

陈子强,李 刚,颜静宛,等. 玉米 CBS 基因家族的鉴定和特征分析[J]. 江苏农业科学,2024,52(4):41-49.
doi:10.15889/j.issn.1002-1302.2024.04.006

玉米 CBS 基因家族的鉴定和特征分析

陈子强,李 刚,颜静宛,郭新睿,胡昌泉,王 月,王 锋,田大刚

(福建省农业科学院生物技术研究所/福建省农业遗传工程重点实验室,福建福州 350003)

摘要:为研究玉米 CBS(*ZmCBS*) 基因家族的特征,探讨其功能,对 *ZmCBS* 基因家族进行鉴定,分析其染色体定位、基因结构、保守结构域及进化关系,并对 *ZmCBS* 家族基因在玉米组织中的表达情况及其在非生物胁迫下的表达变化进行分析。基于玉米基因组数据库,鉴定到 37 个 CBS 基因家族成员,它们不均匀地分布在玉米的 8 条染色体上。该家族蛋白除保守的 CBS 结构域外,多数成员还包含其他结构域,包括电压门控的氯化物通道蛋白(voltage gated CLC)、Phox/Bem1(PBI)结构域、五肽重复序列(pentapeptide repeat,PPR)和 E-SET 结构域等。按照基因结构和进化关系可分为 6 个亚类,且同一进化分支上的基因具有相似的外显子结构和 CDS 长度,但内含子长度差异很大。启动子分析结果表明,*ZmCBS* 基因启动子上含有响应激素和非生物胁迫的顺式作用元件。转录组数据的基因表达谱分析显示,*ZmCBS* 基因家族部分成员在叶、花粉或胚中优势表达。*ZmCBS3* 和 *ZmCBS14* 受冷胁迫诱导表达,*ZmCBS14* 和 *ZmCBS22* 受干旱胁迫诱导表达,暗示这些基因在玉米响应外界胁迫中起重要作用。分析结果将为进一步研究 *ZmCBS* 家族基因奠定基础。

关键词:玉米;CBS 基因家族;生物信息学;生物胁迫;非生物胁迫;特征分析

中图分类号:S513.01 **文献标志码:**A **文章编号:**1002-1302(2024)04-0041-09

胱硫醚 β 合成酶(cystathionine β -synthase, CBS)结构域是广泛存在于所有物种中的保守结构域。其由 β - α - β - β - α 二级结构组成,并进一步折叠成球状三级结构^[1]。该结构域最初是在古细菌蛋白中被发现的,随后在人类的胱硫醚 β 合成酶中被发现^[2-3]。对 CBS 中的 CBS 结构域进行点突变会引起人类几种遗传性疾病,证实 CBS 结构域的重要性^[3]。研究表明,CBS 结构域可结合腺苷化合物,如单磷酸腺苷(AMP)、二磷酸腺苷(ADP)、三磷酸腺苷(ATP)和 S-腺苷蛋氨酸,并在细胞能量感知中发挥功能^[4-6]。除 CBS 外,CBS 结构域还存在于其他蛋白中,通过与 CBS 结构域蛋白(CBS domain-containing proteins, CDCPs)的其他功能区域相结合来调节生物过程。如肌苷-5'-单磷酸脱氢酶

(inosine-5'-monophosphate dehydrogenase, IMPDH)结构域,AMP 激活蛋白激酶(AMP-activated protein kinase, AMPK)结构域和电压门控氯通道(voltage-gated chloride channels, CLC)结构域等^[7-9]。

过去对 CBS 结构域在植物中的研究也有一些报道。前人通过全基因组分析已经从拟南芥、水稻、小麦、大豆中鉴定了 34、59、66、71 个 CDCPs^[1,10]。其中部分成员被证实一些植物生长发育以及响应生物和非生物胁迫过程中发挥着重要功能。拟南芥 *AtCBSX1* 和 *AtCBSX2* 在花(尤其是花药)中高表达,其在叶绿体中通过激活硫氧还蛋白-Trx 系统中的硫氧还蛋白(thioredoxins, Trx)来帮助调节细胞中的 H_2O_2 水平,进而调控花药的开裂^[5,11]。*AtCBSX3* 定位于线粒体,并调控线粒体内 NADP-Trx 系统中的 Trx 成员^[5]。敲除 *AtCBSX3* 会引起花药内壁次生壁增厚,造成花药不裂,表现出雄性不育^[12]。Mou 等发现,水稻接种稻瘟病菌、外源施用水杨酸(SA)或茉莉酸甲酯(MeJA)均可导致 *OsCBSX3* 转录本上调,过表达 *OsCBSX3* 赋予水稻对稻瘟病菌的抗性^[13]。*OsCBSX4* 在模式植物烟草中过表达时表现出抗氧化、抗盐和抗重金属的特性^[14]。大豆 *GmCBS14* 在结节感染区和维管束中特

收稿日期:2023-04-25

基金项目:福建省科技计划公益类专项(编号:2022R1027002、2021R1027001);福建省农业高质量发展超越“5511”协同创新工程项目(编号:XTXGC2021002)。

作者简介:陈子强(1988—),男,福建泰宁人,博士,助理研究员,从事玉米分子育种研究。E-mail:402253716@qq.com。

通信作者:田大刚,博士,副研究员,从事玉米分子育种研究, E-mail:tdg@fjage.org;王 锋,博士,研究员,主要从事水稻生物技术研究, E-mail:wf@fjage.org。

异性表达,并调节大豆结节生物固氮能力^[6]。另外研究发现,一些含有其他功能域的 CDCPs 在非生物(盐、干旱、高温或低氮等)胁迫过程中同样发挥了重要作用,如 *OsCBSCBSPB4*、*GmCBSDUF3* 和 *GmCBS21*^[15-17]。

目前,玉米是全球第一大谷物,也是我国种植面积位居第一的粮食作物,对国家粮食安全尤其是饲料粮安全至关重要^[18]。然而,随着环境条件的日益恶化,其受到的非生物胁迫更加频繁、严重。因此,对玉米非生物胁迫相关基因功能以及胁迫应答机制的研究具有重要意义。以往的研究表明,CDCPs 在植物生长发育以及响应生物和非生物胁迫过程中发挥着重要作用^[5,11-17]。但目前尚未对玉米 CBS(*ZmCBS*)基因家族成员进行鉴定,其功能研究尚未见报道。本研究采用生物信息学方法对 *ZmCBS* 基因家族进行鉴定,分析其染色体定位、基因结构、保守结构域及进化关系,并对 *ZmCBS* 家族基因在玉米组织中的表达情况及其在非生物胁迫下的表达变化进行分析,了解 *ZmCBS* 家族基因的特征,以期为研究 *ZmCBS* 基因的功能奠定基础。

1 材料与方法

1.1 试验时间与地点

本试验于 2023 年 2 月在福建省农业科学院遗传工程重点实验室内完成。

1.2 *ZmCBS* 基因家族成员的鉴定及理化性质分析

通过 NCBI 数据库(<https://www.ncbi.nlm.nih.gov/protein>)获取 *ZmCBS* 蛋白序列。使用 Cluster X (v2.0)生成多序列比对,删除冗余的序列。利用 NCBI 中的 Batch CD - Search (<https://www.ncbi.nlm.nih.gov/Structure/Bwrpsb/bwrpsb.cgi>)验证 CBS 结构域是否存在。最后,将筛选出的基因按照基因在染色体上的位置进行命名。将所得 *ZmCBS* 基因家族蛋白序列提交在线网站 ProtParam (<https://web.expasy.org/protparam/>)上,分析各 *ZmCBS* 蛋白的氨基酸数量、分子量、等电点等理化性质。

1.3 *ZmCBS* 基因家族成员染色体定位

在 NCBI 数据库中获取 *ZmCBS* 基因家族各成员在染色体上的位置信息及玉米各染色体长度信息,利用 TBtools 工具绘制 *ZmCBS* 家族基因的染色体定位图。

1.4 *ZmCBS* 基因结构、蛋白结构域及系统发育分析

从 NCBI 数据库下载 *ZmCBS* 基因家族成员基

因组信息文件(GFF3 格式)和编码序列,并分析 CBS 基因的外显子/内含子结构,通过 TBtools 处理得到可视化图形。

将 *ZmCBS* 家族所有成员的蛋白序列导入 NCBI 中的 Batch CD - Search 网站,分析 *ZmCBS* 家族蛋白保守基序,并通过 TBtools 处理得到可视化图形。

采用 MEGA - X 软件的邻接法(neighbor - joining, NJ)构建系统发育树,bootstrap 设置为 1 000^[19]。采用 iTOL v6 在线网站(<http://itol.embl.de/>)对系统发育树进行美化。

1.5 *ZmCBS* 基因家族成员启动子顺式元件分析

下载 *ZmCBS* 家族基因上游 2 000 bp 的启动子序列,通过 PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>)分析启动子的顺式作用元件^[20]。使用 TBtools 对顺式作用元件在基因上的位置进行可视化处理。

1.6 *ZmCBS* 基因家族成员表达模式分析

从 Maize GDB (<https://qteller.maizegdb.org/>)网站下载玉米 *ZmCBS* 基因家族成员在初生根、成熟叶、6~7 节间、花丝、成熟花粉、胚(20 DAP)和胚乳(12 DAP)等组织中的转录组数据,以及玉米正常情况下和受冷、热、盐、干旱胁迫的转录组数据,使用 TBtools 工具绘制转录组数据热图。

2 结果与分析

2.1 *ZmCBS* 基因家族的鉴定及理化性质分析

从 NCBI 上得到的 *ZmCBS* 基因家族蛋白的氨基酸序列,去除同一基因的冗余序列和不同转录本,随后利用 NCBI 中的 Batch CD - Search 工具检验 CBS 候选基因,删除不具有 CBS 保守结构域的序列,最终得到 37 个 *ZmCBS* 基因。根据它们在染色体上的排列顺序,将其命名为 *ZmCBS1* ~ *ZmCBS37*。*ZmCBS* 家族蛋白的氨基酸数量为 114 ~ 893 个,分子量为 12.92 ~ 94.86 ku。18 个 *ZmCBS* 蛋白的等电点偏酸性(4.64 ~ 7.10),19 个 *ZmCBS* 蛋白的等电点偏碱性(7.11 ~ 10.52)(表 1)。

2.2 *ZmCBS* 基因家族染色体定位

为了进一步了解 *ZmCBS* 基因家族各成员在染色体上的具体分布,根据 *ZmCBS* 在染色体的位置信息(表 1),利用 TBtools 处理得到可视化图形。图 1 显示,除了第 7、第 9 条染色体无基因分布外,*ZmCBS* 在其余染色体上均有分布,其中第 1、第 2、第 3 条染色体上各有 7 个基因,第 4、第 8 条染色体上各有 4 个

表 1 玉米 ZmCBS 基因家族成员信息

基因名称	基因座	染色体位置 (B73_refGen_v4)	转录 ID	转录本长 度 (bp)	CDS 长度 (bp)	外显子 数量 (个)	氨基酸数 量 (个)	蛋白大小 (ku)	理论等电 点 (pI)	描述
ZmCBS1	LOC100281677	Chr1:150305354..150310132	NM_0011154597.2	2 112	1 638	14	545	58.54	6.59	CBS 结构域蛋白 CBSCBSPB2-like
ZmCBS2	LOC100279218	Chr1:187960341..187966288	NM_001353877.1	2 718	2 238	9	745	79.53	6.45	氯离子通道蛋白 CLC-f
ZmCBS3	LOC100280253	Chr1:232438237..232440426	NM_0011153181.1	1 469	1 128	2	375	39.04	5.47	CBS 结构域蛋白 CBSX5-like
ZmCBS4	LOC109943752	Chr1:235203149..235206458	XM_023301409.2	747	354	4	117	13.50	9.39	蔗糖非发酵蛋白 SNF4-like
ZmCBS5	LOC103643844	Chr1:269855616..269882712	XM_008667011.3	2 946	2 346	23	781	85.81	8.05	氯离子通道蛋白 CLC-d
ZmCBS6	LOC100283846	Chr1:280473742..280478439	NM_0011156744.1	1 119	618	6	205	22.50	9.84	CBS 结构域蛋白 CBSX3
ZmCBS7	LOC541867	Chr1:306463718-306471893	NM_0011115535.2	2 288	1 494	13	497	55.04	6.55	AKINβγ-1 蛋白激酶
ZmCBS8	LOC100281050	Chr2:105994789-106003626	NM_0011153969.3	998	684	7	227	24.47	9.84	胱硫醚β合成酶
ZmCBS9	LOC100192993	Chr2:1113134347..1113136421	XM_020547384.2	1 400	570	5	189	19.90	10.52	未鉴定蛋白
ZmCBS10	LOC103647023	Chr2:140373681-140384539	XM_008671598.2	2 257	1 734	13	577	62.27	6.46	CBS 结构域蛋白 CBSCBSPB3
ZmCBS11	LOC100278924	Chr2:187908409..187925116	NM_001279411.1	2 029	1 863	13	620	68.92	8.98	未鉴定蛋白
ZmCBS12	LOC100191503	Chr2:6108832..6115484	XM_008668816.4	3 309	2 439	7	812	88.31	8.28	氯离子通道蛋白 CLC-g
ZmCBS13	LOC100284094	Chr2:64189878..64204377	XM_020547778.2	2 522	1 353	13	450	50.07	5.59	蔗糖非发酵蛋白 SNF4
ZmCBS14	LOC100280507	Chr2:75432208..75437323	NM_001291619.1	1 731	1 272	6	423	46.57	5.21	SNF1 相关蛋白激酶调节亚单位 γ-1 (CKSy-1)
ZmCBS15	LOC100191828	Chr3:162745581..162749975	NM_0011137252.1	1 881	1 224	2	407	43.46	7.47	CBS 结构域蛋白 CBSX6
ZmCBS16	LOC100502544	Chr3:177560355..177565236	XM_020550405.2	3 914	2 367	6	788	84.36	7.79	氯离子通道蛋白 CLC-e
ZmCBS17	LOC103651228	Chr3:204795356..204798873	XM_020551019.2	944	678	8	225	26.02	6.68	未鉴定蛋白
ZmCBS18	LOC103651283	Chr3:210324313..210345002	XM_008676907.3	2 738	2 172	7	723	77.32	5.24	离子通道蛋白 CLC-e
ZmCBS19	LOC103651458	Chr3:221793307..221796249	XM_008677091.2	1 564	1 224	7	407	44.51	5.43	SNF1 相关蛋白激酶调节亚单位 γ-1 (CKSy-1)
ZmCBS20	LOC100286384	Chr3:226886250..226889253	NM_0011159271.1	1 787	1 356	5	451	48.42	4.64	未鉴定蛋白
ZmCBS21	LOC103651948	Chr3:55877395..55879569	XM_020550704.2	1 332	456	5	151	15.90	10.00	未鉴定蛋白
ZmCBS22	LOC100282160	Chr4:176416310..176418140	NM_0011155072.1	963	669	5	222	23.84	6.75	未鉴定蛋白
ZmCBS23	LOC100279220	Chr4:19423034..19426369	NM_0011152244.2	1 845	1 656	13	551	59.10	7.21	CBS 结构域蛋白 CBSCBSPB1-like
ZmCBS24	LOC100216607	Chr4:210801830..210807644	XM_008678433.3	2 369	1 638	14	545	58.32	6.46	CBS 结构域蛋白 CBSCBSPB2
ZmCBS25	LOC100282782	Chr4:68296676..68312881	NM_0011155688.1	1 264	699	8	232	24.61	9.07	CBS 结构域蛋白 CBSX1
ZmCBS26	LOC100383583	Chr5:108542805..108572765	NM_001362092.1	2 739	2 382	6	793	86.65	8.76	未鉴定蛋白
ZmCBS27	LOC542114	Chr5:177701748..177710625	XM_035958850.1	3 400	2 616	8	871	94.95	8.73	未鉴定蛋白
ZmCBS28	LOC100275204	Chr5:211724939..211731039	XM_023302441.2	3 527	2 682	9	893	94.86	8.28	未鉴定蛋白
ZmCBS29	LOC100273954	Chr5:25587324..25589546	NM_0011148343.1	1 479	1 125	2	374	38.97	5.25	CBS 结构域蛋白 CBSX5
ZmCBS30	LOC103626636	Chr5:83797272..83799248	XM_008647039.4	1 792	1 278	3	425	44.92	5.19	SNF1 相关蛋白激酶调节亚单位 γ (CKSy-like)
ZmCBS31	LOC103629516	Chr6:86754531..86760231	XM_035959965.1	3 408	639	7	212	23.42	7.11	蔗糖非发酵蛋白 SNF4-like
ZmCBS32	LOC100382044	Chr8:144150005..144155178	XM_008656966.3	2 056	861	7	286	31.84	9.11	蔗糖非发酵蛋白 SNF4-like
ZmCBS33	LOC100273720	Chr8:167736966-167741663	NM_001360302.1	2 149	1 638	14	545	58.86	7.48	CBS 结构域蛋白 CBSCBSPB1
ZmCBS34	LOC100384442	Chr8:167938106..167940462	XM_008657104.4	2 066	1 311	4	436	47.42	5.23	SNF1 相关蛋白激酶调节亚单位 γ-1 (CKSy-1)
ZmCBS35	LOC103635439	Chr8:80402024..80402986	XM_020542595.1	392	345	4	114	12.92	10.48	蔗糖非发酵蛋白 SNF4-like
ZmCBS36	LOC100279316	Chr10:114815824..114823435	NM_001360964.1	1 697	1 347	13	448	49.88	6.11	蔗糖非发酵蛋白 SNF4
ZmCBS37	LOC100384199	Chr10:121555116..121568952	NM_001353414.1	3 016	2 388	7	795	86.30	6.92	未鉴定蛋白

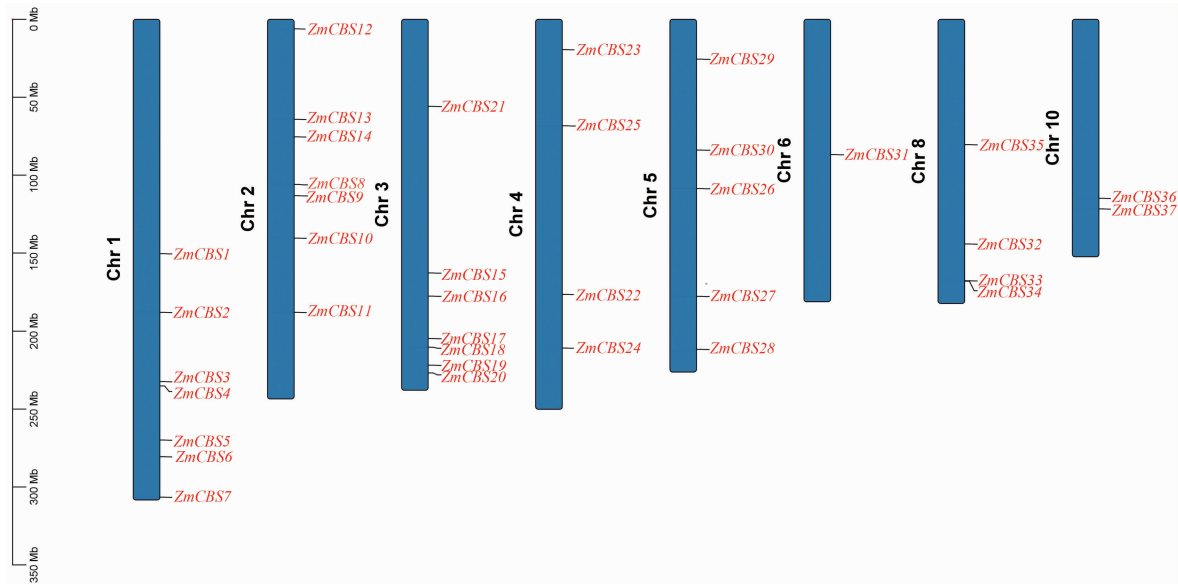


图1 ZmCBS 基因的染色体定位

基因,第5、第6、第8条染色体上分别有5、1、2个基因。

2.3 ZmCBS 基因结构、蛋白结构域及进化关系分析

为了了解玉米 CBS 基因家族各成员的系统进化关系和蛋白保守结构域,使用 NJ 法构建系统发育树,并利用 Batch CD - Search 网站分析了 ZmCBS 家族蛋白保守基序。从图 2 可以看出,ZmCBS 基因家族蛋白大致分为 6 个亚家族(Ⅰ、Ⅱ、Ⅲ、Ⅳ、Ⅴ和Ⅵ),每个亚家族都有明显的特征。Ⅰ组包括 9 个 ZmCBS 蛋白,它们除了 CBS 结构域外,还存在 1 个氯化物通道蛋白(CIC)结构域;Ⅱ组包括 11 个 ZmCBS 蛋白,它们仅含有 CBS 结构域(ZmCBS20 含有 2 个 CBS 结构域);Ⅲ组包括 5 个 ZmCBS 蛋白,它们由 2 个 CBS 结构域和 1 个 Phox/Bem1(PB1)结构域构成;Ⅳ组只有 2 个 ZmCBS 蛋白,其仅含有 1 个 CBS 结构域;Ⅴ组仅有 1 个 ZmCBS 蛋白,由 1 个五肽重复序列(pentatricopeptide repeat, PPR)和 CBS 结构域构成;Ⅵ组有 9 个 CBS 蛋白,它们是蔗糖非发酵 4 相关蛋白(sucrose nonfermenting 4-like protein)(表 1),这类蛋白存在 1~2 个 CBS 结构域,其中 ZmCBS13、ZmCBS14、ZmCBS32 和 ZmCBS36 还含有一个 SET 结构域。

内含子和外显子的结构,以及内含子类型与数量也是一个基因家族典型的进化印迹^[1]。基因结构分析结果显示,同一进化分支上的基因内含子与外显子数量相近且结构组成相似,但内含子长度和

数量差异很大,如Ⅰ组的 ZmCBS27 和 ZmCBS37,Ⅲ组的 ZmCBS24 和 ZmCBS10,以及Ⅵ组的 ZmCBS13 和 ZmCBS36(图 2)。

2.4 ZmCBS 家族基因启动子顺式作用元件分析

对 ZmCBS1~ZmCBS37 的启动子区域进行启动子顺式作用元件分析,分析结果如图 3 所示。ZmCBS1~ZmCBS37 启动子区域普遍存在茉莉酸甲酯、脱落酸(abscisic acid)、光(light)和缺氧(anaerobic induction)等响应元件。除此之外,一些基因启动子还存在其他与外界响应相关的调控元件。Ⅰ、Ⅱ组大多数成员都含有干旱响应元件(drought-inducibility responsiveness)和低温响应元件(low-temperature responsiveness)。Ⅲ组成员含有干旱响应元件和防御及应激响应元件(defense and stress responsiveness),此外防御及应激响应元件也存在于Ⅱ、Ⅳ组的很多成员中。在Ⅵ组的很多成员中存在胚乳表达相关元件(endosperm-expressed element)。这预示着 ZmCBS1~ZmCBS37 在响应外界胁迫过程中发挥着不同的调控作用。

2.5 ZmCBS 基因表达模式分析

基于玉米发育过程中的转录组数据^[21],分析 ZmCBS 基因的组织表达特性。结果显示,Ⅰ组成员 ZmCBS16,Ⅱ组 ZmCBS15,以及Ⅲ组 ZmCB23 和 ZmCB33 在成熟的花粉中表达量明显高于其他组织;而Ⅰ组 ZmCBS5 和 ZmCBS18,Ⅱ组 ZmCBS8、ZmCBS25、ZmCBS9 和 ZmCB22,Ⅲ组 ZmCBS1 和 ZmCBS10 及Ⅴ组 ZmCBS11 在成熟的叶片中表达量

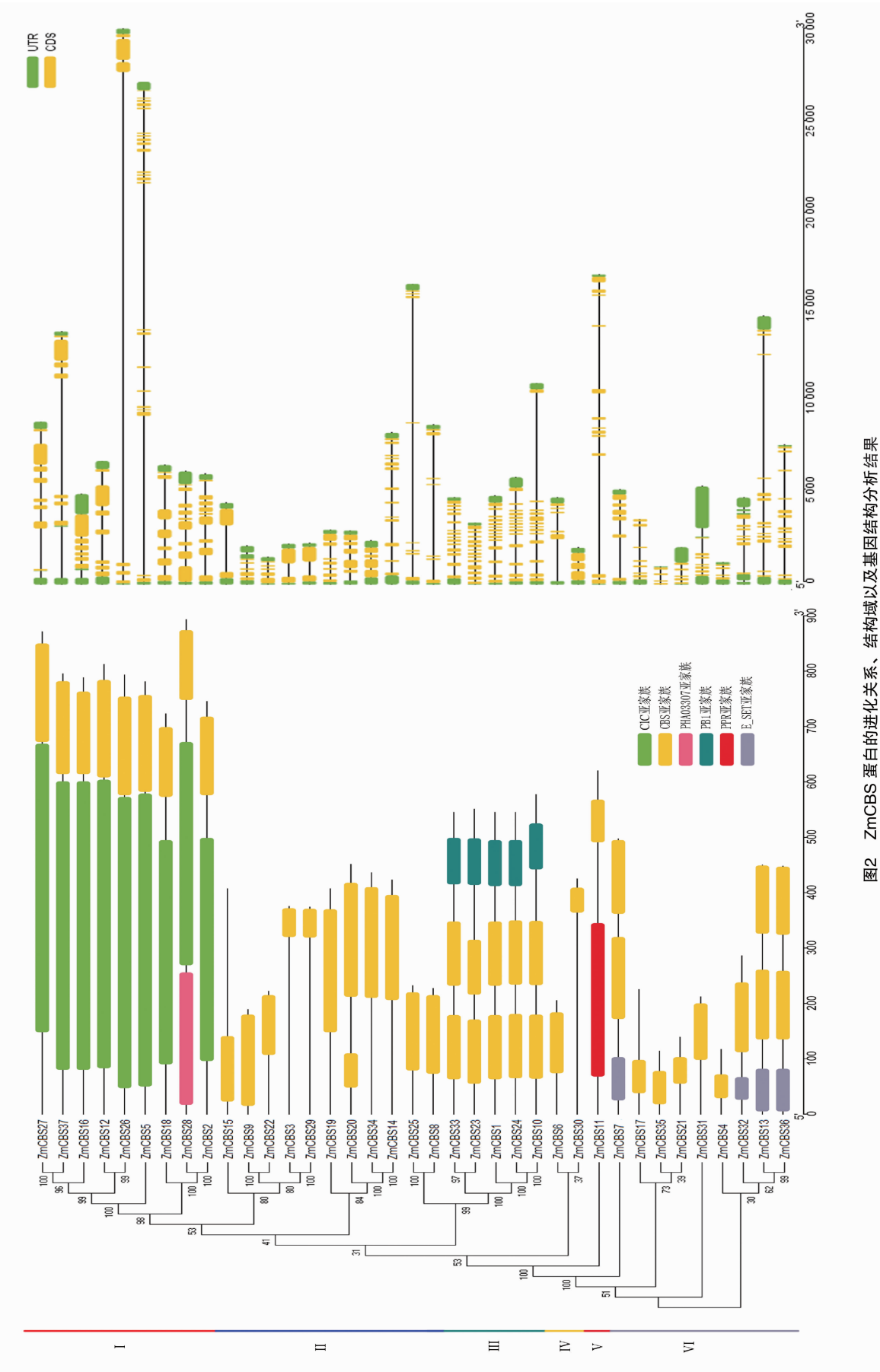


图2 ZnCBS 蛋白的进化关系、结构域以及基因结构分析结果

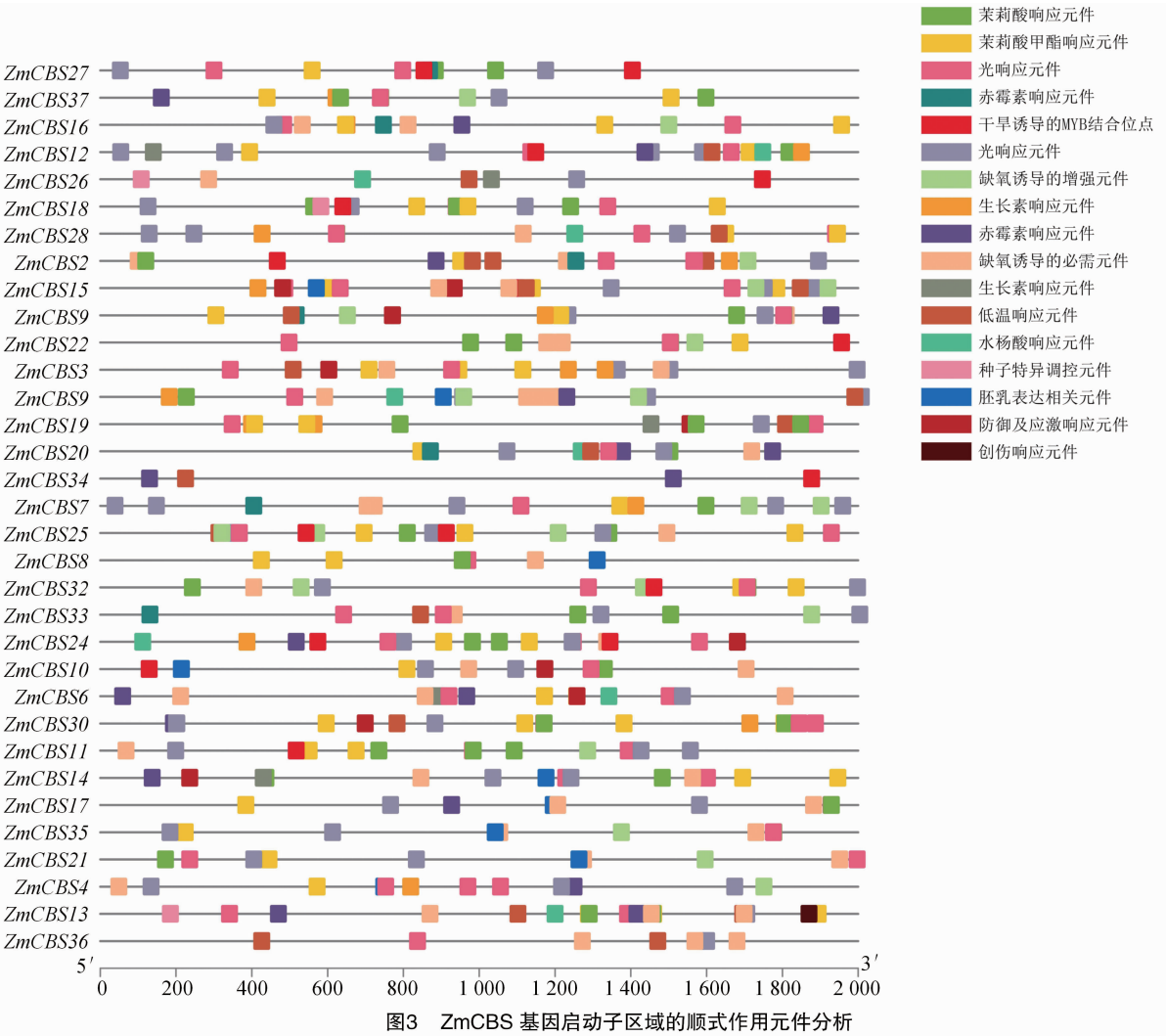


图3 ZmCBS 基因启动子区域的顺式作用元件分析

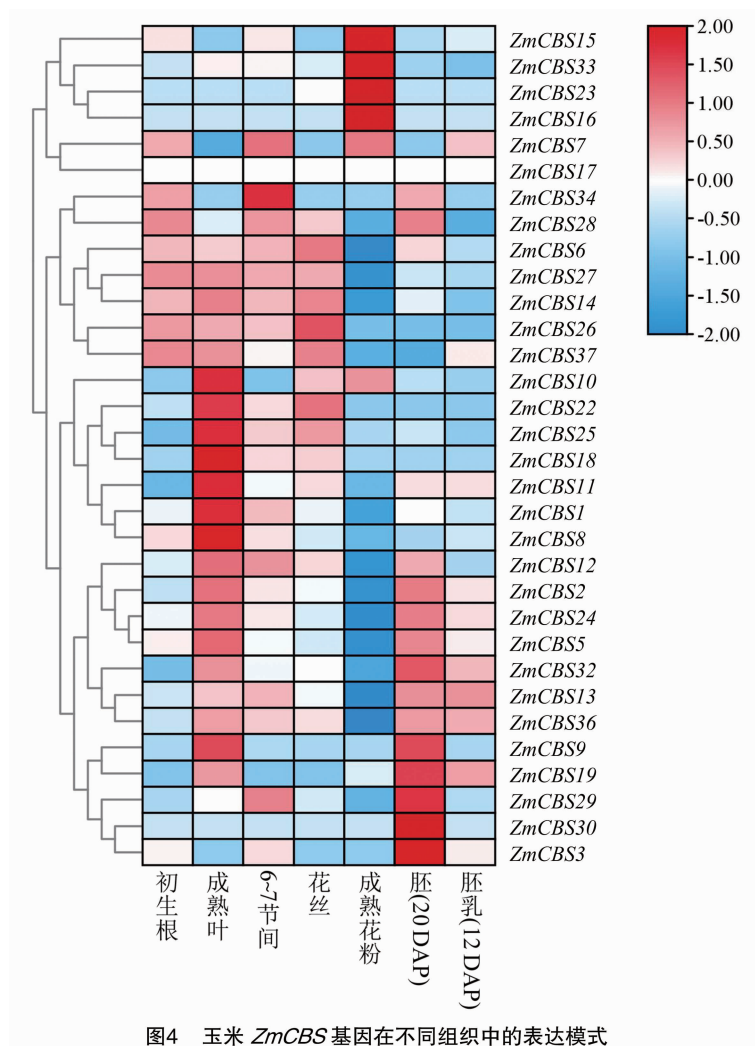
最高;在胚中表达量最高的为Ⅱ组的 *ZmCBS3*、*ZmCBS9*、*ZmCBS19* 和 *ZmCBS29*,Ⅳ组的 *ZmCBS30*,及Ⅵ组的 *ZmCBS32*,这些基因都只含有 1 个 CBS 结构域;在茎(6~7 节间)和花丝中表达量最高的为Ⅱ组 *ZmCBS34* 和Ⅰ组 *ZmCBS26*;另外, *ZmCBS6*、*ZmCBS27*、*ZmCBS14*、*ZmCBS26* 和 *ZmCBS37* 在根、节间、叶和花丝等营养器官中均具有较高的表达水平(图 2 和图 4)。上述 ZmCBS 家族成员差异表达模式表明,ZmCBS 家族可能在调控玉米不同组织器官发育过程中执行功能。

基于玉米受冷和热以及干旱和盐胁迫的转录组数据^[22-23],对 ZmCBS 家族基因进行分析。从图 5 可以看出,除 *ZmCBS26*,其他成员的表达水平受热胁迫后均有不同程度下调。*ZmCBS3* 和 *ZmCBS14* 受冷胁迫后表达水平明显增加(图 5 - A)。干旱胁迫明显上调了 *ZmCBS22* 和 *ZmCBS14* 基因的表达,盐胁迫对 ZmCBS 家族基因的影响不明显(图 5 -

B)。从这些结果推测,大多数 ZmCBS 基因对热胁迫更加敏感,少数基因在干旱和低温胁迫过程中发挥功能。

3 结论与讨论

CBS 基因家族是由一类含有保守的 CBS 结构域的 CDCPs 构成,这类家族几乎存在于所有物种中。到目前为止,对 CDCPs 的研究主要集中在人类或其他动物^[24-26]。植物 CDCPs 报道较少,其主要是通过与 Trxs 相互作用来维持氧化还原稳态,但相关作用机制仍需进一步探讨^[3]。近年来,随着基因组学和生物信息学的发展,相继从拟南芥、水稻、小麦和大豆中鉴定出 CDCPs^[1,10,27]。这些蛋白根据序列中的其他结构域可进一步分成多个亚类: CBSX, 仅含单个 CBS 结构域; CBSDUF21, 携带 CBS 和 1 个未知功能结构域(DUF21); CBSCIC, CBS + CIC; CBSPPR, CBS + PPR; CBSCBSPB, CBS + CBS + PB1;



CBSIMPDH, CBS + IMPDH 等^[10,27]。本研究中,从玉米基因组中鉴定出 37 个 CBS 基因,并分析了这些 CDCPs 的进化关系和保守结构域。基于上述的分类方法,大致将其分为 6 类(图 2)。分类结果没有发现 CBSDUF21 亚类的 CDCPs 存在,可能是因为玉米该亚类成员在 Batch CD - Search 验证 CBS 结构域中被剔除。从基因结构上看,同一进化分支上的基因内含子与外显子数量相近且结构组成相似,但内含子长度和数量差异很大(图 2),推测内含子的插入可能促进了玉米 *CBS* 基因家族的进化^[28]。上述分析结果推测,CBS 各成员间基因、蛋白序列和保守结构域的差异驱动了 CBS 基因家族功能的分化。

之前的研究显示,植物 CBS 家族基因在很多生物过程中都发挥了重要功能,包括花药发育过程,对生物和非生物胁迫的抗性^[5,6,11-17]。一些 CDCPs 的表达具有组织特异性,或者因植物受外界胁迫发

生变化。例如, *AtCBSX1*、*AtCBSX2* 和 *AtCBSX3* 在花粉中高表达, 并且参与了拟南芥的花粉发育^[5, 11-12]。基于 Walley 等的玉米转录组数据, 可以发现Ⅱ组 *ZmCBS15* 以及Ⅲ组 *ZmCB23* 和 *ZmCB33* 在成熟的花粉中表达量明显高于其他组织(图4), 这暗示除 CBX 亚类成员外, CBSClC 和 CBSCBSPB 亚类成员也可能参与了玉米花粉发育过程。此外, 一些只含 CBS 结构域的成员(*ZmCBS3*、*ZmCBS9*、*ZmCBS19*、*ZmCBS29*、*ZmCBS30* 和 *ZmCBS32*) 在胚中表达量最高(图4), 暗示它们在胚发育过程中可能发挥功能。

研究表明,拟南芥受冷胁迫下, *AtCBSDUF3*、*AtCBSCLC7*、*AtCBSCBSPB4*、*AtCBSCBSPB5* 在根胁迫 24 h 时表达上调;在盐胁迫下,氧化胁迫诱导部分 CBS 基因在根部胁迫 3 h 时表达,而在芽中几乎所有差异表达的 CBS 基因在氧化胁迫下均上调表达;

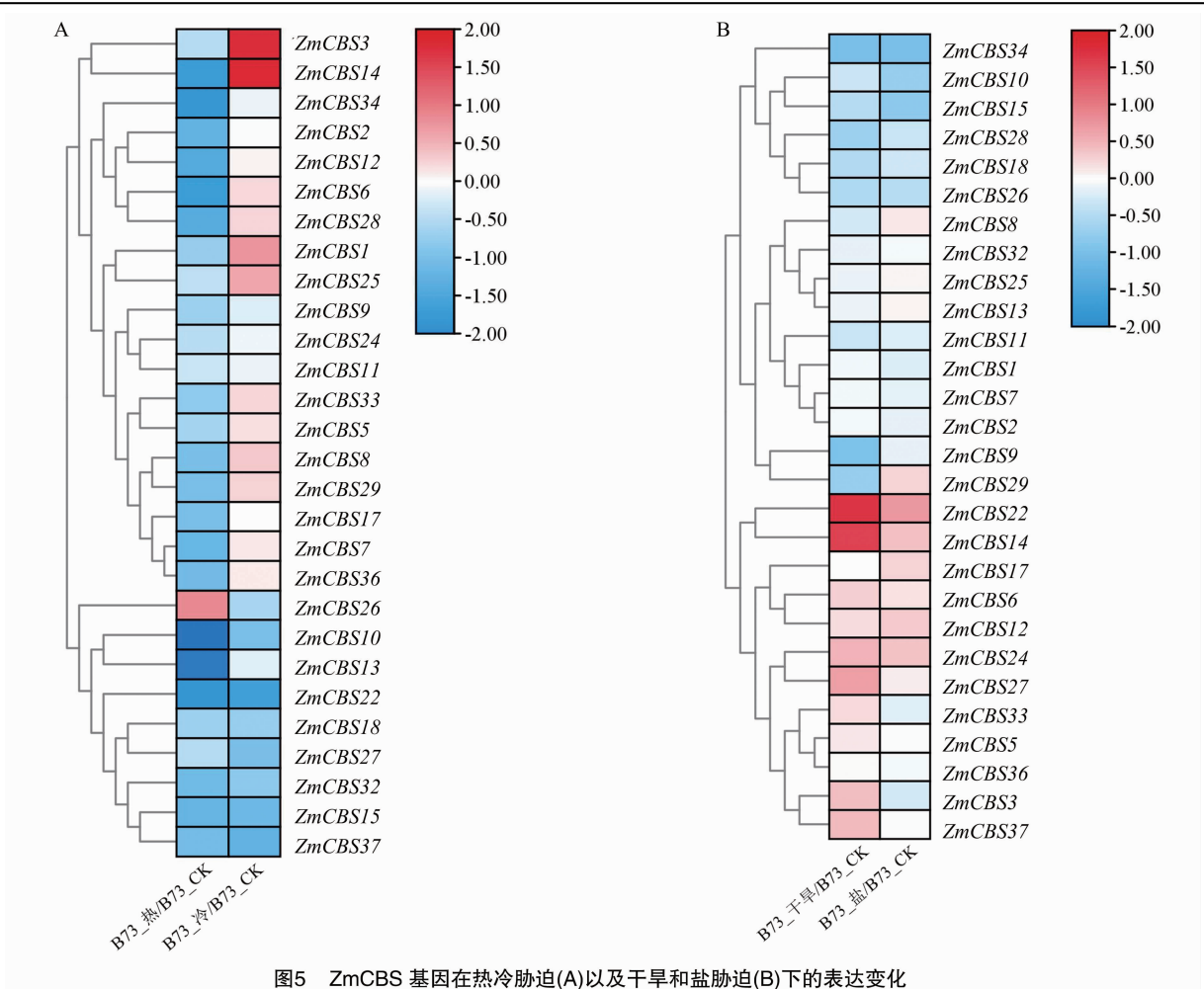


图5 ZmCBS 基因在热冷胁迫(A)以及干旱和盐胁迫(B)下的表达变化

在干旱胁迫下,所有 CDCPs 在根中不同时间点的表达量相当,而在茎中几乎全部的 CDCPs 在干旱胁迫 24 h 时表达上调^[27]。郭富桦等发现,小麦在 5 ℃低温下大多数 CBS 基因表达量并没有明显的变化,但是在 - 10、- 25 ℃下, *TaCBS8s*、*TaCBSCBS1s* 和 *TaCBS10s* 表达量均有明显的升高^[1]。本研究发现,玉米 B73 受外界胁迫后,只有少部分 CDCPs 的表达发生了显著改变,推测这与取样的时间和组织相关。其中, *ZmCBS3* 和 *ZmCBS14* 的表达水平受冷胁迫显著增加, *ZmCBS22* 和 *ZmCBS14* 的表达水平受干旱胁迫显著增加(图 5 - A 和图 5 - B),暗示这些基因在干旱和低温胁迫过程中发挥功能。此外启动子元件分析结果显示,玉米 CBS 基因的启动子区域存在多种与环境响应密切相关的调控元件(图 3),如茉莉酸甲酯、脱落酸、光、缺氧、干旱和低温等响应元件^[29-31]。这些结果都预示着 ZmCBS 基因可能与生长发育和非生物胁迫响应相关基因的表达调控有关。

参考文献:

[1]郭富桦,苍晶,卢秋巍,等. 六倍体小麦 CBS 基因家族全基因组分析[J]. 麦类作物学报,2020,40(4):421-433.

[2]Bateman A. The structure of a domain common to archaeobacteria and the homocystinuria disease protein[J]. Trends in Biochemical Sciences,1997,22(1):12-13.

[3]Ignoul S J. CBS domains: structurefunctionand pathology in human proteins[J]. American Journal of Physiol Cell Physiology,2005,289(6):1369-1378.

[4]Xiao B,Heath R,Saiu P,et al. Structural basis for AMP binding to mammalian AMP - activated protein kinase[J]. Nature,2007,449(7161):496-500.

[5]Yoo K,Ok S,Jeong B C,et al. Single cystathionine β - synthase domain - containing proteins modulate development by regulating the thioredoxin system in *Arabidopsis* [J]. Plant Cell,2011,23(10):3577-3594.

[6]Ke X L,Xiao H,Peng Y Q,et al. Phosphoenolpyruvate reallocation links nitrogen fixation rates to root nodule energy state[J]. Science,2022,378(6623):971-977.

[7]Sintchak M,Fleming M,Futer O,et al. Structure and mechanism of

- inosine monophosphate dehydrogenase in complex with the immunosuppressant mycophenolic acid [J]. Cell, 1996, 85 (6): 921–930.
- [8] Woods A, Cheung P, Smith F, et al. Characterization of AMP-activated protein kinase beta and gamma subunits; assembly of the heterotrimeric complex *in vitro* [J]. Journal Biological Chemistry, 1996, 271 (17): 10282–10290.
- [9] Schmidt-Rose T, Jentsch T. Reconstitution of functional voltage-gated chloride channels from complementary fragments of CLC-1 [J]. Journal Biological Chemistry, 1997, 272 (33): 20515–20521.
- [10] Liu H Z, Wang Q, Xie L Y, et al. Genome-wide identification of cystathionine beta synthase genes in wheat and its relationship with anther male sterility under heat stress [J]. Frontier in Plant Science, 2022, 13: 1061472.
- [11] Jung K, Kim Y, Yoo K, et al. A cystathionine- β -synthase domain-containing protein CBSX2 regulates endotheical secondary cell wall thickening in anther development [J]. Plant Cell Physiology, 2013, 54 (2): 195–208.
- [12] Shin J S, So W M, Kim S Y, et al. CBSX3-Trxo-2 regulates ROS generation of mitochondrial complex II (succinate dehydrogenase) in *Arabidopsis* [J]. Plant Science, 2020, 294: 110458.
- [13] Mou S L, Shi L P, Lin W, et al. Over-expression of rice CBS domain containing protein OsCBSX3 confers rice resistance to *Magnaporthe oryzae* inoculation [J]. International Journal of Molecular Science, 2015, 16 (7): 15903–15917.
- [14] Singh A, Kumar R, Pareek A, et al. Overexpression of rice CBS domain containing protein improves salinity oxidative and heavy metal tolerance in transgenic tobacco [J]. Molecular Biotechnology, 2012, 52 (3): 205–216.
- [15] Kumar R, Subba A, Kaur C, et al. OsCBSCBSPB4 is a two cystathionine- β -synthase domain-containing protein from rice that functions in abiotic stress tolerance [J]. Current Genomics, 2018, 19 (1): 50–59.
- [16] Hao Q N, Yang Y Y, Shan Z H, et al. Genome-wide investigation and expression profiling under abiotic stresses of a soybean unknown function (DUF21) and cystathionine- β -synthase (CBS) domain-containing protein family [J]. Biochemistry Genetics, 2021, 59 (1): 83–113.
- [17] Hao Q N, Shang W, Zhang C J, et al. Identification and comparative analysis of CBS domain-containing proteins in soybean (*Glycine max*) and the primary function of *GmCBS21* in enhanced tolerance to low nitrogen stress [J]. International Journal of Molecular Science, 2016, 17 (5): 620.
- [18] 赵久然, 王元东, 邢锦丰, 等. 高产优质、多抗广适玉米品种京科 968 的培育与应用 [J]. 中国科技成果, 2021, 22 (1): 70–71.
- [19] Kumar S, Stecher G, Li M, et al. MEGA X: molecular evolutionary genetics analysis across computing platforms [J]. Molecular Biology and Evolution, 2018, 35 (6): 1547–1549.
- [20] Lescot M, Déhais P, Moreau Y, et al. PlantCARE: a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences [J]. Nucleic Acids Research, 2002, 30 (1): 325–327.
- [21] Walley J, Sartor R, Shen Z, et al. Integration of omic networks in a developmental atlas of maize [J]. Science, 2016, 353 (6301): 814–818.
- [22] Waters A, Makarevitch I, Noshay J, et al. Natural variation for gene expression responses to abiotic stress in maize [J]. Plant Journal, 2017, 89 (4): 706–717.
- [23] Forestan C, Aiese-Cigliano R, Farinati S, et al. Stress-induced and epigenetic-mediated maize transcriptome regulation study by means of transcriptome reannotation and differential expression analysis [J]. Scientific Report, 2016, 6: 30446.
- [24] Yuan Y Q, Wang Y L, Yuan B S, et al. Impaired CBS-H2S signaling axis contributes to MPTP-induced neurodegeneration in a mouse model of Parkinson's disease [J]. Brain Behavior Immunity, 2018, 67: 77–90.
- [25] 李金平, 李丽帆, 方显明. 胱硫醚 β -合酶研究进展 [J]. 西部医学, 2006, 18 (5): 657–659.
- [26] Bertoni G. CBS domain proteins regulate redox homeostasis [J]. Plant Cell, 2011, 23 (10): 3562.
- [27] Kushwaha H R, Singh A K, Sopory S K, et al. Genome wide expression analysis of CBS domain containing proteins in *Arabidopsis thaliana* (L.) Heynh and *Oryza sativa* L. reveals their developmental and stress regulation [J]. BMC Genomics, 2009, 10: 200.
- [28] Xu G X, Guo C, Shan H Y, et al. Divergence of duplicate genes in exon-intron structure [J]. PNAS, 2012, 109 (4): 1187–1192.
- [29] 张新宁, 邢真真, 李静, 等. 小麦 MBD 基因家族的鉴定和特征分析 [J]. 西北植物学报, 2021, 41 (5): 746–756.
- [30] Fu M, Wu C, Li X, et al. Genome-wide identification and expression analysis of CsCaM/CML gene family in response to low-temperature and salt stresses in *Chrysanthemum seticuspe* [J]. Plants, 2022, 11 (13): 1760.
- [31] Zhang Z J, Huang B, Chen J L, et al. Genome-wide identification of JRL genes in moso bamboo and their expression profiles in response to multiple hormones and abiotic stresses [J]. Frontier in Plant Science, 2022, 12: 809666.