

玉米 *COBRA* 家族成员全基因组鉴定与表达模式分析

潘倚天, 黄敏

(长江大学生命科学学院, 湖北 荆州 434025)

摘要: *COBRA* 基因编码糖基磷脂酰肌醇锚定的植物特异性蛋白, 在初生和次生细胞壁的纤维素生物合成中发挥重要作用。通过生物信息学方法, 从玉米基因组中发现了9个 *COBRA* 家族基因, 并对其基因结构、系统发育和表达模式等进行了分析。结果表明, 9个 *COBRA* 家族基因都含有CCVS保守结构域, 并且均定位于细胞膜上。系统进化分析结果显示, 该家族可以分为2个亚族, 每个亚族内的基因具有相似的基因结构和理化性质。基因表达分析结果表明, 所有 *COBRA* 家族成员响应多种非生物胁迫, 且当植株受到紫外照射处理时, *ZmCOBL1~ZmCOBL3* 和 *ZmCOBL7~ZmCOBL9* 这6个基因均表现出上调应答。

关键词: 玉米; *COBRA* 基因家族; 纤维素; 表达模式; 生物信息学

中图分类号: S513

文献标志码: A

文章编号: 1001-1463(2022)02-0067-06

doi:10.3969/j.issn.1001-1463.2022.02.017

Genome-wide Identification and Expression Pattern Analysis of *COBRA* Family Members in Maize

PAN Yitian, HUANG Min

(College of Life Sciences, Yangtze University, Jingzhou Hubei 434025, China)

Abstract: The *COBRA* gene encodes a glycosyl phosphatidyl inositol-anchored, plant-specific protein that plays an important role in cellulose biosynthesis in primary and secondary cell walls. In this study, 9 *COBRA* family genes were found from the maize genome by bioinformatics methods, and their gene structures, phylogeny and expression patterns were analyzed. The results showed that the 9 *COBRA* family genes all contained CCVS conserved domains, and all of them were located on the cell membrane. Phylogenetic analysis showed that *COBRA* family could be divided into two subfamilies, and the genes in each subgroup had similar gene structure and physicochemical properties. Gene expression analysis showed that all *COBRA* family members

收稿日期: 2021-12-21

基金项目: 国家自然科学基金面上项目(31771801、32072069); 湖北省高等学校优秀中青年科技创新团队计划项目(T201704)。

作者简介: 潘倚天(1996—), 男, 河南开封人, 硕士研究生, 主要从事玉米涝渍灾害研究工作。联系电话: (0)15872100154。Email: ytpan1996@163.com。

通信作者: 黄敏(1957—), 女, 湖北随州人, 实验员, 研究方向为生命科学。Email: 2381615354@qq.com。

- [8] 汪东晖. 武山县早春茬大棚花椰菜栽培技术[J]. 甘肃农业科技, 2020(7): 83-84.
- [9] 俄胜哲, 黄涛, 袁洁, 等. 渭河上游长期设施种植对土壤盐渍化及硝态氮含量的影响[J]. 甘肃农业科技, 2015(8): 19-22.
- [10] 刘强, 刘雪媛. 天水市秦州区大樱桃果园土壤重金属调查与评价[J]. 天水师范学院学报, 2017(2): 42-45.
- [11] 刘瑞雪, 乔冬云, 王萍, 等. 湘潭县农田土壤重金属污染及生态风险评价[J]. 农业环境科学学报, 2019, 38(7): 1523-1530.
- [12] 生态环境部国家市场监督管理总局. GB15618-2018, 土壤环境质量农用地土壤污染风险管控标准(试行)[S]. 北京: 中国标准出版社, 2018.
- [13] 赵辛金, 吴天生, 钟晓宇, 等. 广西典型岩溶区重金属高背景区农田土壤生态风险综合评价[J]. 江苏农业科学, 2020, 48(22): 252-261.
- [14] 孙英, 周金龙, 曾妍妍, 等. 新疆于田县农田土壤重(类)金属污染及潜在生态风险评价[J]. 新疆农业科学, 2018, 55(12): 2271-2278.
- [15] 贾英, 方明, 吴友军, 等. 上海河流沉积物重金属的污染特征与潜在生态风险[J]. 中国环境科学, 2013, 33(1): 147-153.
- [16] 洪启金, 李少能. 广州市番禺区无公害蔬菜生产存在问题及发展对策[J]. 广东农业科学, 2005(4): 115-116.

responded to a variety of abiotic stresses, and all six genes, *ZmCOBL1~ZmCOBL3* and *ZmCOBL7~ZmCOBL9*, showed up-regulated responses when plants were exposed to UV irradiation.

Key words: Maize (*Zea mays* L.); *COBRA* gene family; Cellulose; Express pattern; Bioinformatics

玉米(*Zea mays* L.)是全世界最重要的粮食作物之一,在全球范围内分布十分广泛。2018年,我国玉米总产量占到全国粮食总产量的39%以上,播种面积占到全国粮食作物播种面积的36%^[1],可见玉米的产量高低与我国粮食安全息息相关。我国的玉米产量常受倒伏影响造成不同程度的减产,因此通过筛选玉米与茎秆机械强度相关基因,增强玉米抗倒伏能力,是提升玉米产量和保证国家粮食安全的重要措施。

COBRA 家族是由 *COBRA-Like*(*COBL*)蛋白组成的一类糖基磷脂酰肌醇(Glycosyl phosphatidyl inositol, GPI)锚定蛋白,在初生和次生细胞壁的纤维素生物合成中发挥核心作用^[2]。*COBRA* 家族首先在拟南芥中被发现,*COBRA* 突变体根异常生长^[3]。*Brittle Stalk2* (*BK2*)是在玉米中发现的首个 *COBRA* 家族成员^[4],其在纤维素合成中发挥功能,并在成熟后的次生细胞壁发育发挥适当的作用^[5]。在大多数情况下,*COBRA* 基因家族突变体表现出纤维素含量降低^[6]。利用生物信息学对玉米 *COBRA* 家族基因进行系统分析有助于预测 *COBRA* 家族成员的功能,将为后续研究该家族成员的生物学功能和应用于玉米分子育种提供重要依据。

1 材料与方法

1.1 *COBRA* 家族成员的鉴定和染色体定位

根据高粱的10个 *COBRA* 基因家族成员的序列^[7],使用 BlastP 在玉米全基因组蛋白序列数据库中搜索,得到候选序列。典型的 *COBRA* 家族成员含有1个 *COBL* 结构域(PF04833)^[8]。通过 BatchCD-search 功能和 Pfam 检验验证序列是否含有保守结构域,并删去重复序列。使用 TBtools 软件将玉米 *COBRA* 家族成员在染色体上的位置可视化^[9],并根据基因在染色体上的位置进行命名。

使用同样的方法分别鉴定出水稻、高粱和拟南芥的所有 *COBRA* 基因家族成员,并根据在染色

体上的位置进行命名。

1.2 玉米 *COBRA* 家族成员序列分析

使用 Expasy 网站预测 *ZmCOBL* 蛋白的分子量和等电点,使用 SignalP-5.0 预测 N 端信号肽,并用 big-PI 预测 C 端疏水末端的锚定位点^[10]。用 MEME 网站分析蛋白序列中的保守基序,使用 Softberry 网站预测亚细胞定位。用 TBtools 分析家族成员的内含子和外显子结构。

1.3 玉米和其他物种的 *COBRA* 蛋白系统发育分析及共线性分析

利用 MEGA11 中的 ClustalW 功能,对4个物种 *COBRA* 蛋白进行多重比对,随后利用最大似然法(Maximum Likelihood)构建系统进化树,Bootstrap 值设置为1000。利用 TBtools 的 One Step MC-ScanX 功能分析玉米与高粱和水稻之间的共线性关系,并用 Dual Systemy Plot 功能进行可视化。

1.4 玉米 *COBRA* 家族成员的表达模式分析

利用 MaizeGDB 中的 qTeller 功能下载 *COBRA* 家族成员的 RNA-Seq 数据^[11],经过均一化处理后使用 TBtools 进行热图绘制。

2 结果与分析

2.1 玉米 *COBRA* 家族成员的鉴定和染色体定位

通过 BlsatP 比对和重复序列的去除,最终确定玉米全基因组中共含有9个 *COBRA* 家族成员。染色体定位结果显示,9个成员在染色体上的分布不均匀,共分布在5条染色体上。其中1号染色体和7号染色体各分布有3个成员,而2号染色体、5号染色体和9号染色体各含有1个。根据其在染色体上的位置,将这些成员分别命名为 *ZmCOBL1~ZmCOBL9*(图1)。

2.2 玉米 *COBRA* 家族成员序列分析

对玉米 *COBRA* 家族成员进行序列分析,结果见表1。该家族成员的分子质量为 25.734 59 ~ 75.169 48 kD,平均值为 51.999 79 kD。蛋白质等电点为 6.03 ~ 9.19,其中大于7的有7个成员,小

于7的只有2个成员，说明该基因家族成员多为碱性蛋白。在9个成员中，除 *ZmCOBL2* 和 *ZmCOBL7* 外，其余都含有 N 端信号位点。C 端 GPI 锚定位点预测结果表明，除 *ZmCOBL2* 之外，其余

成员均至少含有 1 个 GPI 锚定位点。亚细胞定位预测显示，所有的 *ZmCOBL* 家族成员均定位在细胞膜上，与其他物种相同^[1]。基因结构分析显示，有 4 个 *ZmCOBL* 家族成员含有 5 个内含子和 6 个外

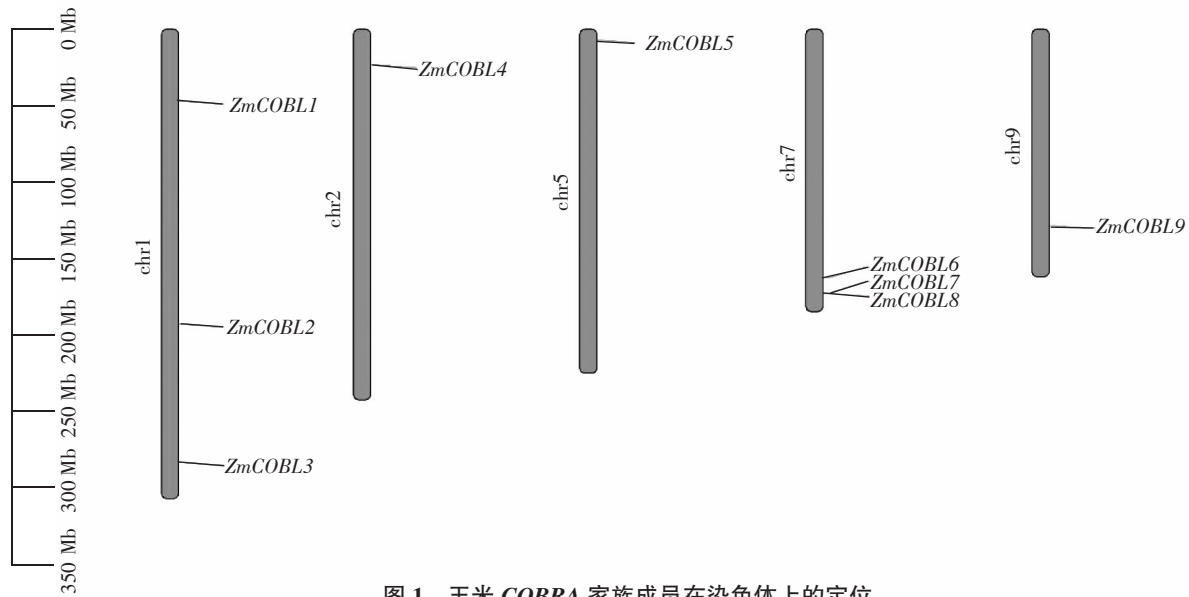


图 1 玉米 COBRA 家族成员在染色体上的定位

表 1 玉米 COBRA 家族成员序列特征

基因名称	基因代码	分子质量 /kD	等电点	N端信号肽位置	潜在ω位点	亚细胞定位	内含子/外显子数 /个
<i>ZmCOBL1</i>	<i>Zm00001eb014120</i>	71.377 40	6.03	23~24	S(640)	细胞膜	0/1
<i>ZmCOBL2</i>	<i>Zm00001eb035300</i>	25.734 59	9.19			细胞膜	2/3
<i>ZmCOBL3</i>	<i>Zm00001eb057580</i>	49.599 87	8.88	30~31	N(422)	细胞膜	5/6
<i>ZmCOBL4</i>	<i>Zm00001eb075290</i>	49.102 96	6.69	22~23	N,G(417,416)	细胞膜	5/6
<i>ZmCOBL5</i>	<i>Zm00001eb214720</i>	49.841 22	8.91	30~31	N(422)	细胞膜	5/6
<i>ZmCOBL6</i>	<i>Zm00001eb323040</i>	75.169 48	8.99	30~31	S,G(646,653)	细胞膜	1/2
<i>ZmCOBL7</i>	<i>Zm00001eb326990</i>	49.090 85	8.83		N,A(415,416)	细胞膜	4/5
<i>ZmCOBL8</i>	<i>Zm00001eb327000</i>	49.046 28	9.04	23~24	N,A(420,426)	细胞膜	5/6
<i>ZmCOBL9</i>	<i>Zm00001eb393070</i>	49.035 50	8.87	27~28	N(422)	细胞膜	2/3

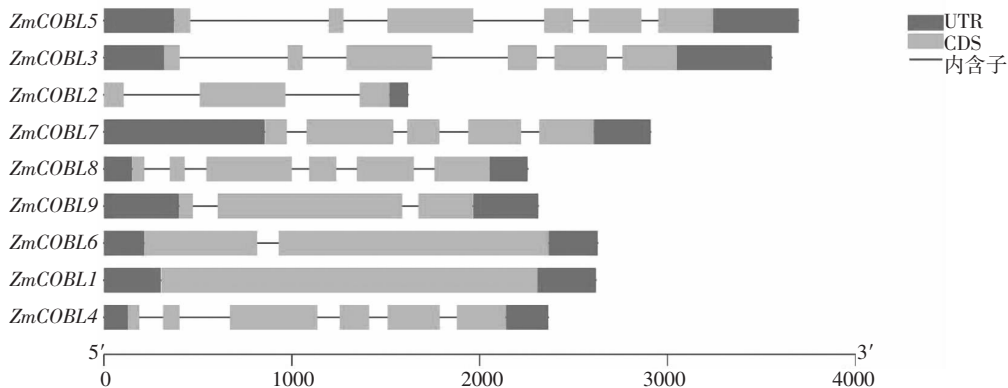


图 2 玉米 COBRA 基因结构

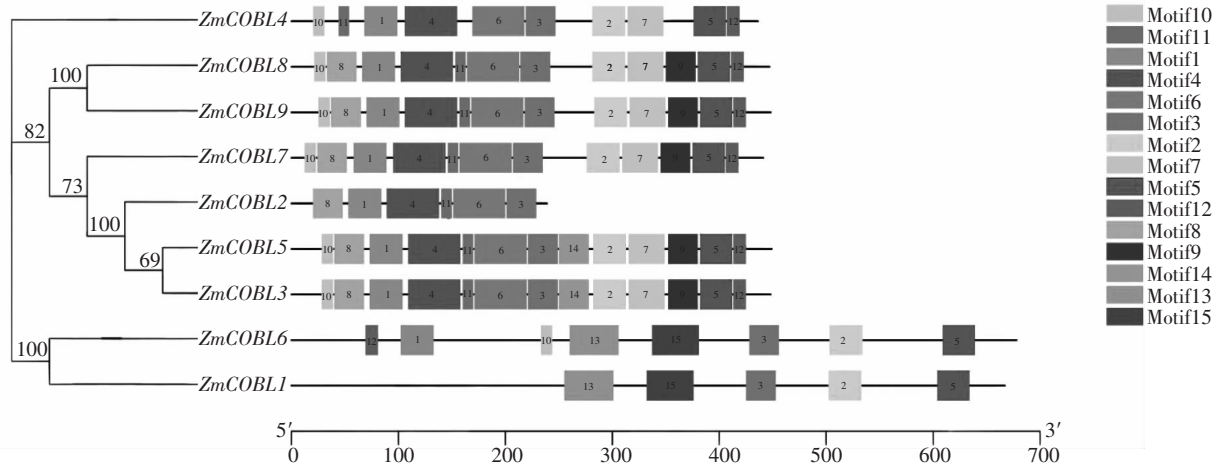


图3 玉米 COBRA 基因家族进化树和 motif 预测结果

显子(图2),说明基因结构在一定程度上较为保守。对保守基序分析后发现,9个 ZmCOBL 家族成员均含有 motif3 (图3),该保守基序含有1个高度保守的序列 CCVS (图4)。此外 motif2 和 motif5 均存在除 ZmCOBL2 之外的8个成员上。

Species/Abbrv					*	*	*						*		
1. ZmCOBL1	-	A	K	P	K	C	C	V	T	F	S	A	Y	Y	N
2. ZmCOBL2	Q	K	T	L	S	C	C	V	S	L	S	S	F	Y	N
3. ZmCOBL3	Q	K	T	P	S	C	C	V	S	L	S	S	F	Y	N
4. ZmCOBL4	S	A	A	P	S	C	C	V	S	M	T	T	F	Y	S
5. ZmCOBL5	Q	K	T	P	S	C	C	V	S	L	S	S	F	Y	N
6. ZmCOBL6	-	R	A	S	K	C	C	V	S	F	S	A	Y	Y	N
7. ZmCOBL7	Q	R	S	P	T	C	C	V	S	L	S	S	F	Y	N
8. ZmCOBL8	S	K	Y	P	S	C	C	V	S	F	S	S	F	Y	N
9. ZmCOBL9	S	R	Y	P	S	C	C	V	S	F	S	S	F	Y	N

图4 玉米 COBRA 家族 CCVS 保守结构域

2.3 玉米和其他物种的 COBRA 蛋白系统发育分析及共线性分析

为了分析 COBRA 蛋白的系统发育关系,将玉米、拟南芥、水稻和高粱等4个物种的 COBRA 蛋白序列,用最大似然法构建进化树(图5)。根据其聚类结果,将41个基因分为2个亚族。玉米 COBRA 家族成员在2个亚族内的数量分别是7个(ZmCOBL2、ZmCOBL3、ZmCOBL4、ZmCOBL5、ZmCOBL7、ZmCOBL8、ZmCOBL9)和2个(ZmCOBL6和ZmCOBL1)。高粱和水稻也出现这种分布不均的现象,亚族1有8个成员,亚族2只有3个成员。拟南芥在2个亚族内的分布较为均匀,分别

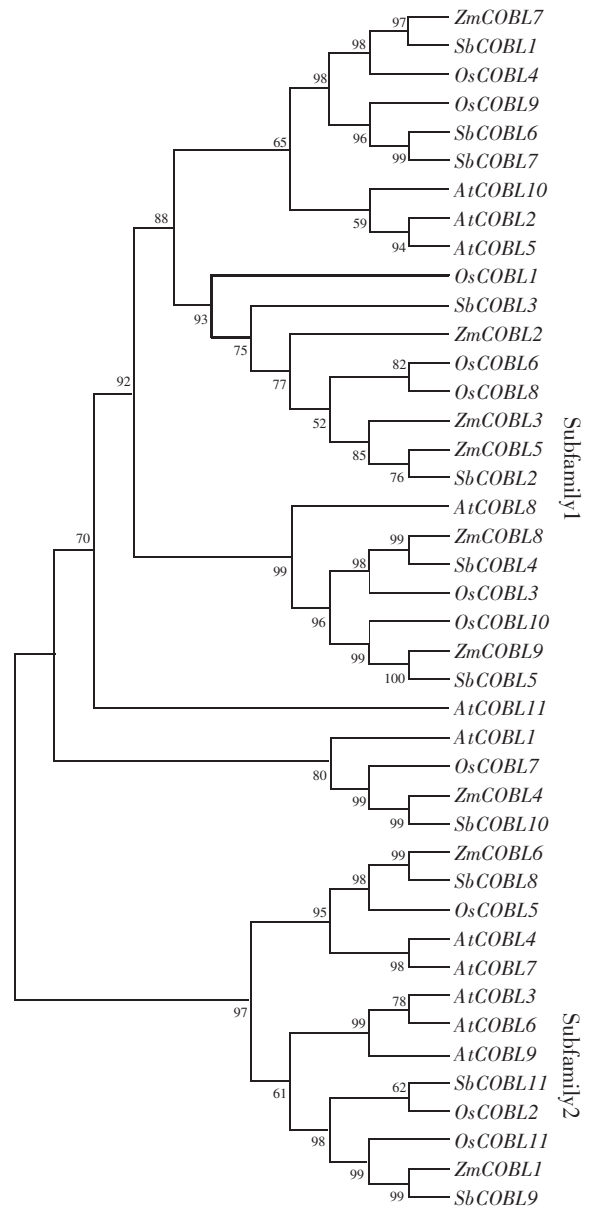


图5 不同物种 COBRA 基因家族进化树

是6个和5个。

为进一步了解 COBRA 家族的进化机制，选取另外 2 个禾本科植物和 1 个模式植物拟南芥，利用 MCSanX 功能分析 COBRA 家族基因间的共线性(图 6)。结果显示，该基因家族在拟南芥和玉米之间没有共线性(图中未显示)。在禾本科植物中，玉米与水稻、高粱分别有 2 个和 5 个共线性基因对，其中 *ZmCOBL1* 在高粱基因组内有 2 个对应基因，*SbCOBL3* 在玉米基因组内也有 2 个对应基因。

2.4 玉米 COBRA 家族成员的表达模式分析

根据 MaizeGDB 公布的 RNA-Seq 数据，筛选出玉米 COBRA 基因家族在不同胁迫处理下的表达量数据，除 *ZmCOBL4* 和 *ZmCOBL6* 之外，其他的 7 个基因均有相对应的表达量。根据上述数据制作表达量热图(图 7)。从图中看出，7 个基因在紫外辐射胁迫时，除 *ZmCOBL5* 外均表现明显的上调响应，而在面对高温胁迫和盐胁迫时，除 *ZmCOBL2* 和 *ZmCOBL8* 外均表现出显著的下调响应，说明这些基因在不同程度上都参与了非生物胁迫的调控。值得一提的是，玉米面对 4 种非生物胁迫时，*ZmCOBL5* 基因均表现出下调应答，而其他基因的表达模式不同。

3 小结与讨论

通过 BlastP 检索，去重和保守结构域验证，共鉴定到 9 个玉米 COBRA 成员，通过其在玉米染色体上的位置分布，分别命名为 *ZmCOBL1*~*ZmCOBL9*。对基因家族成员的序列分析发现，所有蛋白序列都存在 CCVS 保守结构域，8 个蛋白包含 N 端信号肽，7 个含有 GPI 锚定位点，这些特征均

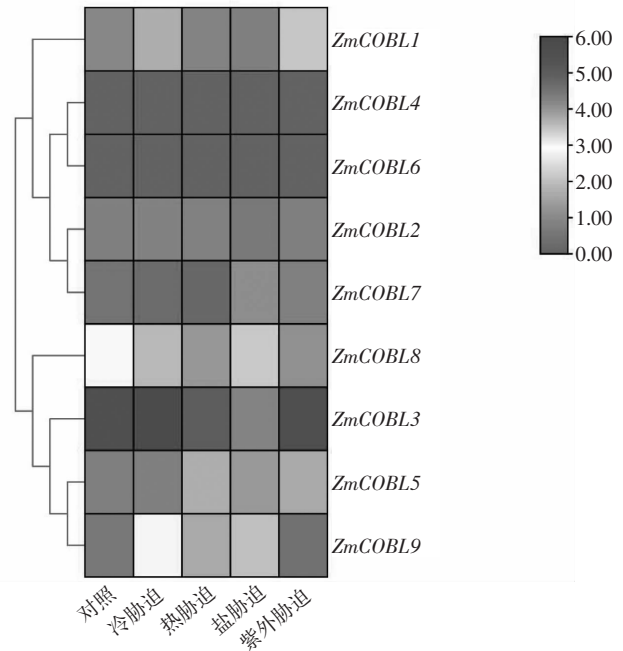


图 7 玉米 COBRA 基因在不同胁迫下表达谱的热图

表明其属于典型的 COBRA 家族蛋白^[12]。经过与拟南芥、高粱和水稻 3 个物种共同构建进化树，发现玉米 COBRA 蛋白和高粱的同源性最高，其次是水稻，而与双子叶植物拟南芥的同源性最低，说明在单子叶植物中 COBRA 蛋白更加保守。此外还发现，这些成员可以分为 2 个亚族，其中 *ZmCOBL6* 和 *ZmCOBL1* 在第二亚族，其他蛋白聚类在第一亚族。与其他禾本科作物一样，玉米也存在成员分布不均的情况，说明 COBRA 蛋白序列在单双子叶植物之间存在一定差异，其差异很有可能是在单双子叶植物分化之后才形成。共线性分析表明，*ZmCOBL1* 同时存在于高粱和水稻 2 个物种的共线性基因对中，表明该基因在单子叶植物

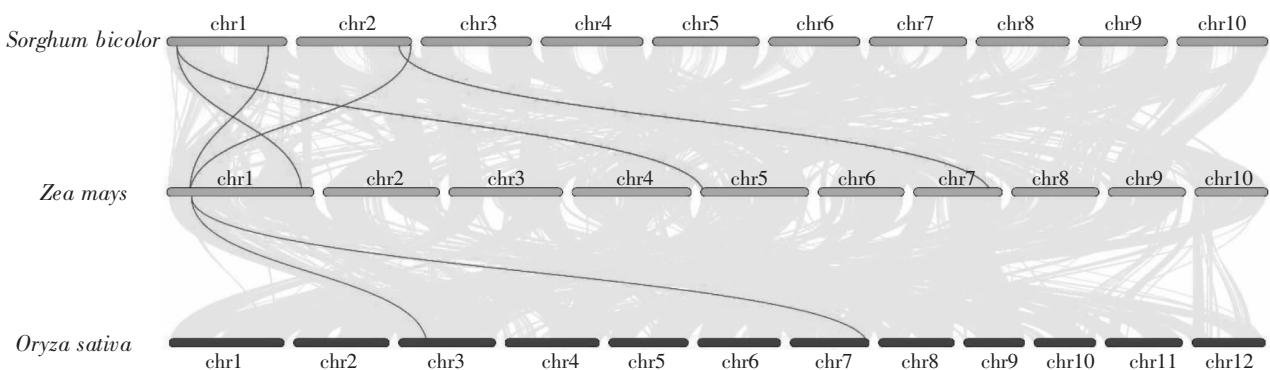


图 6 玉米与其他物种的共线性分析

内同时存在,但不存在于双子叶植物间,可能是该基因形成于物种分化之后。玉米和高粱内各有1个基因在种间形成2个基因对,说明玉米和高粱具有最大的相似遗传关系,和拟南芥相比,水稻和玉米间的共线性关系也较高。任昂彦等^[8]发现,通过分析谷子(*Setaria italica* L.)与玉米、水稻和拟南芥的蛋白进化关系,认为谷子与玉米相似性最高,其次为水稻。这说明禾本科作物之间具有更近的亲缘关系^[13]。

根据对从公共数据库获得的 *ZmCOBL* 家族成员的表达数据分析,可以看出这些基因对非生物胁迫都表现出不同程度的响应。*ZmCOBL* 家族成员是参与植物纤维素合成的重要基因,其表达量的高低直接关系到玉米成熟时的次生细胞壁发育和组织弹性的指标^[14],因此可推测当玉米受到紫外处理时,其纤维素的相关合成途径会受到促进;而当玉米受到高温、盐害等非生物胁迫时,其木质素含量和次生细胞壁发育会受影响,从而导致玉米茎秆强度下降等现象。

本研究通过对玉米 *COBRA* 家族成员进行了鉴定和分析,阐明了该家族成员在序列特征、亚细胞定位、与其他物种的进化关系和应对非生物胁迫时的表达模式,系统的展现了该家族成员的特征,为将来利用分子设计育种手段提高玉米茎秆强度和应对非生物胁迫提供了思路和理论基础^[15]。

参考文献:

- [1] 李方圆. 玉米脆性基因主效位点挖掘及连锁分子标记的开发[D]. 泰安: 山东农业大学, 2020.doi: 10.27277/d.cnki.gsdnu.2020.000985.
- [2] LI P, LIU Y R, TAN W Q, et al. *Brittle Culm 1* encodes a COBRA-like protein involved in secondary cell wall cellulose biosynthesis in sorghum [J]. *Plant & Cell Physiology*, 2018, 60(4): 788-801.
- [3] BENFEY P N, LINSTED P J, ROBERTS K. Root development in Arabidopsis: four mutants with dramatically altered root morphogenesis[J]. *Development*, 1993 119(1): 57-70.
- [4] CHING A, DHUGGA K S, APPENZELLER L, et al. *Brittle stalk 2* encodes a putative glycosylphosphatidylinositol-anchored protein that affects mechanical strength of maize tissues by altering the composition and structure of secondary cell walls[J]. *Planta*, 2006, 224(5): 1174-1184.
- [5] SINDHU A, LANGEWISCH T, OLEK A, et al. Maize *Brittle stalk2* encodes a COBRA-like protein expressed in early organ development but required for tissue flexibility at maturity[J]. *Plant Physiology*, 2007, 145(4): 1444-1459.
- [6] DAI X, YOU C, CHEN G, et al. *OsBCIL4* encodes a COBRA-like protein that affects cellulose synthesis in rice[J]. *Plant Molecular Biology*, 2011, 75(4): 333-345.
- [7] 元志成, 柯余生, 吴富进, 等. 高粱 *COBRA* 基因家族全基因组的鉴定和表达分析[J]. *河南农业科学*, 2020, 49(10): 33-41.
- [8] 任昂彦, 陈环宇, 王根平, 等. 谷子 *COBRA* 家族成员的全基因组鉴定及表达模式分析[J]. *分子植物育种*, 2021, 10(2): 1-10.
- [9] CHEN C J, CHEN H, ZHANG Y, et al. TBtools: An integrative toolkit developed for interactive analyses of big biological data[J]. *Molecular Plant*, 2020, 13(8): 1194-1202.
- [10] 齐燕妮, 李闻娟, 王利民, 等. 亚麻 *COBRA* 基因家族的鉴定与生物信息学分析[J]. *甘肃农业科技*, 2019(9): 33-38.
- [11] WATERS A J, MAKAREVITCH I, NOSHAY J, et al. Natural variation for gene expression responses to abiotic stress in maize[J]. *The Plant Journal*, 2017, 89(4): 706-717.
- [12] 崔永耀. 毛果杨 *COBRA* 基因家族及 *COBRA3/11* 功能分析[D]. 哈尔滨: 东北林业大学, 2018.
- [13] 姜玉梅, 姜如云, 李俊畅, 等. 小麦 *BESI* 基因家族的比较基因组学分析[J]. *麦类作物学报*, 2021, 41(10): 1181-1188.
- [14] JULIUS B T, MCCUBBIN T J, MERTZ R A, et al. Maize *Brittle Stalk2-Like3*, encoding a COBRA protein, functions in cell wall formation and carbohydrate partitioning[J]. *The Plant Cell*, 2021, 33(10): 3348-3366.
- [15] 刘忠祥, 何海军, 石晓娟, 等. 玉米叶夹角突变体 FU1603 的选育及遗传分析[J]. *甘肃农业科技*, 2019(12): 1-4.