

基于 GBS、DArT-array 和 SSR 标记构建 普通小麦高密度遗传图谱

赵春杰¹ 李慧慧² 刘德梅³ 兰彩霞¹

1. 华中农业大学植物科学技术学院, 武汉 430070; 2. 中国农业科学院作物科学研究所, 北京 100081;
3. 中国科学院西北高原生物研究所/青海省作物分子育种重点实验室, 西宁 810008

摘要 以印度小麦品种 Sujata 与澳大利亚小麦品系 Avocet 杂交获得 148 个家系的重组自交系群体为材料, 利用 6 397 对基于二代测序技术的 GBS 标记、705 对 DArT-array 标记和 164 对 SSR 标记构建普通小麦高密度遗传图谱。结果显示: 该图谱覆盖小麦的 21 条染色体, 分为 25 个连锁群, 总长度 6 104.4 cM, 标记间平均距离为 0.84 cM。本研究构建的高密度连锁图为后续 QTL 定位、图位克隆和分子标记辅助选择等研究奠定基础。

关键词 小麦; 连锁图谱; GBS 标记; DArT-array 标记; SSR 标记; 重组自交系

中图分类号 S 512.103.2 **文献标识码** A **文章编号** 1000-2421(2019)06-0056-06

小麦是重要的粮食作物, 它是异源六倍体 (AABBDD, $2n=42$) 植物, 基因组结构复杂。小麦农艺性状遗传基础的研究很大程度上依赖于遗传图谱的绘制。随着技术的不断发展, 多种标记种类在小麦遗传作图中被广泛应用, 其中包括早期依赖于酶切的限制性片段多态性 (RFLP) 标记^[1]、基于 PCR 技术的随机扩增多态性 (RAPD)^[2] 标记、扩增片段长度多态性 (AFLP)^[3] 标记和简单重复序列 (SSR)^[4-6] 标记。

Chao 等^[1] 在 1988 年首次报道了基于 RFLP 标记的小麦遗传图谱。但小麦中 RFLP 标记数量少, 多态性较低, 所构建的遗传图谱标记密度低; 此外, RFLP 标记操作步骤繁琐, 费时费力, DNA 用量大, 因而在小麦中应用并不广泛^[7]。1998 年, Röder 等^[4] 利用 230 对 SSR 标记对小麦遗传群体 Opata85×W7984 进行基因型分析, 构建小麦第 1 个 SSR 标记遗传图谱。随后, Somers 等^[8] 将 4 个作图群体的结果整合到一起绘制了 1 个包含 1 235 对 SSR 标记的高密度遗传连锁图, 全长 2 569 cM, 标记间平均间隔 2.2 cM。Penner^[9] 利用 620 对 AFLP 标记及 42 对 SSR 标记, 以 Garent×Saunders 群体为材料, 结合中国春缺体四体材料, 构建

了普通小麦 AFLP 标记遗传连锁图。

尽管利用这些标记已构建了比较完整的图谱, 但其标记饱和度还不足以指导后续更加深入的遗传学研究, 如基因精细定位及图位克隆; 随着二代测序的普及以及测序成本的降低, 基于全基因组重测序的基因型分析方法应运而生。GBS (genotyping by sequencing) 标记平台是通过对作图群体进行低覆盖的全基因组测序, 获取其中的 SNP 基因型信息, 进而用于遗传连锁图构建。与传统方法相比, 该方法具有标记密度高、工作量小和可重复性好等优点^[10]。Poland 等^[11] 利用 GBS 的方法对来源于国际玉米小麦改良中心的 254 个高世代家系进行分析, 获得了 41 371 对 SNP 标记。Li 等^[12] 以 3 个重组自交系群体 PBW343×Kingbird、PBW343×Kenya Swara 和 PBW343×Muu 为材料, 利用 28 644 对 GBS 标记, 构建 1 个标记间平均长度为 0.88 cM 的高密度遗传连锁图, 并利用该连锁图验证了多个锈病相关的抗性基因/QTL。Yang 等^[13] 利用 2 917 对 GBS 标记, 以 CM28×CM28TP 的 F₂ 代群体为材料, 构建了一个包含 21 个连锁群, 总长为 2 371.4 cM 的遗传连锁图, 并将控制雄蕊数目的基因 *Pis1* 定位于 2D 染色体上。

收稿日期: 2019-03-28

基金项目: 国家自然科学基金国际(地区)合作与交流项目(31861143010); 华中农业大学自主科技创新基金项目; 中央高校基本科研业务费专项(2662019PY009); 青海省作物分子育种重点实验室开放课题(2017-ZJ-Y14)

赵春杰, 硕士研究生, 研究方向: 小麦分子遗传学. E-mail: 429216195@qq.com

通信作者: 兰彩霞, 博士, 教授, 研究方向: 小麦抗病分子育种. E-mail: cxlan@mail.hzau.edu.cn

本研究以印度抗条锈病、叶锈病小麦品种 Sujata 和澳大利亚极度感病小麦品系 Avocet 为亲本构建的重组自交系为材料^[14], 利用 6 397 对 GBS 标记、705 对 DArT-array (diversity arrays technology, DArT-array) 标记和 164 对传统的 SSR 标记对该群体进行基因型分析, 构建普通小麦的高密度遗传图谱, 以期通过多种标记相结合的策略, 利用更多的遗传标记, 构建更长遗传距离的连锁图谱, 为后续遗传分析工作奠定基础。

1 材料与方法

1.1 试验材料

供试亲本为印度小麦品种 Sujata 和澳大利亚小麦品系 Avocet。印度小麦品种 Sujata 在田间表现出早熟、高产、白粒、出粉率高、高秆等特点, 并且对条锈病和叶锈病保持抗性 50 多年, 在生产实践中大面积推广应用^[14]; 而澳大利亚小麦品系 Avocet 则是公认的条锈病和叶锈病感病亲本, 该材料在田间表现株高中等、极度感病、生育期比 Sujata 晚 7 d 左右。本研究以 Sujata 和 Avocet 为亲本, 通过单粒传的方法^[15], 构建 148 份 F₃ 代重组自交系群体。

1.2 DNA 提取及数据分析

利用 CTAB 法抽提亲本和重组自交系群体的

DNA, 每个重组自交系家系 DNA 来源约 20 个单株混合取样。首先, 利用 1 450 对 DArT-array 标记和 520 对 SSR 标记对 2 个双亲和 140 份重组自交系进行基因型分析。然后, 利用 6 397 对 GBS 标记对该群体再次进行基因型分析。群体基因型数据可视化通过 Onemap 软件包 (<https://github.com/augusto-garcia/onemap/blob/master/README.md>) 实现, 通过 Joinmap 4.1 结合 R/QTL 软件包^[16] 构建遗传连锁图谱。

2 结果与分析

2.1 SSR 和 GBS 标记的筛选及分布

通过对亲本 Avocet 和 Sujata 进行多态性筛选, 共获得 164 对 SSR 多态性标记和 705 对 DArT-array 标记。其中 DArT-array 标记在不同染色体上分布不均衡, 如 6A 染色体上分布的多态性标记最多, 有 113 对; 而 5D 染色体上分布最少, 仅有 1 对 (表 1)。通过二代测序的方法, 对亲本 Avocet 和 Sujata 整个群体进行基因型分析, 共获得 6 397 对高质量 GBS 多态性标记。与 DArT-array 标记类似, GBS 标记在不同染色体上的分布也并不均衡; 如 3B 和 5B 染色体的多态性标记数目均超过 1 000 对, 而 3D 和 5D 染色体则少于 20 对 GBS 标记 (表 1)。

表 1 RIL 群体遗传连锁图谱统计信息

Table 1 Statistical information of linkage map for RIL population

染色体 Chromosome	连锁群数目 No. of groups	GBS 标记 数目 No. of GBS markers	DArT-array 标记数目 No. of DArT- array markers	SSR 标 记数目 No. of SSR markers	连锁群 长度/cM Length of linkage group	标记间平 均距离/cM Average interval distance	偏分离标 记数目 No. of distorted markers
1A	1	418	56	10	418.68	0.87	12
2A	1	326	19	10	407.62	1.15	49
3A	1	66	14	4	204.11	2.43	7
4A	1	118	25	15	286.42	1.81	15
5A	1	368	8	9	227.87	0.59	6
6A	1	31	113	14	177.33	1.12	72
7A	1	221	37	0	348.16	1.35	25
1B	1	454	35	15	346.26	0.69	33
2B	1	342	27	7	374.84	1.00	3
3B	1	1 012	60	11	435.07	0.4	18
4B	1	87	21	6	228.93	2.01	7
5B	2	1 052	37	13	311.62	0.28	57
5B/7B	1	852	1	2	284.81	0.33	56
6B	1	569	57	9	393.52	0.62	9
7B	1	55	30	14	266.01	2.69	9
1D	1	32	15	0	111.64	2.38	4

续表 1 Continued Table 1

染色体 Chromosome	连锁群数目 No. of groups	GBS 标记 数目 No. of GBS markers	DArT-array 标记数目 No. of DArT- array markers	SSR 标 记数目 No. of SSR markers	连锁群 长度/cM Length of linkage group	标记间平 均距离/cM Average interval distance	偏分离标 记数目 No. of distorted markers
2D	1	86	14	5	267.26	2.55	8
3D	1	14	99	4	158.58	1.36	52
4D	1	35	3	8	119.11	2.59	37
5D	1	11	1	3	88.17	5.88	0
6D	2	76	9	2	412	4.74	31
7D	2	172	24	3	236.39	1.19	24
小计 Total	25	6 397	705	164	6 104.40	0.84	534
A	7	1 548	272	62	2 070.19	1.10	186
B	9	4 423	268	77	2 641.06	0.55	192
D	9	426	165	25	1 393.15	2.26	156

2.2 遗传图谱的绘制

利用 6 397 对高质量 GBS 标记,结合 705 对 DArT-array 标记和 164 对 SSR 标记,对本研究群体进行基因型分析。通过 R 语言程序包 onemap,将群体个体的基因型进行可视化展示(图 1),结果表明 7 266 对分子标记可以高密度覆盖小麦基因组。利用这些标记,构建小麦遗传连锁图谱,包含

25 个连锁群,覆盖小麦 21 条染色体,其中 5B、6D 和 7D 染色体各包含 2 个连锁群,其他染色体均只有 1 个连锁群(表 1)。连锁图谱总长为 6 104.4 cM,标记间平均间距为 0.84 cM(表 1,图 3)。有 3 926 对标记间遗传距离小于 1 cM,占总数目的 54%,另有 95.7%的标记间遗传距离小于 5 cM(图 2)。

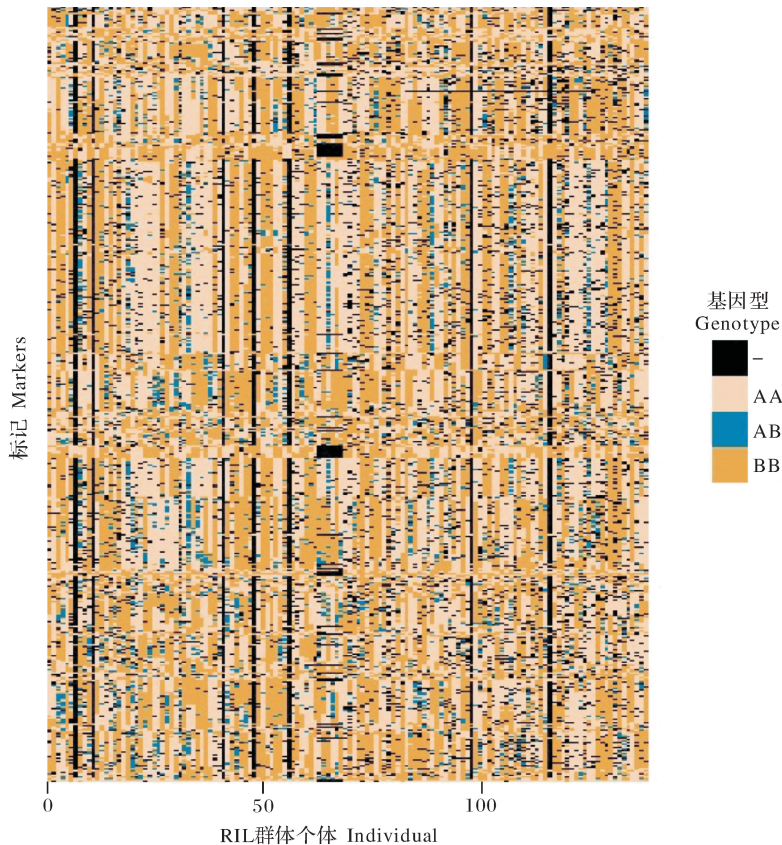


图 1 重组自交系群体标记基因型热图

Fig.1 Heatmap of the genotype for all markers in the recombinant inbred lines

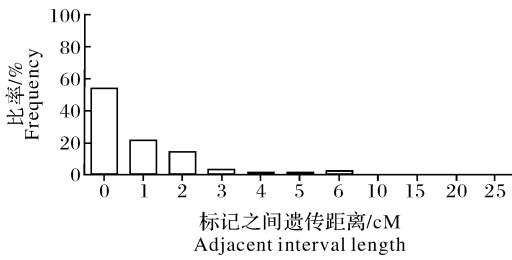


图 2 相邻标记间遗传距离在全基因组的分布情况
Fig.2 Frequency of markers on genetic linkage map according to adjacent interval length

染色体组统计分析结果显示：B 组染色体上有效标记数目最多，包含 4 423 对 GBS 标记、268 对 DArT-array 标记和 77 对 SSR 标记，分别占 3 种标记总数的 69.1%、38.0% 和 47.0%；此外，B 组染色体的连锁群长度最长(2 641.06 cM)，标记间平均遗传距离最小(0.55 cM)；A 组染色体有 1 548 对 GBS

标记、272 对 DArT-array 标记和 62 对 SSR 标记，连锁群全长 2 070.19 cM，标记间平均遗传距离为 1.1 cM；D 染色体组的这 3 类标记个数最少，有 426 对 GBS 标记、165 对 DArT-array 标记和 25 对 SSR 标记，其中 GBS 标记仅占总数的 6.7%；此外，D 组染色体的连锁群总长最短(1 393.15 cM)，且标记间平均遗传距离最长(2.26 cM)(表 1, 图 3)。

进一步分析不同染色体的标记情况，GBS 标记在 B 组染色体的 3B 和 5B 上分布较多，分别为 1 012 对和 1 052 对，同时 5D 染色体含有最少的 GBS 标记，仅 11 对(0.2%)。3B 染色体标记总数最多，连锁图遗传距离最长(435.07 cM)，标记间平均遗传距离最短(0.4 cM)；而 5D 染色体标记数目最少，连锁图最短(88.17 cM)，标记间平均遗传距离最大(5.88 cM)(表 1, 图 3)。

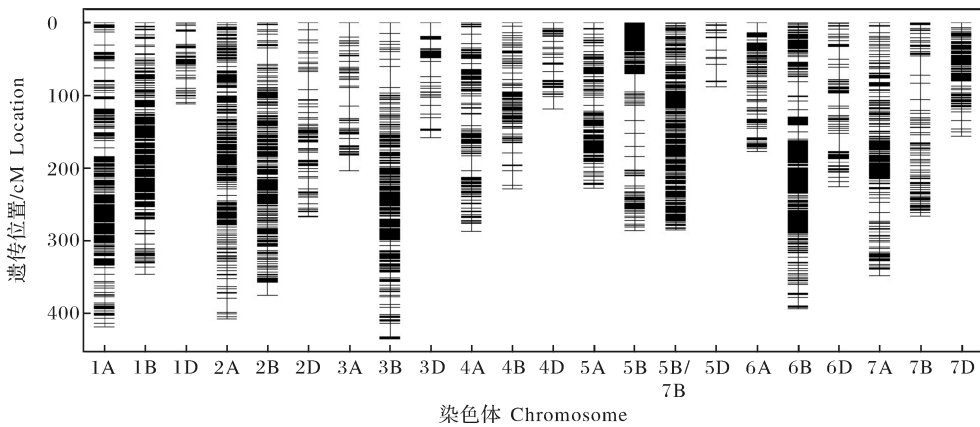


图 3 利用 Avocet 和 Sujata 的 148 个重组自交系构建的小麦高密度遗传连锁图

Fig.3 High-density linkage map constructed with 148 RILs derived from the cross between Avocet and Sujata

通过与近年小麦中构建的遗传连锁图比较，发现研究结果都有明显的提高；同时，在标记间平均遗传距离上，本研究结果与前人结果相当(表 2)。

表 2 遗传连锁图的比较

Table 2 Comparison of linkage maps between present study and previous reports

作图群体 Type of population	标记数目 No. of markers	标记类型 Type of markers	遗传图谱长度/cM Length of linkage map	标记间平均遗传距离/cM Adjacent distance	连锁群数目 No. of linkage group	文献或来源 Reference
RIL ¹⁾	3 757	GBS	3 302.45	0.88	29	[12]
F ₂	29 17	GBS	2 371.40	0.81	21	[13]
F ₃	3 083	GBS	4 293.94	0.80	31	[19]
RIL	10 172	iSelect SNP array	4 676.10	0.39	29	[20]
RIL	6 312	iSelect SNP array, SSR	3 049.40	1.10	24	[22]
DH	1 854	SNP array, SSR	4 082.44	2.20	30	[22]
RIL²⁾	7 266	GBS, DArT, SSR	6 104.40	0.84	25	

注：1) 3 个 RIL 群体为 PBW343 × Kingbird、PBW343 × Kenya Swara 和 PBW343 × Muu。2) 黑色加粗行为本研究结果。
Note: 1) Three RIL populations are PBW343 × Kingbird, PBW343 × Kenya Swara and PBW343 × Muu; 2) The line of black bold is the results of this study.

2.3 标记偏分离分析

作图群体双亲本间多态性标记的 2 种基因型在群体中理论上应当符合 1:1 的比例。对连锁图谱上的 7 266 对标记的 2 种基因型在重组自交系群体中的分布进行卡方测验表明,以 0.01 概率水平为准,534 对标记表现偏分离,占总标记数的 7.3% (表 1)。在不同亚基因组具有偏分离的标记数目差异不大,A、B 和 D 基因组上分别有 186、192 和 156 对偏分离标记。但是,不同染色体上的分布又存在明显的差异。6A 染色体上偏分离的标记最多(72 对,13.5%),而 5D 染色体上没有偏分离标记(表 1)。

3 讨 论

本研究利用 GBS、DArT-array 和 SSR 标记相结合的方法,共计 7 266 对标记用于连锁图谱构建,覆盖小麦所有染色体,总遗传距离长度为 6 104.4 cM,标记间平均遗传距离为 0.84 cM。

由于小麦为异源六倍体植物,其 A、B、D 3 个染色体组之间具有一定的同源性,导致传统分子标记在不同材料之间的多态性较低,限制了高密度小麦连锁图谱的构建。张帆等^[17]利用 126 对 SSR 标记构建了总长为 1 773.2 cM 的遗传连锁图,标记间平均遗传距离为 14.07 cM。李卫华等^[18]用 184 对 SSR 标记,获得了类似的标记密度(16.32 cM)。近期,随着新技术不断开发应用,以二代测序技术为基础的 GBS 标记和 SNP 芯片对构建小麦高密度遗传连锁图发挥了重要作用。多个研究结果证明,不管是 GBS 标记^[12-13,19]还是 SNP 芯片^[20-22],在不同类型群体中的应用,都获得了比较稳定的结果(表 2)。

本研究结果表明小麦遗传连锁图的构建还存在一定优化空间,如 5B、6D 和 7D 等多条染色体无法形成一个单独的连锁群(表 1)。一方面,虽然本研究群体中的多态性分子标记较多,但是对于小麦基因组(17G)来说,这些标记的覆盖率还是很低。与此结果类似,Kumar 等^[20]利用更多数目的标记(10 172 对),得到了 29 个连锁群。另一方面,双亲在这 3 条染色体上多态性标记较少,也可能导致无法构建连锁图。再者,在不同染色体上的标记数目分布不均衡,其中 GBS 标记尤为严重。GBS 标记分布较多的染色体是 3B 和 5B,它们分别具有 1 012 (15.8%)和 1 052 (16.4%)对标记;而与之相对分布极少的染色体是 3D 和 5D,分别只有 14 和 11 对标

记。除此之外,其他标记数目少于 100 对(占比少于 1.6%)的染色体还较多:如 3A、6A、4B、7B、1D、2D、4D 和 6D。这种状况造成部分染色体标记间平均遗传距离较大,如 5D 染色体为 5.88 cM。后续研究需要对 GBS 标记的使用进行适当的优化。

综上,本研究采用 GBS、DArT-array 和 SSR 3 种标记相结合的方法构建普通小麦高密度遗传图谱,一定程度上弥补了单种标记在标记密度、覆盖度等方面的缺点,为后续小麦高密度遗传图谱构建提供了新方法。

致 谢: 特别感谢国际玉米小麦改良中心普通小麦育种课题组 Ravi Singh 博士团队友情提供本研究所需重组自交系群体!

参 考 文 献

- [1] CHAO S, SHARP P, WORLAND A, et al. RFLP-based genetic maps of homoeologous group 7 chromosomes of wheat and barley[J]. *Theoretical applied genetics*, 1989, 78: 495-504.
- [2] WILLIAMS J G, KUBELIK A R, LIVAK K J, et al. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers[J]. *Nucleic acids research*, 1990, 18 (22): 6531-6535.
- [3] VOS P, HOGERS R, BLEEKER M, et al. AFLP: a new technique for DNA fingerprinting[J]. *Nucleic acids research*, 1995, 23(21): 4407-4414.
- [4] RÖDER M S, KORZUN V, WENDEHAKKE K, et al. A microsatellite map of wheat[J]. *Genetics*, 1998, 149(4): 2007-2023.
- [5] PESTSOVA E, GANAL M W, RÖDER M S. Isolation and mapping of microsatellite markers specific for the D genome of bread wheat[J]. *Genome*, 2000, 43(4): 689-697.
- [6] GUPTA P, BALYAN H, EDWARDS K, et al. Genetic mapping of 66 new microsatellite (SSR) loci in bread wheat[J]. *Theoretical applied genetics*, 2002, 105(3): 413-422.
- [7] LIU Y G, TSUNEWAKI K. Restriction fragment length polymorphism (RFLP) analysis in wheat. II. Linkage maps of the RFLP sites in common wheat[J]. *The Japanese journal of genetics*, 1991, 66(5): 617-633.
- [8] SOMERS D J, ISAAC P, EDWARDS K. A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.)[J]. *Theoretical applied genetics*, 2004, 109(6): 1105-1114.
- [9] PENNER G A. An AFLP based genome map of wheat (*Triticum aestivum*) [C]. San Diego: [s.n.], 1998: 163.
- [10] HE J, ZHAO X, LAROCHE A, et al. Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding[J]. *Frontiers in plant science*, 2014, 5: 484-492.
- [11] POLAND J, ENDELMAN J, DAWSON J, et al. Genomic selec-

- tion in wheat breeding using genotyping-by-sequencing[J]. The plant genome, 2012, 5(3):103-113.
- [12] LI H, VIKRAM P, SINGH R P, et al. A high density GBS map of bread wheat and its application for dissecting complex disease resistance traits[J]. BMC genomics, 2015, 16(1):216-230.
- [13] YANG Z, CHEN Z, PENG Z, et al. Development of a high-density linkage map and mapping of the three-pistil gene (*Pis1*) in wheat using GBS markers[J]. BMC genomics, 2017, 18(1):567-575.
- [14] LAN C, ZHANG Y, HERRERA-FOESSEL S A, et al. Identification and characterization of pleiotropic and co-located resistance loci to leaf rust and stripe rust in bread wheat cultivar Sujata[J]. Theoretical applied genetics, 2015, 128(3):549-561.
- [15] BASNET B R, SINGH R P, HERRERA-FOESSEL S A, et al. Genetic analysis of adult plant resistance to yellow rust and leaf rust in common spring wheat Quaiu 3[J]. Plant disease, 2013, 97(6):728-736.
- [16] BROMAN K W, WU H, SEN S, et al. R/qtl: QTL mapping in experimental crosses[J]. Bioinformatics, 2003, 19(7):889-890.
- [17] 张帆, 臧明丽. 利用 SSR 分子标记构建小麦遗传图谱[J]. 南方农机, 2017, 48(16):1-3.
- [18] 李卫华, 刘伟, 尤明山, 等. 利用多种 SSR 引物构建小麦遗传连锁图谱及其多态性分析[J]. 麦类作物学报, 2007, 27(1):1-6.
- [19] LI Q, GUO J, CHAO K, et al. High-density mapping of an adult-plant stripe rust resistance gene *YrBai* in wheat landrace Baidatou using the whole genome DArTseq and SNP analysis [J]. Frontiers in plant science, 2018, 9:1120-1129.
- [20] KUMAR A, MANTOVANI E, SEETAN R, et al. Dissection of genetic factors underlying wheat kernel shape and size in an Elite×Nonadapted cross using a high density SNP linkage map [J]. The plant genome, 2016, 9(1):81-103.
- [21] SU Q, ZHANG X, ZHANG W, et al. QTL detection for kernel size and weight in bread wheat (*Triticum aestivum* L.) using a high-density SNP and SSR-based linkage map[J]. Frontiers in plant science, 2018, 9:1484-1497.
- [22] 李龙, 彭智, 毛新国, 等. 小麦高密度遗传图谱构建及抗旱相关生理性状的遗传解析[J]. 植物遗传资源学报, 2018, 19(3):531-538.

Constructing high density genetic linkage map of bread wheat with GBS, DArT-array and SSR markers

ZHAO Chunjie¹ LI Huihui² LIU Demei³ LAN Caixia¹

1. College of Plant Science & Technology, Huazhong Agricultural University, Wuhan 430070, China ;
 2. Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081, China ;
 3. Northwest Institute of Plateau Biology, Chinese Academy of Sciences/Qinghai Provincial Key Laboratory of Crop Molecular Breeding, Xining 810008, China

Abstract Genetic linkage map based on classical molecular marker platforms with low density cannot meet the requirements for wheat functional genomics studies due to its huge and complex genome. 148 recombinant inbred lines (RILs) derived from the crossing Sujata (Indian wheat variety) with Avocet (Australia wheat variety) were used to construct the high-density genetic linkage map with 6 397 genotyping-by-sequencing (GBS) markers, 705 DArT-array markers and 164 simple sequence repeat (SSR) markers. 21 chromosomes were covered by 25 linkage groups with a total length of 6 104.4 cM. It will provide an important resource for QTL mapping, map-based cloning and marker assisted selection in wheat breeding.

Keywords wheat (*Triticum aestivum* L.); linkage map; GBS marker; DArT-array markers; SSR marker; recombinant inbred lines (RILs)

(责任编辑: 张志钰)