

高粱耐冷基因全基因组鉴定及进化分析

张国琴, 葛玉彬, 张正英

(甘肃省农业科学院作物研究所, 甘肃 兰州 730070)

摘要: 利用公布的高粱全基因组数据, 以及 BLAST、Pfam、MEGA6.0 等生物信息学软件对高粱相关耐冷基因数量、系统进化树进行分析。结果表明, 在高粱全基因组中共有 AP2 家族基因 203 个、*ICE1* 基因 252 个、*EIN3* 基因 15 个、*SNAC2* 基因 133 个、*AMP1* 基因 3 个、*COLD1* 基因 3 个、*qLGT3-1* 基因 53 个、*ctb1* 基因 90 个。通过进化树分析, 高粱中 XP002457016.1、XP021-317625.1; XP002463411.2、XP002467935.1、XP021317527.1; XP002447111.1、XP021-315473.1、XP002447110.1、XP021318475.1、XP021318476.1; XP002449576.1、XP0024 59022.1; XP002466328.1; XP021317902.1; XP002446618.1、XP002452340.1 和 XP02 1314112.1 基因与模式作物中 AP2 基因家族、*AtEIN3*、*Oscb1*、*OsqLGT3-1*、*AtAMP1*、*OsCOLD1*、*OsSNAC2* 亲缘关系最近, 主要分布在高粱 1、2、3、4、5、6 号染色体。综合分析, 在高粱全基因组中耐冷转录因子 (*AP2*、*EIN3*、*ICE1*、*SNAC2*) 拷贝数较多, 与模式作物耐冷基因亲缘关系最近的基因主要分布在高粱 1、3、5 号染色体, 耐冷功能基因 (*COLD1*、*qLGT3-1*、*ctb1*、*AMP1*) 拷贝数较少, 与模式作物耐冷基因亲缘关系最近的基因主要分布在 1、4、6 号染色体。

关键词: 高粱; 耐冷基因; 全基因组鉴定; 进化分析

中图分类号: S514 **文献标志码:** A **文章编号:** 1001-1463(2020)09-0037-05

doi: 10.3969/j.issn.1001-1463.2020.12.010

Genome-wide Identification and Evolutionary Analysis of Cold Tolerance Genes in *Sorghum bicolor*

ZHANG Guoqin, GE Yubin, ZHANG Zhengying

(Institute of Crops, Gansu Academy of Agricultural Sciences, Lanzhou Gansu 730070, China)

Abstract: In this study, The whole sorghum genome data and bioinformatics software which BLAST, MEGA and Pfam were used to carry out the genome-wide analysis of sorghum (*Sorghum bicolor*) cold-tolerance genes, including gene number, phylogenetic analysis, of sorghum cold-tolerance gene. The results show that we have found 203 AP2 genes, 252 *ICE1* genes, 15 *EIN3* genes, 133 *SNAC2* genes, 3 *AMP1* genes, 3 *COLD1* genes, 53 *qLGT3-1* genes, 90 *ctb1* genes in sorghum genome; According to phylogenetic analysis, these XP 002457016.1, XP 021317625.1; XP 002463411.2, XP 002467935.1, XP 021317527.1; XP 002447111.1, XP 021315473.1, XP 002447110.1, XP 021318475.1, XP 021318476.1; XP 002449576.1, XP 002459-022.1; XP002466328.1; XP021317902.1; XP002446618.1, XP002452340.1 and XP021314112.1 cold-tolerance genes genetic relationship are closest to AP2 gene family, *AtEIN3*, *Oscb1*, *OsqLGT3-1*, *AtAMP1*, *OsCOLD1* and *OsSNAC2* in model crops, this are mainly distributed in chromosome 1, 2, 3, 4, 5 and 6 of sorghum. Comprehensive analysis showed that there were cold tolerance transcription factor (*AP2*, *EIN3*, *ICE1*, *SNAC2*) copy number is more in sorghum genome, which genetic relationship are closest to the model crop cold tolerance genes are mainly distributed in chromosome 1, 3, 5, cold tolerance functional genes (*COLD1*、

收稿日期: 2020-06-24

基金项目: 甘肃省农业科学院院地科技合作项目(2017GAAS59); 甘肃省农业科学院科技支撑计划项目(2017GAAS44); 甘肃省农业科学院生物技术育种专项(2019GAAS07)。

作者简介: 张国琴(1987—), 女, 甘肃靖远人, 助理研究员, 硕士, 主要从事高粱分子育种工作。Email: zhanggq0828@126.com。

通信作者: 葛玉彬(1979—), 男, 甘肃靖远人, 助理研究员, 主要从事高粱遗传育种工作。Email: gyb1996@126.com。

qLTG3-1、*ctb1*、*AMPI*) copy number is less, which genetic relationship are closest to the model crop cold tolerance genes are mainly distributed in chrom -osome 1, 4, and 6.

Key words: Sorghum; Cold-tolerance genes; Genome-wide identification; Evolution analysis

高粱 [*Sorghum bicolor* (L.) Moench] 是一种喜温 C₄ 作物, 具有生物产量高、抗旱、耐涝、耐盐碱等特性, 是干旱和半干旱地区重要的禾谷类作物。高粱春季播种后突遇低温时常常使种子在土壤中“粉种”或霉烂, 导致萌发率、出苗率降低, 出苗后植株长势较弱且生长缓慢, 从而减少作物产量^[1-2]。在高纬度温带地区, 提早播期可延长高粱生长周期, 增加生物产量, 但早播遇到冷害的胁迫, 种子不能发芽或者表现出较差的幼苗活力, 限制高粱早期生长, 降低生物量和产量, 从而限制高粱在全球更多地区的推广利用^[3-4]。

许多植物通过低温驯化增加耐冷性。关于在低温胁迫下参与低温调节和信号转导的基因功能及特性的报道颇多。植物通过调节各种基因的表达来适应低温胁迫, 根据功能不同可将耐冷功能基因分为两大类。一类是功能性基因, 即合成膜脂相改变的基因, 如不饱和脂肪酸的基因、提高细胞膜流动性基因等; 另一类基因则参与整个信号传导过程, 在整个生理生化系统中起着调控作用, 在模式作物拟南芥中转录因子 CBF/DREB1 蛋白家族和 *COR* 基因调控低温诱导基因的表达, 从而增加植物的耐冷性^[5-6]。已鉴定出一些直接调控 *CBF* 基因的转录因子, 例如激活型转录因子螺旋结构转录因子 (*ICE1*) 和钙结合转录激活子 (CAMTAs), 以及抑制性型转录因子 *MYB15*、*EIN3* 和 *PIF3/4/7*^[7-9]; 在水稻中已克隆的耐冷基因有 *Os-COLD1*、*OsSNAC2*、*OsDREB1B*、*OsqLTG3-1* 等^[10-12]。我们通过 NCBI、国家水稻数据中心等数据库, 对模式作物拟南芥和水稻中的耐冷关键基因进行搜集, 并在高粱基因组中进行同源性比较, 旨在为高粱耐冷候选基因的挑取提供参考。

1 材料与方法

1.1 供试材料数据来源

利用 NCBI 数据库、模式植物数据库 (www.modelcrop.org) 搜索已报道的拟南芥耐冷基因, 从国家水稻数据中心数据库搜索水稻中已克隆耐冷基因和转录因子。登录高粱全基因组蛋白数据库 (<https://www.ncbi.nlm.nih.gov/genome>), 从该网站下载所需的 DNA 和蛋白质数据。

1.2 高粱耐冷基因的鉴定

利用 pfam 数据库 (<http://pfam.janelia.org/search/sequence>) 中的隐马尔可夫模型 (HMM) 拟南芥、水稻耐冷基因结构域的氨基酸序列与高粱全基因组蛋白数据库进行 BLASTp 序列比对, 阈值小于等于 1E-5, 初步筛选出候选基因序列, 得到候选序列后用 Pfam 对结构域进行确认^[13]。

1.3 创建高粱 *ARF* 基因家族进化树

利用 MEGA6.0 软件中的 ClustalW, 采用默认参数, 对从高粱全基因组中搜索到的耐冷基因和模式作物耐冷基因蛋白的氨基酸序列进行多重比对, 比对结果采用 NJ (Neighbor-Joining) 法构建系统进化树, Bootstrap 校验值为 1 000 次。

2 结果与分析

2.1 模式作物耐冷基因的搜集

利用 NCBI 数据库、模式植物数据库 (www.modelcrop.org) 对已报道的拟南芥耐冷基因进行搜索, 结果如表 1 所示, 搜集到的拟南芥主要耐冷基因有 *AtCBF1* (AT4G25470)、*AtCBF2* (AT4G25480)、*AtCBF3* (AT4G25490)、*AtICE1* (AT1G12860)、*AtEIN3* (AT3G20770)、*AtAMPI* (AT3G54720) 等, 分别编码 C-端脱水效应结合因子蛋白、碱性螺旋-环-螺旋 DNA 结合蛋白、乙烯不敏感蛋白、肽酶 M28 家族蛋白。从国家水稻数据中心数据库

搜索水稻中主要耐冷基因, 搜索到 *OsCOLD1* (Os04g51180)、*OsSNAC2* (Os04g0460600)、*OsqLTG3-1* (Os03g0103300)、*Osctb1* (Os04g0619300)、*OsDREB1B* (Os09g0522000)、*OsDREB1F* (Os01g0968800)等, 分别编码鸟嘌呤核苷酸结合蛋白、无顶端分生组织蛋白、未知蛋白、环氧氯丙烷相关蛋白、脱水效应元件结合蛋白。

2.2 高粱耐冷基因鉴定

通过 BLASTp 模式作物耐冷基因与高粱全基因组蛋白数据库进行搜索, 通过 Pfam 对结构域进行确认, 结果如表 1 所示。鉴定出高粱的 AP2 家族(*CBF1*、*CBF2*、*CBF3*、*DREB1B*、*DREB1F*)基因有 203 个、*ICE1* 基因有 252 个、*EIN3* 基因有 15 个、*SNAC2* 基因有 133 个、*AMPI* 基因有 3 个、*COLD1* 基因有 3 个、*qLTG3-1* 基因有 53 个、*ctb1* 基因有 90 个。从结果中得出高粱全基因组中耐冷转录因子基因的拷贝数除 *EIN3* 乙烯信号通路中转录因子拷贝数少(13 个), 其余的拷贝数较多(AP2 家族 203 个、*ICE1* 252 个、*SNAC2* 133 个); 耐冷功能型基因拷贝数较少(*COLD1* 3 个、*qLTG3-1* 53 个、*ctb1* 90 个、*AMPI* 3 个), 这有助于为高粱耐冷候选基因的筛选提供依据。

2.3 高粱耐冷基因的进化研究

将鉴定出的高粱耐冷基因与模式作物相关基因利用 MEGA6.0 软件构建系统进化树, 统计结果如表 2 所示, 高粱全基因组的 *EIN3* 基因家族被聚为五大类, 其中 XP 002463411.2、XP 002467935.1、XP 021317527.1 与拟南芥中 *AtEIN3* 亲缘关系近, 聚为一类, 分布在高粱 1 号和 2 号染色体。高粱全基因组的 AP2 家族被聚为十五大类, 其中 XP 002457016.1、XP 021317625.1 与拟南芥和水稻中 AP2 家族(*AtCBF1*、*AtCBF2*、*AtCBF3*、*OsDREB1B*、*OsDREB1F*)亲缘关系近, 聚为一类, 分布在高粱 3 号和 5 号染色体。高粱全基因组的 *ctb1* 基因家族被聚为十三大类, 其中 XP 002447111.1、XP 021315473.1、XP 002447110.1、XP 021318475.1、XP 021318476.1 与水稻 *Osctb1* 亲缘关系近, 被聚为一类, 分布在 4 号和 6 号染色体。高粱全基因组 ICE 家族被分为十大类, 其中 XP 002449576.1、XP 002459022.1 与水稻中 *AtICE1* 亲缘关系近, 聚为一类, 分布在高粱 3 号和 5 号染色体。高粱全基因组中的 *LTG3-1* 基因被分为十三大类, 其中 XP021312137.1 与水稻中 *OsqLTG3-1* 亲缘关系最近, 聚为一类, 分布在 1 号染色体。高

表 1 模式作物耐冷基因及在高粱全基因组中拷贝数

模式作物耐冷基因	ID号	编码蛋白	高粱中拷贝数 /个
AP2家族(<i>AtCBF1</i> 、 <i>AtCBF2</i> 、 <i>AtCBF3</i> 、 <i>OsDREB1B</i> 、 <i>OsDREB1F</i>)	AT4G25470 AT4G25480 AT4G25490 Os09g0522000 Os01g0968800	C-端脱水效应结合因子 脱水效应元件 结合蛋白	AP2家族(<i>CBF1</i> 、 <i>CBF2</i> 、 <i>CBF3</i> 、 <i>DREB1B</i> 、 <i>DREB1F</i>) 203
<i>AtICE1</i>	AT1G12860	碱性螺旋-环-螺旋DNA结合蛋白	<i>ICE1</i> 252
<i>AtEIN3</i>	AT3G20770	乙烯不敏感蛋白	<i>EIN3</i> 15
<i>OsSNAC2</i>	Os04g0460600	无顶端分生组织蛋白	<i>SNAC2</i> 133
<i>AtAMPI</i>	AT3G54720	肽酶M28家族蛋白	<i>AMPI</i> 3
<i>OsCOLD1</i>	Os04g51180	鸟嘌呤核苷酸结合蛋白	<i>COLD1</i> 3
<i>OsqLTG3-1</i>	Os03g0103300	未知蛋白	<i>qLTG3-1</i> 53
<i>Osctb1</i>	Os04g0619300	环氧氯丙烷相关蛋白	<i>ctb1</i> 90

表 2 高粱中与模式作物耐冷基因系统进化结果统计分析

模式作物耐冷基因名称	高粱全基因组耐冷基因系统进化被聚类别数	与模式作物耐冷基因亲缘关系近的基因	染色体位置
<i>AtEIN3</i>	5	XP 002463411.2 XP 002467935.1 XP 021317527.1	2号染色体 1号染色体 1号染色体
<i>AP2</i> (<i>AtCBF1</i> 、 <i>AtCBF2</i> 、 <i>AtCBF3</i> 、 <i>OsDREB1B</i> 、 <i>IF</i>)	15	XP 002457016.1 XP 021317625.1	3号染色体 5号染色体
<i>Osctb 1</i>	13	XP 002447111.1 XP 021315473.1 XP 002447110.1 XP 021318475.1 XP 021318476.1	6号染色体 4号染色体 6号染色体 6号染色体 6号染色体
<i>AtICE 1</i>	10	XP 002449576.1 XP 002459022.1	5号染色体 3号染色体
<i>OsqLTG3-1</i>	13	XP021312137.1	1号染色体
<i>AtAMPI</i>	2	XP002466328.1	1号染色体
<i>OsCOLD1</i>	2	XP021317902.1	6号染色体
<i>OsSNAC2</i>	11	XP002446618.1 XP002452340.1 XP021314112.1	6号染色体 4号染色体 3号染色体

高粱全基因组中 *AMPI* 被分为 2 类，其中 XP002466328.1 与拟南芥 *AtAMPI* 亲缘关系最近，聚为一类，分布在 1 号染色体。高粱全基因组中 *COLD1* 基因被分为二类，其中 XP021317902.1 与 *OsCOLD1* 亲缘关系最近，聚为一类，分布在 6 号染色体。高粱全基因组中 *SNAC2* 基因家族被分为十一大类，其中 XP002446618.1、XP002452340.1、XP021314112.1 与 *OsSNAC2* 亲缘关系最近，聚为一类，分布在 3 号、4 号和 6 号染色体。分析以上结果，高粱中的与模式作物耐冷基因亲缘关系近的耐冷基因主要分布在高粱 1、2、3、4、5、6 号染色体上，其中分布在 1 号和 6 号染色体上的最多，从而为找到高粱耐冷苗头基因提供依据。

3 小结和讨论

对已报道模式作物中部分耐冷基因及转录因子，通过对高粱全基因组进行扫描，得到高粱中的 *AP2* (*AtCBF1*、*AtCBF2*、*AtCBF3*、*OsDREB1B*、*OsDREB1F*) 基因有 203 个，*ICE1* 基因有 252 个，*EIN3* 基因有 15 个，*SNAC2* 基因有 133 个，*AMPI* 基因有 3 个，*COLD1* 基因有 3 个，*qLTG3-1* 基因有 53 个，*ctb1* 基因有 90 个。通过进化树分析，得出在高粱中 XP 002457016.1 和 XP

021317625.1 与模式作物 *AP2* 基因家族亲缘关系最近，分布在 3 号和 5 号染色体；XP 002463411.2、XP 002467935.1、XP 021317527.1 与 *AtEIN3* 亲缘关系最近，分布在 1 号和 2 号染色体；XP 002447111.1、XP 021315473.1、XP 002447110.1、XP 021318475.1 和 XP 021318476.1 与 *Osctb1* 亲缘关系最近，分布在 4 号和 6 号染色体；XP 002449576.1 和 XP 002459022.1 与 *Osctb1* 亲缘关系最近，分布在 3 号和 5 号染色体；XP021312137.1 与 *OsqLTG3-1* 亲缘关系最近，分布在 1 号染色体；XP002466328.1 与 *AtAMPI* 亲缘关系最近，分布在 1 号染色体；XP021317902.1 与 *OsCOLD1* 亲缘关系最近，分布在 6 号染色体；XP002446618.1、XP002452340.1、XP021314112.1 与 *OsSNAC2* 亲缘关系最近，分布在 3 号、4 号和 6 号染色体。

目前高粱耐冷研究主要涉及 QTL 初步定位。但已定位 QTL 效应值足够大的较少，还需要进一步研究验证这些 QTL 的效应，还不能直接用于高粱早期耐低温分子标记辅助选择育种^[3-4, 14-15]，有必要通过其他方法加快高粱耐冷性研究，例如 RNA-seq 等，近年来，RNA-seq 技术广泛应用于植物应答低温胁迫，水稻中，通过 RNA-seq 对耐冷

型水稻转录组分析鉴定出 121 个低温诱导基因, 并且发现一个活性氧(ROS)-bZIP1 调节子对早期低温胁迫响应起着重要作用^[16]; 利用 RNA-seq 技术对 2 个甜玉米自交系在低温胁迫进行转录组分析, 分别鉴定出 357 和 455 差异表达基因, 鉴定出与冷害响应共同差异表达基因 94 个, 进一步分析发现 74 个差异表达转录因子, 其中大多数参与新陈代谢和激素调节子^[17]。我们对模式作物已报到的耐冷基因通过对高粱全基因组分析, 得出在高粱全基因组中耐冷基因的苗头基因及大概所在染色体分布, 从而为高粱耐冷快速研究方法(RNA-seq、GWAS 等)分析得到差异表达基因获得目标耐冷候选基因提供依据。

参考文献:

- [1] YU J, TUINSTRA M, CLAASSEN M M, et al. Analysis of cold tolerance in sorghum under controlled environment conditions[J]. Field Crops Research, 2004, 85: 21-30.
- [2] 成慧娟, 严福忠, 马尚耀, 等. 影响高粱种子出苗率下降的原因及预防措施[J]. 内蒙古农业科技, 2011(4): 112.
- [3] BUROW G, BURKE J, XIN Z G, et al. Genetic dissection of early-season cold tolerance in sorghum (*Sorghum bicolor* L.) Moench [J]. Molecular Breeding, 2011, 28: 391-402.
- [4] KNOLL J, GUNARATNA N F, EJETA G. QTL analysis of early-season cold tolerance in sorghum[J]. Theoretical and applied genetics, 2008, 116(4): 577-587.
- [5] AHAMED A, MURAI-HATANO M, JUNKO I S, et al. Cold stress-induced acclimation in rice is mediated by root specific aquaporins[J]. Plant Cell Physiology, 2012, 53: 1445-1456.
- [6] LATA C, PRASAD M. Role of DREBs in regulation of abiotic stress responses in plants[J]. Journal Experimental Botany, 2011, 60: 121-135.
- [7] JIANG B, SHI Y, ZHANG X Y, et al. PIF3 is a negative regulator of the CBF pathway and freezing tolerance in Arabidopsis[J]. PNAS, 2017, 114: 6695-6702.
- [8] CHINNUSAMY V, OHTA M, KANRAR S, et al. ICE1: a regulator of cold-induced transcriptome and freezing tolerance in Arabidopsis [J]. Genes. Dev., 2003, 17: 1043-1054.
- [9] CHIN-MEI LEE, MICHAEL F, THOMAS H. Photoperiodic regulation of the C-repeat binding factor (CBF) cold acclimation pathway and freezing tolerance in Arabidopsis thaliana [J]. PNAS, 2012, 109(37): 15054-15059.
- [10] MA Y, DAI X, LIN X, et al. COLD1 confers chilling tolerance in rice [J]. Cell, 2015, 160(6): 1209-1221.
- [11] KENJI FUJINO, HIROSHI SEKIGUCHI, YASUYUKI MATSUDA, et al. Molecular identification of a major quantitative trait locus, qLTG3-1, controlling low-temperature germinability in rice [J]. PNAS, 2008, 105(34): 12623-12628.
- [12] DUBOUZET J G, SAKUMA Y, YUSUKE ITO, et al. DREB genes in rice, encode transcription activators that function in drought high: altand cold-responsive gene expression [J]. Plant Journal, 2003, 33: 751-763.
- [13] 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[J]. 2019(9): 33-36.
- [14] KNOLL J, EJETA G. Marker-assisted selection for early-season cold tolerance in sorghum: QTL validation across populations and environments [J]. Theoretical and Applied Genetics. 2008, 116: 541-53.
- [15] 张国琴, 葛玉彬, 张正英, 等. 高粱抗旱研究综述[J]. 甘肃农业科技, 2018(6): 67-72.
- [16] WANG D Z, JIN Y N, DING X H, et al. Genetic regulation and signal transduction in the ICE CBF COR signaling pathway during cold stress in plants [J]. Biochemistry Moscow, 2017, 82(10): 1103-1117.
- [17] MAO J H, YU Y T, YANG J, et al. Comparative transcriptome analysis of sweet corn seedlings under low-temperature stress [J]. Crop Journal, 2017, 5: 396-406.

(本文责编: 陈伟)