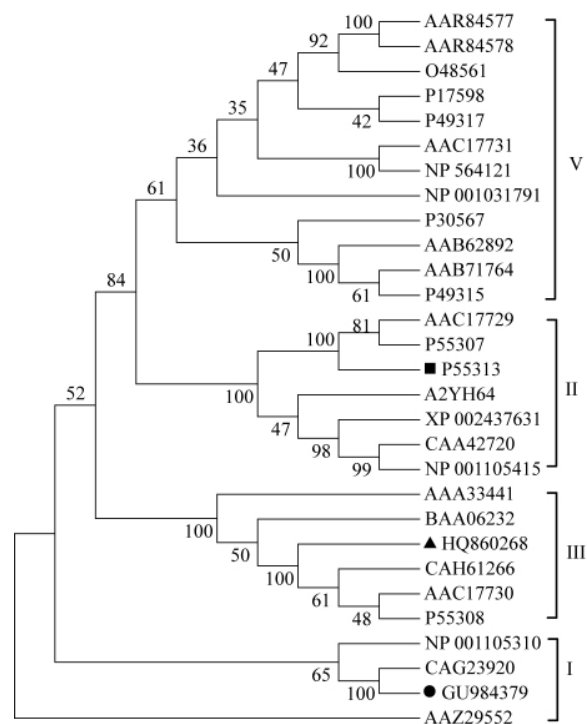


图 9 植物 CAT 系统发育树

Fig. 9 Phylogenetic tree of plant catalases

注:根据小麦 *CAT* 家族成员的发现顺序进行 3 个分支命名;圆、方框及上三角形分别标示 TaCAT1、TaCAT2 和 TaCAT3;节点处数值表示 bootstrap 分析的置信度。

Notes: The three groups are denoted based on the sequenced isolation of TaCAT. The circle, box, and upper triangle represent TaCAT1, TaCAT2, and TaCAT3, respectively. The numbers at the nodes are the indications of the level of confidence for the major branches determined by bootstrap analysis.

GU984379: TaCAT1, *T. aestivum*; P55313: TaCAT2, *T. aestivum*; HQ860268: TaCAT3, *T. aestivum*; BAA06232: CatA, *O. sativa*; A2YH64: CatB, *O. sativa*; AAZ29552: CatC, *O. sativa*; NP_001105415: Cat1, *Z. mays*; NP_001105310: Cat2, *Z. mays*; AAA33441: Cat3, *Z. mays*; P55307: Cat1, *H. vulgare*; P55308: Cat2, *H. vulgare*; AAB62892: Cat1, *Nicotiana glutinosa*; AAB71764: Cat1, *N. tabacum*; P49315: Cat1, *N. plumbaginifolia*; P49317: Cat3, *N. plumbaginifolia*; AAC17729: Cat1, *H. vulgare*; AAC17730: Cat2, *H. vulgare*; AAC17731: Cat1, *A. thaliana*; NP_001031791: Cat2, *A. thaliana*; AAR84578: Cat1a, *Lotus japonicus*; AAR84577: Cat1b, *L. japonicus*; CAA42720: Cat1, *Z. mays*; CAG23920: Cat, *F. arundinacea*; CAH61266: Cat, *S. cereale*; NP_564121: Cat1, *A. thaliana*; O48561: Cat4, *G. max*; P17598: Cat1, *G. hirsutum*; P30567: Cat2, *G. hirsutum*; XP_002437631: Cat, *S. bicolor*

4 结论

根据已报道的水稻及玉米 *CAT* 序列,基于 *in silico* 技术进行小麦 EST 序列比对、拼接和克隆,成功分离到一个新的小麦 *CAT* 同工酶基因,命名为 *TaCAT3*。借助生物信息技术

对已有的小麦 3 种 *TaCAT* 基因及其编码产物进行理化及生物学特性分析, 结果表明, *TaCAT3* 与 *TaCAT1* 和 *TaCAT2* 不仅在基因组结构上具有较大差异, 在亚细胞定位方面也存在差异, 但其编码产物的氨基酸组成、二级结构及高级结构具有较高的相似性。聚类分析表明, 3 种 *TaCAT* 同工酶亚型能形成 3 个独立的进化分支。

参考文献 (References)

- [1] Willekens H, Chammongpol S, Davey M, *et al.* Catalase is a sink for H₂O₂ and is indispensable for stress defence in C3 plants [J]. *The EMBO Journal*, 1997, 16(16): 4806–4816.
- [2] Luna C M, Pastori G M, Driscoll S, *et al.* Drought controls on H₂O₂ accumulation, catalase (CAT) activity and *CAT* gene expression in wheat [J]. *Journal of Experimental Botany*, 2005, 56(411): 417–423.
- [3] Iwamoto M, Maekawa M, Saito A, *et al.* Evolutionary relationship of plant catalase genes inferred from exon–intron structures: Isozyme divergence after the separation of monocots and dicots [J]. *Theoretical and Applied Genetics*, 1998, 97(1): 9–19.
- [4] Scandalios J G, Guan L, Polidoros A N. Catalases in plants: gene structure, properties, regulation, and expression [C]/Scandalios J G. Oxidative stress and the molecular biology of antioxidant defenses[M]. New York: Cold Spring Harbor Laboratory Press, 1997, 343–406.
- [5] Higo K, Higo H. Cloning and characterization of the rice *CatA* catalase gene, a homologue of the maize *Cat3* gene [J]. *Plant Molecular Biology*, 1996, 30(3): 505–521.
- [6] Iwamoto M, Higo H, Higo K. Strong expression of the rice catalase gene *CatB* promoter in protoplasts and roots of both a monocot and dicots[J]. *Plant Physiology and Biochemistry*, 2004, 42(3): 241–249.
- [7] Williamson J D, Scandalios J G. Differential response of maize catalases to abscisic acid: Vp1 transcriptional activator is not required for abscisic acid–regulated *Cat1* expression [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1992, 89(18): 8842–8846.
- [8] Bethards L A, Skadsen R W, Scandalios J G. Isolation and characterization of a cDNA clone for the *Cat2* gene in maize and its homology with other catalases[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1987, 84(19): 6830–6834.
- [9] Scandalios J G, Tong W F, Roupakias D G. *Cat3*, a third gene locus coding for a tissue–specific catalase in maize: Genetics, intracellular location, and some biochemical properties [J]. *Molecular and General Genetics*, 1980, 179(1): 33–41.
- [10] Scandalios J G. Subunit dissociation and recombination of catalase isozymes [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1965, 53(5): 1035–1040.
- [11] Iwamoto M, Higo H, Higo K. Differential diurnal expression of rice catalase genes: The 5′–flanking region of *CatA* is not sufficient for circadian control[J]. *Plant Science*, 2000, 151(1): 39–46.
- [12] Suzuki M, Ario T, Hattori T, *et al.* Isolation and characterization of two tightly linked catalase genes from castor bean that are differentially regulated[J]. *Plant Molecular Biology*, 1994, 25(3): 507–516.
- [13] Niebel A, Heungens K, Barthels N, *et al.* Characterization of a pathogen–induced potato catalase and its systemic expression upon nematode and bacterial infection [J]. *Molecular Plant–Microbe Interactions*, 1995, 8(3): 371–378.
- [14] Skadsen R W, Schulze–Lefert P, Herbst J M. Molecular cloning, characterization and expression analysis of two catalase isozyme genes in barley[J]. *Plant Molecular Biology*, 1995, 29(5): 1005–1014.
- [15] Zhong H, McClung C. The circadian clock gates expression of two *Arabidopsis* catalase genes to distinct and opposite circadian phases[J]. *Molecular and General Genetics*, 1996, 251(2): 196–203.
- [16] Saruyama H, Matsumura T. Cloning and characterization of a cDNA encoding catalase in wheat[J]. *DNA Sequence*, 1999, 10(1): 31–35.
- [17] Vranova E, Inze D, Van Breusegem F. Signal transduction during oxidative stress [J]. *Journal of Experimental Botany*, 2002, 53(372): 1227–1236.
- [18] Neill S J, Desikan R, Clarke A, *et al.* Hydrogen peroxide and nitric oxide as signalling molecules in plants [J]. *Journal of Experimental Botany*, 2002, 53(372): 1237–1247.
- [19] Papadakis A, Roubelakis–Angelakis K. Oxidative stress could be responsible for the recalcitrance of plant protoplasts[J]. *Plant Physiology and Biochemistry*, 2002, 40(6–8): 549–559.
- [20] Szechyńska–Hebda M, Skrzypek E, Dąbrowska G, *et al.* The role of oxidative stress induced by growth regulators in the regeneration process of wheat[J]. *Acta Physiologiae Plantarum*, 2007, 29(4): 327–337.
- [21] Rajeswari V, Paliwal K. Peroxidase and catalase changes during *in vitro* adventitious shoot organogenesis from hypocotyls of *Albizia odoratissima* L. f. (Benth)[J]. *Acta Physiologiae Plantarum*, 2008, 30(6): 825–832.
- [22] Lucau–Danila A, Laborde L, Legrand S, *et al.* Identification of novel genes potentially involved in somatic embryogenesis in chicory (*Cichorium intybus* L.)(J]. *BMC Plant Biology*, 2010, 10(1): 122–136.
- [23] Zubko E, Adams C J, MacHáčeková I, *et al.* Activation tagging identifies a gene from *Petunia hybrida* responsible for the production of active cytokinins in plants[J]. *The Plant Journal*, 2002, 29(6): 797–808.
- [24] Low R K, Prakash A P, Swarup S, *et al.* A differentially expressed *bZIP* gene is associated with adventitious shoot regeneration in leaf cultures of *Paulownia kawakamii*[J]. *Plant Cell Reports*, 2001, 20(8): 696–700.
- [25] Koornneef M, Beda J, Hanhart C, *et al.* Characterization and mapping of a gene controlling shoot regeneration in tomato [J]. *The Plant Journal*, 1993, 3(1): 131–141.
- [26] Schmidt E D, Guzzo F, Toonen M A, *et al.* A leucine–rich repeat containing receptor–like kinase marks somatic plant cells competent to form embryos[J]. *Development*, 1997, 124(10): 2049–2062.
- [27] Banno H, Ikeda Y, Niu Q W, *et al.* Overexpression of *Arabidopsis* *ESR1* induces initiation of shoot regeneration[J]. *The Plant Cell*, 2001, 13(12): 2609–2618.
- [28] Nishimura A, Ashikari M, Lin S, *et al.* Isolation of a rice regeneration quantitative trait loci gene and its application to transformation systems [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2005, 102(33): 11940–11944.
- [29] Zhang S, Wong L, Meng L, *et al.* Similarity of expression patterns of *knotted1* and *ZmLEC1* during somatic and zygotic embryogenesis in maize (*Zea mays* L.)(J]. *Planta*, 2002, 215(2): 191–194.
- [30] Komatsuda T, Annaka T, Oka S. Genetic mapping of a quantitative trait locus (QTL) that enhances the shoot differentiation rate in *Hordeum vulgare* L.)(J]. *Theoretical and Applied Genetics*, 1993, 86(6): 713–720.
- [31] Singla B, Tyagi A K, Khurana J P, *et al.* Analysis of expression profile of selected genes expressed during auxin–induced somatic embryogenesis in leaf base system of wheat (*Triticum aestivum*) and their possible interactions [J]. *Plant Molecular Biology*, 2007, 65(5): 677–692.
- [32] 陈军营, 马平安, 赵一丹, 等. 小麦成熟胚脱分化过程中生长素相关

- 基因的表达分析[J]. 作物学报, 2009, 35(10): 1798-1805.
Chen Junying, Ma Pingan, Zhao Yidan, *et al.* *Acta Agronomica Sinica*, 2009, 35(10): 1798-1805.
- [33] Bordoli L, Kiefer F, Arnold K, *et al.* Protein structure homology modeling using SWISS-MODEL workspace[J]. *Nature Protocols*, 2008, 4(1): 1-13.
- [34] Saruyama H, Matsumura T. Cloning and characterization of a cDNA encoding catalase in wheat[J]. *Mitochondrial DNA*, 1999, 10(1): 31-35.
- [35] Scandalios J G. Response of plant antioxidant defense genes to environmental stress[M]. *Advances in Genetics*, 1990, 28: 1-41.
- [36] Deléage G, Roux B. An algorithm for protein secondary structure prediction based on class prediction[J]. *Protein Engineering*, 1987, 1(4): 289-294.
- [37] Emanuelsson O, Nielsen H, Brunak S, *et al.* Predicting subcellular localization of proteins based on their N-terminal amino acid sequence [J]. *Journal of Molecular Biology*, 2000, 300(4): 1005-1016.
- [38] Matsuda S, Vert J P, Saigo H, *et al.* A novel representation of protein sequences for prediction of subcellular location using support vector machines[J]. *Protein Science*, 2005, 14(11): 2804-2813.
- [39] Bendtsen J D, Nielsen H, von Heijne G, *et al.* Improved prediction of signal peptides: SignalP 3.0[J]. *Journal of Molecular Biology*, 2004, 340(4): 783-795.
- [40] Neuberger G, Maurer-Stroh S, Eisenhaber B, *et al.* Prediction of peroxisomal targeting signal 1 containing proteins from amino acid sequence[J]. *Journal of Molecular Biology*, 2003, 328(3): 581-92.
- [41] Baker D, Sali A. Protein structure prediction and structural genomics[J]. *Science*, 2001, 294(5540): 93-96.
- [42] Sekhar P N, Kishor P B, Reddy L A, *et al.* *In silico* modeling and hydrogen peroxide binding study of rice catalase [J]. *In Silico Biology*, 2006, 6(5): 435-447.
- [43] von Ossowski I, Hausner G, Loewen P C. Molecular evolutionary analysis based on the amino acid sequence of catalase [J]. *Journal of Molecular Evolution*, 1993, 37(1): 71-76.

(责任编辑 吴晓丽)

第五届国际生物数学大会

时间: 2011.06.03—06.05

地点: 南京

主办: 中国数学会

主题: 数学生态学和生物数学

电话: 025-58731160

电子信箱: icmb2011@163.com