



曹士亮,于滔,董玲,等.基于 SLAF-Super BSA 技术的玉米萌发期耐冷性定位的初步研究[J].黑龙江农业科学,2021(2):1-5.

基于 SLAF-Super BSA 技术的玉米萌发期耐冷性定位的初步研究

曹士亮¹,于滔¹,董玲²,刘宝民¹,王成波¹,李文跃¹,杨耿斌¹,武尔娜¹

(1.黑龙江省农业科学院 玉米研究所/农业农村部东北北部玉米生物学与遗传育种重点实验室/黑龙江省玉米种质资源创制与利用重点实验室,黑龙江 哈尔滨 150086;2.东北农业大学农学院,黑龙江 哈尔滨 150030)

摘要:为加快玉米萌发期耐冷性遗传育种研究,在对玉米重组自交系群体进行萌发期耐冷性鉴定的基础上,采用分离群体混合分析(Bulked Segregant Analysis,BSA)方法,对2亲本和2混池进行特异位点扩增片段测序(Specific-locus Amplified Fragment-sequencing,SLAF),初步定位了玉米萌发期耐冷性相关的染色体区段,并对关联区域内的基因进行初步富集分析。结果表明:研究获得了282 563个SLAF标签,每个SLAF标签平均父本测序深度为27.06×,母本平均测序深度为31.85×,耐冷池和冷敏感池的测序深度分别为36.72×和34.17×。获得多态性SLAF标签47 624个,多态性比例为16.85%。通过BSA关联分析在玉米第9号染色体上定位到了一个91.06 Mb大小的玉米萌发期耐冷性关联区域,区域内共关联到3 170个基因。

关键词:玉米;萌发期耐冷性;SLAF;BSA

萌发期冷害导致玉米萌发及出苗缓慢,易造成种子霉变腐烂,给玉米生产带来严重损失。多数研究表明,玉米萌发期耐冷性属于多基因控制的数量性状。研究表明低温下玉米发芽能力的上位性、加性和显性基因效应显著^[1]。马延华^[2]利用156个DH系和105个SSR标记构建玉米遗传连锁图,共检测到5个与玉米发芽期耐冷性有关的QTL,分别位于玉米第2,4,6及10染色体上,这些QTL可以解释表型变异的5.45%~12.85%。Strigens等^[3]采用全基因组关联分析的方法,定位了19个与耐冷相关的分子标记,可以解释玉米苗期生长阶段和叶绿素荧光参数5.7%~52.5%的表型遗传变异,同时发现了多个耐冷相关候选基因。玉米萌发期耐冷性不仅遗传基础复杂,还受到表型鉴定方法、遗传分析方法等因素影响,研究结果不尽相同。基于简化测序的Super-BSA技术,将二代测序(Next Generation Sequencing)的高通量和分离混池分析(Bulked Segregant Analysis,BSA)的简化样本数的优势

相结合,近年来,在玉米果皮纤维素含量相关基因定位^[4]、大豆酸性磷酸酶含量候选基因定位及相关分子标记开发^[5]等方面得到了广泛的应用。本研究在玉米萌发期耐冷性鉴定的基础上,采用特异位点扩增片段测序(Specific-locus Amplified Fragment-sequencing,SLAF)的分离混池分析(Bulked Segregant Analysis,BSA)方法,定位玉米萌发期耐冷性相关的染色体关联区域,并初步对关联区域内的基因进行富集分析,期望为玉米萌发期耐冷性遗传育种研究提供一定的理论依据和技术参考。

1 材料与方法

1.1 材料

试验材料为来自于玉米自交系B73和Mo17杂交后代的重组自交系,通过对这些重组自交系进行萌发期耐冷性鉴定,采用5℃低温处理4d后转入25℃培养3d的试验条件,筛选出耐冷和冷敏感极端材料各30份。

1.2 方法

1.2.1 SALF文库构建及高通量测序 对玉米重组自交系耐冷性存在差异的极端表型单株分别取等量的叶片,混合得到耐冷混池(30份)和冷敏感混池(30份),提取亲本自交系B73、Mo17和两个混池的DNA。选择限制性内切酶HaeⅢ对样品进行酶切,分别回收414~444 bp和464~

收稿日期:2020-10-22

基金项目:黑龙江省农业科学院“农业科技创新跨越工程”(HNK2019CX22-13);黑龙江省农业科学院杰出青年基金项目(2019JCQN004);国家博士后项目(2017M621318);黑龙江省博士后项目(LBH-Z17205);黑龙江省玉米超高产田建设(2017YFD0300506-3);国家自然科学基金项目(31771814)。

第一作者:曹士亮(1980—),男,博士,副研究员,从事玉米分子育种研究。E-mail:caoshiliang2003@126.com。

494 bp大小的酶切片段构建测序文库,文库质检合格后采用 Illumina HiSeqTM 2500 进行测序。以水稻(*Oryza sativa*)测序数据为对照,评估监控试验过程是否正常,确定酶切方案的有效性。

1.2.2 SLAF 标签及 SNP 标记开发 对不同亲本及抗感池材料进行测序,根据产生 reads 的相似性进行聚类,聚类到一起的 reads 被称为一个 SLAF 片段(SLAF 标签)。不同样本间存在差异的 SLAF 标签被定义为多态性 SLAF 标签。接下来对玉米各染色体上 SLAF 标签及其分布进行分析。采用 GATK 软件和 samtools 软件进行 SNP 的检测。为了保证 SNP 检测的准确性,取两个软件检测的交集得到最终的 SNP 位点集。

1.2.3 BAS 关联分析 BAS 关联分析参考杜龙岗等^[4]的研究。利用获得的高质量 SNP 分子标记识别萌发期耐冷性混池间存在差异的位点。再根据欧氏距离(Euclidean Distance,ED)确定染色体上与目标性状相关的区域。其原理是利用混池间差异 SNP 频率的差异,计算两混池间 SNP 的 ED 值^[6],如果标记与目标性状存在较大关联,关联区域附近的 SNP 位点会倾向于在两混池间存在较大的差异,因此目标区域内的 ED 值较大,可通过作图在目标关联区域附近会出现较明显的峰。对获得的原始 ED 值采用 5 次方处理以消除背景噪音干扰。采用 LOESS 算法拟合回归曲线^[6],并确定最终关联区域。

1.2.4 关联基因注释 使用 BLAST 软件将关联区域内的基因分别与 GO、COG、KEGG 数据库比对,获得区域内基因的注释信息。

2 结果与分析

2.1 RILs 分离群体的耐冷性混池的建立与 DNA 准备

采用 5℃低温处理 4 d 后转入 25℃培养 3 d 的试验条件对 280 份 B73×Mo17 群体和两个亲本进行了室内萌发耐冷性鉴定,耐冷性材料 B73 的发芽率为 92.5%,冷敏感自交系 Mo17 的发芽率为 4.5%,30 份抗池材料平均发芽率为 93.5%,30 份冷敏感材料平均发芽率为 6.75%,确定玉米萌发期耐冷性存在差异的两个混池。

2.2 SLAF 文库的构建及评价与标记开发

2.2.1 SLAF 文库的构建及评价 对玉米参考基因组序列进行电子酶切预测,确定限制性内切酶为 *Hae* III,酶切片段长度在 414~444 bp 及 464~494 bp 的序列定义为 SLAF 标签,分别预

测可得到 100 573 和 108 130 个 SLAF 标签,SLAF 标签基本均匀分布在基因组上,两种酶酶切后,分别由 12.63%和 14.49%的 SLAF 标签位于重复序列区域内,大部分落在非重复区域。能满足基因定位和进一步开发分子标记的要求。本研究中,用于质控的对照为水稻(*Oryza sativa*),其基因组大小为 374.31 Mb,下载地址:<http://rice.plantbiology.msu.edu>。对照的测序 reads 与其参考基因组比对结果显示比对效率基本正常,双端比对效率为 83.94%,且对照的酶切效率为 95.85%,表明酶切反应也比较正常。

对亲本 B73 和 Mo17、耐冷性极端混池共 4 个样品进行酶切和测序。共获得 45.01 Mb reads 数据,测序平均 Q30 为 84.22%,平均 GC 含量为 47.02%。亲本和 2 个混池的 read 数和数据质量如表 1 所示。结果表明,所获得的玉米测序数据符合要求,可以进一步应用于 SLAF 标记开发和关联分析。

表 1 各样品测序数据统计

样品名称	总 Reads 数	Q30 百分数/%	GC 含量/%
Mo17	8975644.00	84.47	47.06
B73	9192310.00	84.43	47.02
耐冷池	13821321.00	84.21	46.93
冷敏感池	13022798.00	83.75	47.07
平均	11253018.25	84.22	47.02
建库对照(水稻)	417130.00	80.22	43.53

2.2.2 SLAF 标签的确定及其在玉米染色体上的分布 通过 SOAP 软件将 SLAF 序列定位到玉米(B73)参考基因组(下载地址:https://www.ncbi.nlm.nih.gov/assembly/GCF_000005005.1/)上,SLAF 标签开发亲本平均测序深度为 29.46×,混池平均测序深度为 35.45×(表 2)。

表 2 SLAF 标签统计

样品名称	SLAF 数目	总深度	平均深度
Mo17	257454.00	6966705.00	27.06×
B73	200339.00	6380797.00	31.85×
耐冷池	279480.00	10262506.00	36.72×
冷敏感池	279140.00	9538214.00	34.17×

将 SLAF 标签定位到参考基因组上,统计不同染色体上的 SLAF 标签和多态性 SLAF 标签,共获得 282 563.00 个 SLAF 标签,多态性标签 47 624.00 个,详见表 3。

表 3 SLAF 标签和多态性 SLAF 标签染色体分布统计

染色体 ID	SLAF 标签数	多态性 SLAF 标签数
chr1	40913.00	6796.00
chr2	33078.00	5505.00
chr3	31454.00	5516.00
chr4	34315.00	5955.00
chr5	29206.00	4251.00
chr6	22233.00	3785.00
chr7	24634.00	4622.00
chr8	23366.00	4011.00
chr9	21152.00	3377.00
chr10	20353.00	3632.00
其他	1859.00	174.00
总计	282563.00	47624.00

根据 SLAF 在染色体上的分布,将 SLAF 标签和多态性 SLAF 标签绘制成图,由图 1 和图 2 可知,每一个黄色条带代表一条染色体,按照 1 Mb 的大小对基因组进行了划分,每个 window 内的 SLAF 标签数越多,颜色越深,SLAF 标签数越少,颜色越浅;图中颜色越深的区域即 SLAF 标签集中分布的区域,多态性 SLAF 标签在玉米 10 条染色体上分布较为均匀。

2.2.3 SNP 信息统计 将获得的亲本及耐冷和冷敏感池 SNP 统计信息绘制成表,由表 4 可以看出,试验获得总 SNP 为 515 008.00 个。耐冷池和冷敏感池 SNP 数目比较接近,为 43 万个以上,并且杂合率分别为 46.42%和 47.02%,说明混池中出现了较明显的分离。

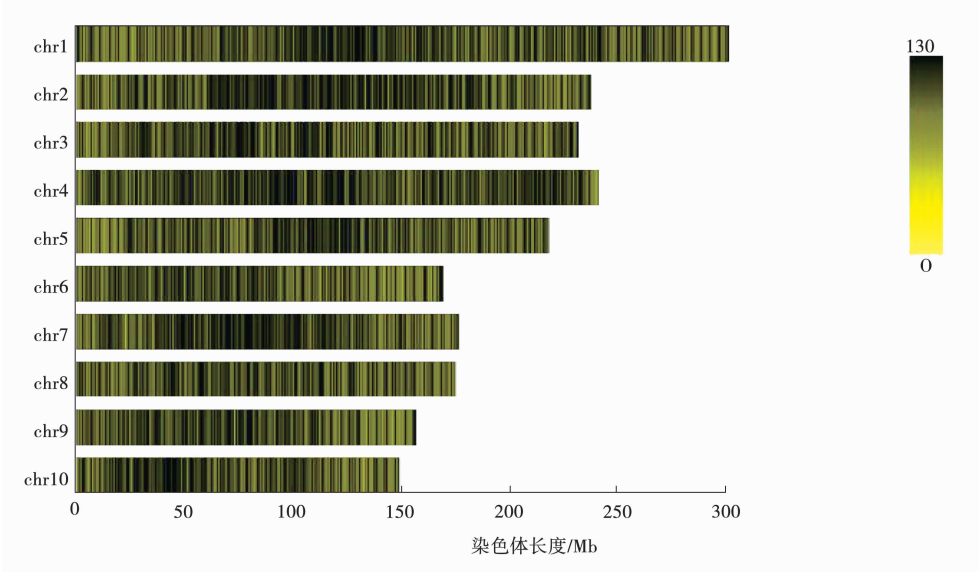


图 1 SLAF 标签的分布

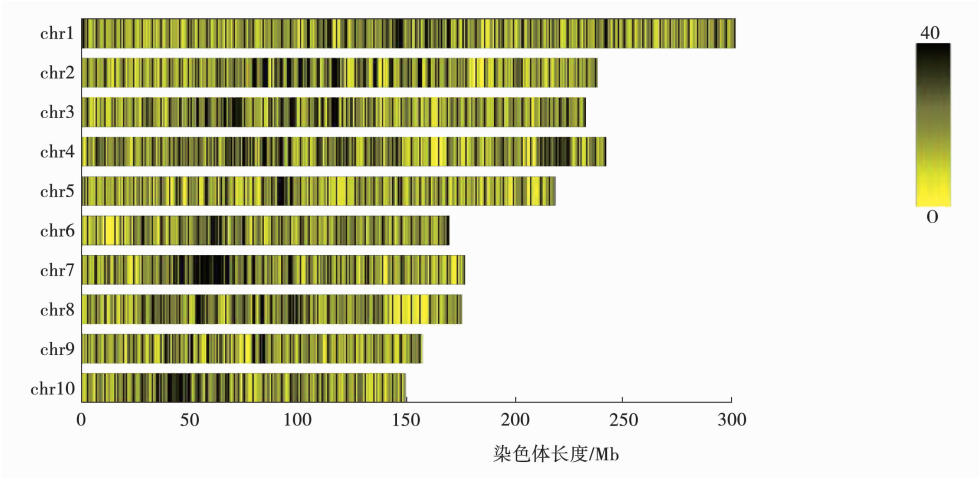


图 2 多态性 SLAF 标签的分布

表 4 样品 SNP 信息统计

样品名称	总 SNP	对应样品 SNP 数目	杂合率/%
Mo17	515008.00	284437.00	3.18
B73	515008.00	425565.00	8.96
耐冷池	515008.00	433233.00	46.42
冷敏感池	515008.00	430509.00	47.02

2.3 关联分析

在关联分析前,首先对任一混池中 reads 支持度小于 3 的 SNP 进行过滤,然后过滤与相应亲本的 SNP 类型差异的位点,最终得到高质量的可信 SNP 位点,并在此基础上识别两混池间差异的

位点。过滤统计及结果见表 5,使用 ED 算法^[6] 计算与目的基因连锁的区域,并通过拟合曲线确定最终关联区域(图 3 和图 4)。由图可知,其中 9 号染色体出现了显著的差异峰。该染色体的结果放大如图 4,取所有位点拟合值的中值加上 3 倍的标准差得到的阈值为 0.083,可得到定位区域为 1 923 152.00~92 980 413.00 bp(91 057 262.00 bp)。

表 5 SNP 过滤及最终位点数统计

Read 支持度 小于 3 的位点	亲本 SNP 类型 差异的位点	混池间一 致的位点	混池间差 异的位点
103388.00	65769.00	168374.00	35057.00

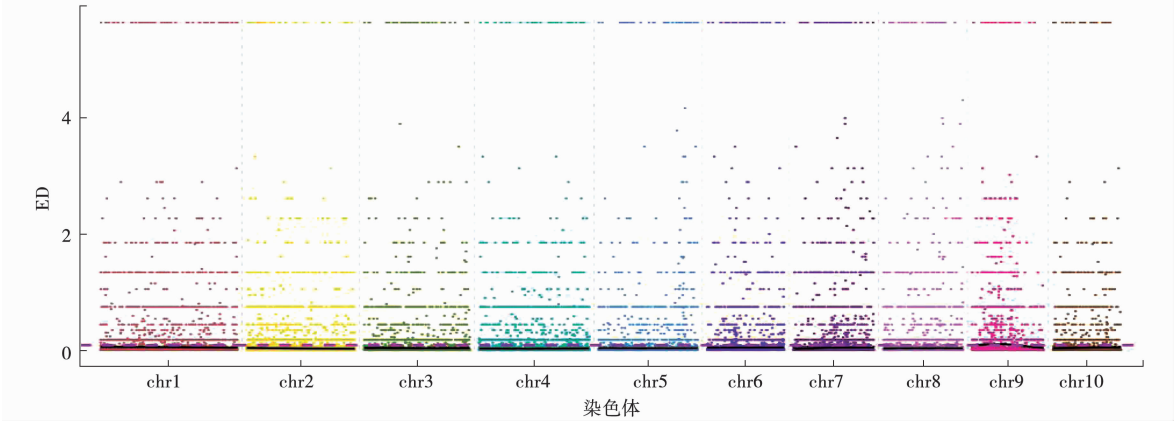


图 3 ED 值经 5 次方以后得到的点以及拟合曲线的分布

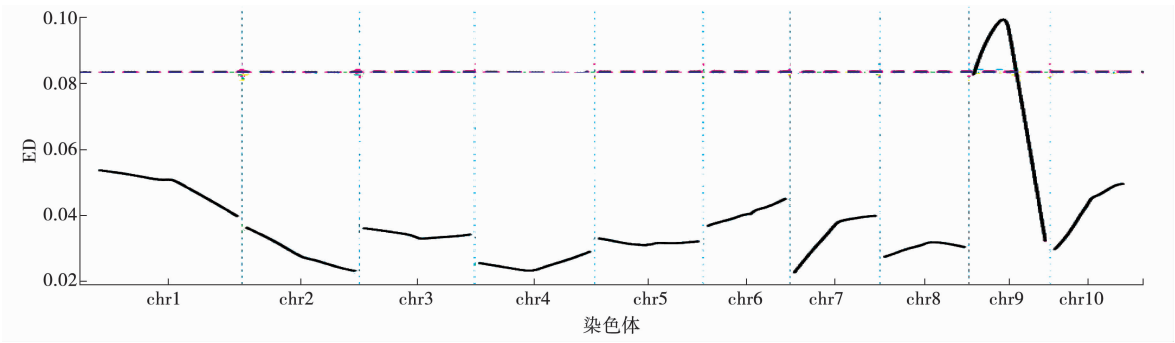


图 4 BSA 分析结果

2.4 关联区域的基因富集分析

通过对关联区域内的基因进行查找,共发现 3 170.00 个基因,对这些基因分别进行 KEGG、GO 和 COG 功能注释分析,关联区域内共有 367.00 个基因注释到 KEGG 通路,1 529.00 个基因注释到 GO Term,679.00 个基因注释到 COG Class。

3 结论与讨论

基因组测序技术的快速发展大大地降低了分子标记开发成本。SLAF-seq 技术是在利用酶切

进行染色体片段选择(SLAF 标签)的基础上进行的基因组测序,属于简化基因组测序技术的一种,具有有效 reads 长、通量高、方案设计灵活等优势。近年来,SLAF-seq 技术被广泛地应用于种质遗传多样性分析^[7]、作物性状相关基因定位^[8] 和分子标记开发^[9] 等方面。本研究选取了 *Hae* III 酶进行酶切,片段长度选择在 414~444 bp 及 464~494 bp 的 SLAF 标签进行建库测序,SLAF 标签开发亲本平均测序深度为 29.46×,混池平均测序深度为 35.45×,共获得 282 563.00 个 SLAF 标签,多态性标签 47 624.00 个。质量控

制分析表明 Q30 值均大于建库对照的值,说明试验较好地实现了质量控制,充分地展现出 SLAF-seq 技术的灵活性和准确性。从获得 SNP 标记来看,耐冷池和冷敏感池 SNP 数目比较接近,为 43 万个以上,但混池的杂合率分别为 46.42% 和 47.02%,说明混池中出现了较明显的分离,较高的杂合率可能对性状的定位造成一定影响。

分离群体混合分析(Bulked Segregant Analysis,BSA)方法是一种简化的定位与目标性状连锁分子标记的方法,由于混池一般来自具有相对性状的两个亲本并经过了向性状的两个方向的选择,因此认为混池间基因型的差异是造成表现型差异的重要原因,即这种基因型的差异就是与目标性状紧密连锁的区域甚至是基因本身。SLAF-seq 技术与 BSA 相结合,可以充分发挥两者的优势,降低成本,缩短周期,实现快速对目标性状的定位^[10-12]。玉米萌发期耐冷性属于多基因控制的数量性状,遗传基础比较复杂。本研究采用 ED (Euclidean Distance,欧氏距离)算法计算与目的基因连锁的区域。最终在玉米第 9 号染色体上定位了一个位置为 1 923 152.00~92 980 413.00 bp (91 057 262.00 bp)区域。从定位的结果来看,定位的区域还比较大,超过了 91 Mb。从关联区域的基因富集分析来看所富集的基因也比较多,共计 3 170 个基因被注释到不同的通路当中。后续可以针对这 91 Mb 区域开展和筛选到的基因进行进一步研究,同时本研究所发的标记可以扩展到其他玉米材料,在更广泛的遗传材料背景下开展玉米萌发期耐冷性的遗传研究。

参考文献:

[1] 张晓聪,雍洪军,张焕欣,等.玉米芽期和苗期耐冷性研究进

展[J].作物杂志,2012(6):8-14.

- [2] 马延华.玉米发芽至苗期耐冷性资源鉴定及遗传分析[D].沈阳:沈阳农业大学,2013.
- [3] Trigen A, Freitag N, Gilbert X, et al. Association mapping for chilling tolerance in elite flint and dent maize inbred lines evaluated in growth chamber and field experiments[J]. Plant, Cell & Environment, 2013, 36: 1871-1887.
- [4] 杜龙岗,王美兴.玉米 SLAF 标记的开发及其在玉米果皮纤维素含量 BSA 分析中的应用[J].中国农业科学,2018, 51(8):1421-1430.
- [5] 刘渊,孔佑宾,李喜焕,等.基于 SLAF-BSA 技术挖掘大豆酸性磷酸酶候选基因及标记开发[J].植物遗传资源学报, 2020, 21(1):164-173.
- [6] Gene Ontology Consortium. The Gene Ontology(GO) database and informatics resource[J]. Nucleic Acids Research, 2004, 32(S1):D258-D261.
- [7] 张一中,范昕琦,杨慧勇,等.基于简化基因组测序高粱育种材料亲缘关系的分析[J].生物技术通报,2020, 36(12): 21-33.
- [8] 谭炎宁,余东,盛夏冰,等. SLAF-seq BSA 定位水稻黄叶转绿基因 *grc2* 的效果研究[J].农业生物技术学报,2020, 28(3):381-388.
- [9] 唐晓敏,张春荣,周良云,等.基于 SLAF-Seq 技术的广金钱草 SNP 位点开发及遗传分析[J].分子植物育种,2020, 18(18):6101-6107.
- [10] 王伟,刘凡,任莉,等.采用 SLAF-seq 技术开发甘蓝型油菜霜霉病抗性 SNP 位点[J].中国油料作物学报,2016, 38(5):555-562.
- [11] 姜涛,温春秀,田伟,等.基于 SLAF-seq 技术连翘 SNP 分子标记开发及遗传多样性分析[J/OL].分子植物育种:1-15 [2020-10-22]. <http://kns.cnki.net/kcms/detail/46.1068.S.20200720.1130.010.html>.
- [12] 郁晓敏,金杭霞,杨清华,等.利用 SLAF-seq 结合 BSA 方法发掘大豆种皮色相关基因[J/OL].分子植物育种:1-8 [2020-10-22]. <http://kns.cnki.net/kcms/detail/46.1068.S.20191212.1148.008.html>.

Preliminary Study on Location Cold Tolerance During Maize Germination Stage Based on SLAF-Super BSA Technology

CAO Shi-liang¹, YU Tao¹, DONG Ling², LIU Bao-min¹, WANG Cheng-bo¹, LI Wen-yue¹, YANG Geng-bin¹, WU Er-na¹

(1. Maize Research Institute, Heilongjiang Academy of Agricultural Sciences, Key Laboratory of Maize Biology, Genetics and Breeding in North Northeast China, Ministry of Agriculture and Rural Areas, Key Laboratory of Maize Germplasm Resources Creation and Utilization, Harbin 150086, China; 2. Agricultural College, Northeast Agricultural University, Harbin 150030, China)

Abstract: In order to speed up maize genetics and breeding on cold tolerance during germination stage, Bulk Segregant Analysis (BSA) and Specific-locus Amplified Fragment-sequencing (SLAF) were used to locate the chromosomal regions related to the trait. The results showed that both two parents and two mixed pools were sequenced with SLAF method and 282 563 SLAF tags were obtained, the average paternal sequencing depth of SLAF tag was 27.06× and the average maternal sequencing depth of SLAF tag was 31.85×, sequencing depth of mixed pools were 36.72× and 34.17× respectively. 47 624 polymorphism SLAF tags were obtained with a polymorphism rate of 16.85%. A 91.06 Mb region associated with cold tolerance during germination stage was located on chromosome 9 of maize, in which a total of 3 170 genes were associated.

Keywords: maize; cold tolerance during germination stage; SLAF; BSA