Construction of ultra-high-density genetic linkage map of a sorghum-sudangrass hybrid using whole genome resequencing

Qianqian Lu*, Xiaoxia Yu*, Huiting Wang, Zhuo Yu*, Xia Zhang, Yaqi Zhao
Agricultural College, Inner Mongolia Agricultural University, Hohhot, Inner Mongolia, China

*These authors contributed equally to this work.
*yuzhuo58@sina.com

Abstract

The sorghum-sudangrass hybrid is a vital annual gramineous herbage. Few reports exist on its ultra-high-density genetic map. In this study, we sought to create an ultra-high-density genetic linkage map for this hybrid to strengthen its functional genomics research and genetic breeding. We used 150 sorghum-sudangrass hybrid F_2 individuals and their parents (scattered ear sorghum and red hull sudangrass) for high-throughput sequencing on the basis of whole genome resequencing. In total, 1,180.66 Gb of data were collected. After identification, filtration for integrity, and partial segregation, over 5,656 single nucleotide polymorphism markers of high quality were detected. An ultra-high-density genetic linkage map was constructed using these data. The markers covered approximately 2,192.84 cM of the map with average marker intervals of 0.39 cM. The length ranged from 115.39 cM to 264.04 cM for the 10 linkage groups. Currently, this represents the first genetic linkage map of this size, number of molecular markers, density, and coverage for sorghum-sudangrass hybrid. The findings of this study provide valuable genome-level information on species evolution and comparative genomics analysis and lay the foundation for further research on quantitative trait loci fine mapping and gene cloning and marker-assisted breeding of important traits in sorghum-sudangrass hybrids.

Introduction

Sorghum-sudangrass hybrid is an important annual gramineous forage hybrid of *Sorghum bicolor* (Linn.) Moench (2n = 2x = 20) and *S. sudanense* (Piper) Stapf (2n = 2x = 20) [1] with drought tolerance, lodging resistance, and disease resistance characteristics. Moreover, this hybrid has apparent heterosis. Therefore, it has broad application prospects in environmental protection, animal husbandry, and aquaculture, among others. It is widely planted in the United States [2], including northern Colorado [3], China [4], the southern part of the Korean Peninsula, and most of the dry areas of Jeju Island [5]. Dong et al. [6] reported that the sorghum-sudangrass hybrid is rich in amino acids, sugars, minerals, and other compounds, with high nutritional value and good palatability, and is commonly used as hay, silage, or green
feed. With the rapid development of animal husbandry, the quality and quantity of existing forage grass fails to meet the growing demand for forage material in developing herbivorous animal husbandry.

An insufficient forage supply has become a limiting factor in the livestock industry in the northwestern Himalayan region of India [7], the southern part of the Korean Peninsula and Jeju Island [5], the northern and southeastern United States [8, 9], and China [10]. Therefore, it is necessary to identify key genes that control desirable traits and select new varieties of the sorghum-sudangrass hybrid with high yields and quality. However, it is challenging to identify genotypes with ideal traits using traditional breeding. One way to overcome this challenge is to construct molecular marker maps on chromosomes using population separation data generated by specific hybridization combinations [11, 12]. Developing a high-density genetic map is necessary for mapping quantitative trait loci (QTLs) in crops and is an essential genetic tool in other areas of molecular biology [13]. At present, domestic and foreign studies on sorghum-sudangrass hybrid mainly focus on cultivation management [14, 15], development and utilization of the nutritional and feeding value [4, 7], or breeding of new varieties [16]. However, there are few reports on ultra-high density genetic mapping of the sorghum-sudangrass hybrid.

High-quality molecular genetic linkage maps are prerequisites for the genetic analysis, QTL location of important traits, and efficient molecular breeding (e.g., marker assisted selection [MAS]). The difference types of molecular marker affect the accuracy of genetic linkage maps [17–19]. To date, more than 10 genetic maps of sorghum-sudangrass hybrid were mainly constructed by traditional first- and second-generation types of molecular markers such as amplified fragment length polymorphism (AFLP), restriction fragment length polymorphism (RFLP), simple sequence repeat (SSR), sequence associated amplified polymorphism (SRAP), and randomly amplified polymorphic DNA (RAPD) [20–24]. These maps contained 124–444 molecular markers, spanning 803.13–1468.12 cM, and the average distance between markers ranged from 2.98 cM to 8.84 cM. However, the traditional molecular markers on these genetic maps are based on gel electrophoresis and band counting, which is time-consuming, costly, and prone to human errors. Besides, these maps do not meet the requirements of further fine QTL mapping, as well as annotation and cloning of functional genes, due to the limited number of molecular markers, low map resolution, and large linkage gaps on maps.

In recent years, some high-density genetic linkage maps of plants have been constructed using SNP markers developed by next-generation sequencing (NGS) technology [25]. The rapid development of various genotyping and high-throughput sequencing technologies have led to the development of large-scale SNP markers. In general, the SNP markers used for recently map construction of many plants are mainly based on simplified genome sequencing of NGS technology [26–28]. Jin et al. [29] published a genetic map of sorghum-sudangrass hybrid containing 1,065 markers based on restriction-site-associated DNA sequencing (RAD-seq), which covered a distance of 1,191.7 cM with an average interval of 1.11 cM. Compared with the previously reported maps, the map density here was shortened. However, despite low experimental cost and simple operation of the simplified genome sequencing, some deficiencies still exist, including insufficient sequencing depth, low marker coverage, and the inability to find certain key SNPs loci closely related to the main traits [30, 31].

With the development of NGS and the reduction of sequencing cost, whole genome resequencing (WGRS) technology has become one of the most effective research strategies in breeding and genomics, which has been applied to construct maps of important crops, such as rice [32], tea [31], and potato [33]. However, an ultra-high-density map of sorghum-sudangrass hybrid based on WGRS technology has not been reported.
This study applied WGRS to identify SNPs and build genetic maps of sorghum-sudangrass hybrid, using an F2 population obtained by crossing scattered ear sorghum with red hull sudangrass. This ultra-high-density map can be utilized to map QTLs of important traits, perform comparative genomics, and in marker-assisted breeding of sorghum-sudangrass hybrid.

Materials and methods

The protocol described in this peer-reviewed article is published on protocols.io (http://dx.doi.org/10.17504/protocols.io.14egn2226g5d/v1), and is included for printing as S1 File with this article.

Plant materials and genomic DNA extraction

The genetic map was created from a mapping population of 150 individuals selected at random from the F2 generation population, which were obtained by self-cross-bagging from F1 generation population cross between ‘scattered ear sorghum’ (♀) and ‘red hull sudangrass’ (♂). scattered ear sorghum, a variety with strong drought resistance and high yield, was collected from northeast China. Red hull sudangrass has a high nutritional value and good palatability and is planted widely in Inner Mongolia, China [34]. The experimental materials were planted in the experimental field of the Inner Mongolia Agricultural University, Saihan District, Hohhot (Fig 1). During the early jointing stage, young leaves from the F2 individuals and their parents were flash frozen in liquid nitrogen. Genomic DNA (gDNA) was extracted with the DNA secure Plant Kit from Tiangen Biotech, Beijing, China. The quality of the extracted gDNA was confirmed via electrophoresis in a 0.8% (w/v) agarose gel.

Library construction and genotyping by WGRS

The 152 gDNA samples were randomly sheared into 350 bp fragments in the Covaris breaker (Covaris, Woburn, MA, USA). Libraries were built based on Illumina’s TruSeq Library Construction Kit (Illumina, San Diego, CA, USA). Briefly, the gDNA fragments were processed by end repair, modified by poly-A tail and sequencing adapter addition, purified, and amplified via PCR to construct the library. Paired-end sequencing libraries were sequenced with a read length of 350 bp using an Illumina HiSeqTM PE150 (Illumina, San Diego, CA, USA).

The parental genotypes were sequenced separately at a sequencing depth of 29.71× and 28.77×. Individual F2 plants were sequenced at 9.97× coverage. The raw data were filtered to determine the sequencing read quantities, sequencing error rates, Q20, Q30, GC content, and data output. The filtered reads were compared with the reference genome assemblies of sorghum bicolor and used for SNP identification and genotyping.

SNP calling and genotyping

Sequencing data of the parent and offspring plants were aligned to the reference genome sorghum bicolor (sorghum) (https://phytozome-next.jgi doe.gov/info/Sbicolor_v3_1_1 Accession ID: ABXC03000000) using Burrows-Wheeler Aligner (BWA) (http://bio-bwa sourceforge.net/). Duplicate parts (rmDup) were removed by SAMTOOLS (https://www.htslib.org). Marker development was implemented after genotype detection of the parents of sorghum-sudangrass hybrid. Polymorphic markers were obtained after removing the markers of the same type shared by both parents and those missing in one or both parents. GATK (https://gatk.broadinstitute.org) was used to genotype the obtained polymorphic molecular markers, then the markers were then screened out in line with the mapping population, with the conditions of parental 9× and line 3× or 4×, and count the SNPs. In general, polymorphic
markers are divided into eight different separation modes, which are comprised of \( hk \times hk \), \( cc \times ab \), \( nn \times np \), \( aa \times bb \), \( lm \times ll \), \( ab \times cd \), \( ef \times eg \), and \( ab \times cc \). Using the results of parental genotyping, we developed polymorphic markers between the parents, from which we selected those matching the mapping marker type of the population. The percentage of genomic positions covered with at least one and four reads represents \( 1 \times \) and \( 4 \times \) coverage of the reference genome, respectively. sorghum-sudangrass hybrid, as a diploid species, can have a maximum...
of four alleles per locus. The F₂ plants were derived from two homozygous parents; hence, the subsequent analysis focused only on SNP markers that showed the aa × bb segregation pattern.

**Genomics map construction with ultra-high-density**

SNPs were screened for quality improvement of the genetic map based on the following guidelines: abnormal bases and genotypes were denoted as deletions (indicated by "-"); and Chi-square test was used to identify SNP markers with segregation distortion, which were then filtered out (significance, \( p < 0.001 \)). Linkage maps with 10 linkage groups (LGs) were constructed using Lepmap3 [https://sourceforge.net/projects/lep-map3](https://sourceforge.net/projects/lep-map3). After linkage analysis, the obtained high-quality genetic markers, were divided into 10 LGs according to chromosome sequences with LOD values ranged from 2 to 6. The LGs were ordered along the physical location of the chromosomes using the maximum likelihood algorithm, and the genetic distance between markers was calculated using the Kosambi mapping function [35]. Each LG with stable SNPs was aligned to the sorghum_bicolor_v3 reference genome at the cutoff value of \( 1 \times 10^{-10} \) and sequence coverage rate of >85% using BLASTN [https://blast.ncbi.nlm.nih.gov](https://blast.ncbi.nlm.nih.gov) [36]. We evaluated the genetic map of sorghum-sudangrass hybrid using haplotypes and heat maps. Double hybridization and deletion were manifested in the haplotype map as genotyping and labeling sequence errors [37]. Errors in the order of markers were identified using a heat map, reflecting the recombination relationship among the markers of individual LGs. We defined the region with a higher recombination frequency than other regions as the recombination hot spot [38, 39].

**Results**

**Statistics and quality assessment of WGRS**

A total of 150 F₂ individuals and their parents were resequenced. There were 1,180.66 GB of sequence data filtered to produce 7,841,459,884 bp of clean reads. (S1 Table) Among all the high-quality data, the number of reads in the males and females was 151,469,838 bp and 156,284,244 bp, respectively. Each individual had 50,224,705.35 reads. The Q20 and Q30 scores for bases between progenies were 97.72% and 93.53%, respectively. The average GC content was 44.37%. The high-quality data and the reference genome comparison showed that the male and female rates were 98.75% and 98.80%. The average rates of the F₂ generation was 98.58% (Table 1). The results showed that the sequencing data were of high quality and suitable for further analysis.

The obtained reads were aligned to the sorghum genome, and their genome coverage is shown in Table 1 and S2 Table. Genomic coverage for 1× ranged between 92.33% and 97.05%, with parents covering 1× at 97.05% and 97.04%, and offspring covering 1× at 94.78%. Genome coverage for 4× ranged between 77.94% and 95.12%, with parents having 95.07% and 95.12%, and offspring having 82.56%. The sequencing depth of female scattered ear sorghum and male red hull sudangrass was 28.77× and 29.71× on average, respectively, while the average sequencing depth of the offspring was 9.97×.

Using the GATK tool, we detected 5,199,747 and 5,365,314 SNP markers in scattered ear sorghum and red hull sudangrass, respectively; the number of heterozygous sites was 3,797,357 and 4,082,017, respectively, and the number of homozygous sites was 1,402,390 and 1,283,297, respectively. The heterozygous ratios for scattered ear sorghum and red hull sudangrass were 73.03% and 76.08%, respectively. The average number of SNPs detected in the F₂ offspring was 3,191,158.56, of which 80.27% were heterozygous (S3 Table). The results showed that the base coverage depth was distributed uniformly across the genome and that the sequencing data were of high quality.
SNP marker detection and development

A total of 5,633,474 (31.13%) polymorphic markers were obtained (Table 2), which were comprised of hk × hk, cc × ab, nn × np, aa × bb, lm × ll, ab × cd, ef × eg, and ab × cc (Table 3).

Among them, the hk × hk polymorphism site was heterozygous for both parents, the aa × bb site was homozygous for both parents, the hk × hk, ab × cd, and ef × eg sites were heterozygous for both parents, and the nn × np, lm × ll, nn × np, and cc × ab sites were heterozygous for one and homozygous for the other parent. As shown in Table 3, of the 5,633,474 markers, SNPs were most abundant in the lm × ll type (35.30%), followed by that of the nn × np type (29.86%), and only 8 SNPs were in the ab × cd type. Since the mapping population in this study originated from two homozygous parent hybrid plants, the available marker type for the F2 population was of the aa × bb type. Subsequently, 270,715 aa × bb progenies were screened for genotype deletion, genotyping error, insufficient completeness (gene coverage < 75%), and partially separated markers (p < 0.001) (S4 and S5 Tables); 8,009 valid SNP markers were obtained after the screening. These were used for linkage analysis of the genetic map of sorghum-sudangrass hybrid.

Ultra-high-density genetic map construction

Of the 8,009 SNPs identified, subsequent screening of abnormal bases and genotypes, and segregation distortion of unqualified markers produced 5,656 high-quality markers. Using these SNPs, we constructed a high-density genetic map. Each marker was evenly distributed across the 10 LGs (Fig 2). These markers spanned 2,192.84 cM, with an average distance of 0.39 cM.

Table 1. Statistics of sequencing data for sorghum-sudangrass hybrid.

|                        | Scattered ear sorghum (♀) | Red hull sudangrass (♂) | F2 population |
|------------------------|---------------------------|--------------------------|---------------|
| Clean bases (G)        | 22,720,475,700            | 23,442,636,600           | 7,533,705,802 |
| Clean reads (bp)       | 151,469,838               | 156,284,244              | 50,224,705.3  |
| Mapped reads           | 149,574,247               | 154,407,070              | 49,509,454.1  |
| Mapping rate (%)       | 98.75                     | 98.80                    | 98.58         |
| Q20 (%)                | 97.06                     | 96.98                    | 97.72         |
| Q30 (%)                | 92.16                     | 91.89                    | 93.52         |
| GC content (%)         | 43.70                     | 43.81                    | 44.37         |
| Average depth (×)      | 28.77                     | 29.71                    | 9.97          |
| Coverage 1× (%)        | 97.05                     | 97.04                    | 94.78         |
| Coverage 4× (%)        | 95.07                     | 95.12                    | 82.56         |
| Total SNP              | 5,365,314                 | 5,199,747                | 3,191,158.56  |
| Heterozygosis SNP      | 4,082,017                 | 3,797,357                | 2,561,679.55  |
| Homozygous SNP         | 1,283,297                 | 1,402,390                | 629,479.01    |
| Het rate (%)           | 76.08                     | 73.03                    | 80.27         |

Table 2. Polymorphism locus statistics.

| Marker type            | Numbers | Rate (%) |
|------------------------|---------|----------|
| Polymorphic marker     | 5,633,474 | 31.13    |
| Missing in parents 1/2/both | 4,262,790 | 23.56    |
| Parents with same type | 8,197,958 | 45.31    |
| Number of input file lines | 17,854,886 | 100      |
between adjacent markers. Approximately 401 to 755 markers were found in each LG, covering a distance of 115.39 cM to 264.04 cM. The number of SNP markers, the total genetic distance, and the average genetic distance on each LG are presented in Table 4 and S6 Table. Among the 10 LGs, the longest was LG7 at 264.04 cM, and the shortest was LG5 at 115.39 cM. The lowest density between markers was observed in LG5, containing 447 SNP markers, with a span of 115.39 cM and the average distance of 0.26 cM between the markers. LG1 had the maximum gap distance between markers of 13.72 cM. In the 10 LGs, 46 intervals had a spacing of 5–10 cM, three intervals had a spacing of 10–20 cM, and 5,607 intervals had a spacing of <5 cM. In LG9, up to 99.67% of the intervals between markers was <5 cM. Compared with the map published by Jin et al. [29], our maps are superior in terms of the mapping population size (150 here vs. 126 in [29]), high-throughput sequencing method (WGR-seq vs. RAD-seq), markers number (5,656 vs. 1,065), and sequencing data (7,841,459,884 vs. 233,669,757). Compared to other published genetic linkage maps for sorghum-sudangrass hybrid, our linkage map had the longest total map length (2,192.84 cM) and the shortest mean genetic interval between markers (0.39 cM/marker) (Table 5). There were three gaps with inter-marker genetic distances greater than 10 cM, among which the largest was 13.72 cm (LG10). Gap distances <5 cM accounted for 99.13% of the total LGs. Overall, the marker distribution of the map was relatively uniform, and the quality of the map was high.

Genetic and physical map comparisons and collinearity analyses

Sorghum genome is 3–4 times smaller than that of maize, and is considered the model plant for polyploid sugarcane and diploid crops [40]. The sorghum-sudangrass hybrid genome was compared to the sorghum genome by Yang et al. [41] and Jin et al. [29], who demonstrated that it is feasible to develop SNP markers for sorghum-sudangrass hybrid using the Sorghum genome. Since the sequencing of the sorghum-sudangrass hybrid genome was incomplete, we used the sorghum genome as a reference genome to develop SNP markers for sorghum-sudangrass hybrid.

Quality of a genetic linkage map is judged by the relationship between the SNP marker genetic and physical location [42]. We aligned all 5,656 SNP markers to the reference genome and conducted collinearity analyses. The 10 LGs showed high collinearity, covering almost the entire physical length of the reference genome (Fig 3). The genetic linkage map of most regions matched the reference genome, while short reversals and translocations existed in few areas. Chromosome inversions in individuals that are the product of selfing in a population are important for plant evolution, in particular their adaptation and speciation [37]. This result may guide studies on the utilization of heterosis and identification of hybrids.

Table 3. Segregation distortion for SNP markers.

| Segregant type | Cross type | Female genotype | Male genotype | Count | Percentage |
|---------------|------------|-----------------|---------------|-------|------------|
| Type1         | hk × hk    | Heterozygosis    | Heterozygosis  | 1,674,553 | 29.72%     |
| Type2         | cc × ab    | Heterozygosis    | Heterozygosis  | 951   | 0.02%      |
| Type3         | nn × np    | Heterozygosis    | Heterozygosis  | 1,682,365 | 29.86%     |
| Type4         | lm × ll    | Heterozygosis    | Heterozygosis  | 1,988,538 | 35.30%     |
| Type5         | aa × bb    | Heterozygosis    | Heterozygosis  | 270,715 | 4.81%      |
| Type6         | ab × cc    | Heterozygosis    | Heterozygosis  | 1137  | 0.02%      |
| Type7         | ab × cd    | Heterozygosis    | Heterozygosis  | 8     | 0.00%      |
| Type8         | ef × eg    | Heterozygosis    | Heterozygosis  | 15,207 | 0.27%      |

https://doi.org/10.1371/journal.pone.0278153.t003
Fig 2. Ultra-high-density genetic linkage map of sorghum-sudangrass hybrid.

https://doi.org/10.1371/journal.pone.0278153.g002
Qualitative analysis of ultra-high-density genetic maps

In this study, haplotype maps were constructed for 150 F2 individuals using 5,656 SNP markers, with almost all the markers clearly defined on the LGs (S1 Fig). The probability of double crossover rate and deletions was extremely low, indicating high map quality. We constructed the heat maps of the 10 chromosomes of sorghum-sudangrass hybrid. The sequence of the 10 LG markers showed the same trend (a red diagonal line in S2 Fig), indicating the high quality of the map.

Discussion

An increase in the number and type of markers is an effective way of improving the resolution of a map, accelerating the exploration of new genomic information, and mapping of QTLs for important agronomic traits [38, 43]. WGRS represents a new technique to exploit SNPs in plant genomes efficiently. The linkage maps with high densities of various economically important crops have been extensively constructed using WGRS [44–48].

The advantages of WGRS are two-fold. First, compared with marker-based genotyping, maps can be constructed faster and more accurately [34]. Large-scale SNPs developed by WGRS can compensate for the deficiency of marker-based genotyping methods in map construction. Second, WGRS is superior to simplified genome sequencing technologies regarding marker development. Because it uses ultrasonic technology to randomly shear the genomic DNA, it is not restricted by limited restriction enzyme cutting sites. The marker polymorphism is developed with high density, stable inheritance, and easy automation analysis. With the increasing advancement in DNA sequencing platforms, WGRS technology has become the primary choice of NGS.

Table 4. Characteristics of genetic linkage groups of sorghum-sudangrass hybrid.

| Linkage group | Marker number | Length covered | Average length | Max gap | Gap < 5 cM | Gap 5–10 cM | Gap 10–20 cM | Gap >20 cM | Ratio |
|---------------|---------------|----------------|----------------|---------|-----------|-------------|--------------|-----------|-------|
| LG1           | 531           | 203.34         | 0.38           | 13.72   | 528       | 2           | 1            | 0         | 99.44 |
| LG2           | 569           | 219.04         | 0.38           | 6.39    | 565       | 4           | 0            | 0         | 99.30 |
| LG3           | 480           | 223.58         | 0.47           | 7.15    | 477       | 3           | 0            | 0         | 99.38 |
| LG4           | 755           | 214.69         | 0.28           | 9.52    | 747       | 8           | 0            | 0         | 98.94 |
| LG5           | 447           | 115.39         | 0.26           | 6.39    | 445       | 2           | 0            | 0         | 99.55 |
| LG6           | 401           | 219.86         | 0.55           | 12.85   | 392       | 8           | 1            | 0         | 97.76 |
| LG7           | 679           | 264.04         | 0.39           | 9.52    | 676       | 3           | 0            | 0         | 99.56 |
| LG8           | 634           | 243.72         | 0.38           | 10.33   | 625       | 8           | 1            | 0         | 98.58 |
| LG9           | 598           | 240.17         | 0.40           | 7.16    | 596       | 2           | 0            | 0         | 99.67 |
| LG10          | 562           | 249.01         | 0.44           | 7.93    | 556       | 6           | 0            | 0         | 98.93 |
| Total         | 5656          | 2192.84        | 0.39           | 13.72   | 5607      | 46          | 3            | 0         | 99.13 |

https://doi.org/10.1371/journal.pone.0278153.t004

Table 5. Statistics of the existing map of sorghum-sudangrass hybrid.

| Mapping population | Marker types | Markers number | Length (cM) | Average distance (cM) | Linkages | References |
|--------------------|--------------|----------------|-------------|-----------------------|----------|------------|
| F2:3               | AFLP, RAPD   | 166            | 836         | 5.03                  | 10       | Lu, 2005 [20] |
| F2                 | SSR          | 124            | 1096.2      | 8.84                  | 10       | Li, 2015 [21] |
| F2                 | SSR          | 181            | 803.13      | 4.38                  | 10       | Yu, 2018 [22] |
| F2                 | AFLP         | 444            | 1468.12     | 3.38                  | 10       | Shi, 2017 [23] |
| F2                 | SRAP, SSR    | 427            | 1273.4      | 2.98                  | 10       | Shi, 2018 [24] |
| RIL                | SNP          | 1065           | 1191.7      | 1.11                  | 10       | Jin, 2021 [29] |

https://doi.org/10.1371/journal.pone.0278153.t005
Recently, the genomes of sesame [44], *Brassica napus* [45], peanut [46], and soybean [47] were mapped using WGRS. Lin et al. [48] implemented WGRS to develop a high-density genetic map of wheat and identified 21,197 polymorphic SNP markers. In the present study, we used WGRS to sequence 150 F₂ plants and their parents. The sequencing data were compared with the reference genome, and many SNP markers were detected, greatly improving the map’s quality. After rigorous screening and filtering, 5,656 high-quality polymorphic SNPs were obtained, and a high-density genetic linkage map of sorghum-sudangrass hybrid was constructed. Compared to other published genetic linkage maps for sorghum-sudangrass hybrid, the one presented herein contains the most SNP markers and the highest marker density. Therefore, WGRS is an efficient and accurate strategy for marker development and construction of genetic maps on a large scale.

Several genetic maps were created for sorghum-sudangrass hybrid (Table 5), but important quantitative traits could not be accurately identified due to the limitations pertaining to the available molecular markers and low resolution of the maps. The earliest map comprised 158 AFLP and 8 RAPD markers, covering 836 cM of the sorghum-sudangrass hybrid genome and with the average marker spacing of 5.03 cM [20]. In the present study, we constructed an
ultra-high-genetic linkage map containing 5,656 SNP markers. Compared with previous maps, our maps are superior in terms of the mapping population size, high-throughput sequencing method, markers number, and sequencing data [20–24, 29]. The genome coverage of our map was greatly improved, which may be attributed to several reasons: (1) The large genetic diversity between the two parents used in this study contributed to their significant difference in several important traits (Fig 1). The research showed that parents with extreme differences in one or more phenotypic traits are ideal for mapping populations because it lays a foundation for QTL mapping of important traits and construction of high-density genetic maps [49]. (2) The sequencing depth was deep enough. To obtain sufficient SNP markers, we sequenced the parents and progeny at a depth of 28.77×, 29.71×, and 9.97×. (3) The WGRS approach allowed for a development of a large number of SNP markers in the genome. It increases map density, makes further QTL detection more accurate and accelerates the screening process of candidate genes related to important traits in plants. In addition, this map can provide high-quality references for future research on molecular genetics and genomics of sorghum-sudangrass hybrid.

High-resolution linkage mapping has become an indispensable part of genetic research. It can effectively identify important agronomic QTLs and plays important roles in fine mapping, candidate gene prediction, and marker-assisted breeding [13].

The high-density genetic map is a valuable tool to study plant phenotypic variation and shorten the breeding cycle. It is a bridge between traits and genomes, revealing the relationship between genes and phenotypic variation [50]. In recent years, with the increasing number of available complete genome sequences of crops, high-density genetic maps have been combined with crop genomes to identify QTL loci physically similar to the genome [51]. The high-density genetic maps have been widely used in QTL mapping, fine mapping of important traits, and candidate gene prediction in weed-rice [51], wheat [52], maize [53], rape [54], cotton [55], and others. In addition, high-density genetic maps can also be served as a platform for genome assembly or studying the collinearity between related species and the structural variation of hybrid genomes [40].

The ultra-high-density genetic map constructed in this study has the largest number and highest density of markers that is currently available for the sorghum-sudangrass hybrid. The population sample size directly affects the aquality and accuracy of genetic map. However, due to human and material resource limitations, selecting an appropriate population size is critical, with most studies using 100–200 progeny or individuals [56]. Our study developed an ultra-high-density genetic map with 5,656 SNP markers from 150 plants of an F₂ sorghum-sudangrass hybrid population. Next, we will apply these markers to a larger population to increase the map resolution. This will lay a foundation for fine mapping of QTLs for important traits in sorghum-sudangrass hybrid.

Conclusions

High-density genomic maps can optimize marker-assisted breeding in sorghum-sudangrass hybrid, which utilizes molecular markers linked to or separated from target traits. sorghum-sudangrass hybrid is an important annual gramineous forage that requires further improvement. Marker-assisted breeding is more effective when the genetic map is saturated. We have constructed an ultra-high-density linkage map for sorghum-sudangrass hybrid for the first time, utilizing the SNP markers developed using the WGRS. A map containing 5,656 SNP markers was constructed, covering an area of 2,192.84 cM with an average spacing between markers of 0.39 cM. sorghum-sudangrass hybrid breeding will be greatly improved with this ultra-high-density genetic map, which offers an improvement upon existing maps.
Furthermore, this study is of great importance for developing diploid sorghum-sudangrass hybrids. The application of WGRS will provide a valuable reference for other diploid graminoid crops.

**Supporting information**

**S1 Fig.** Haplotype maps for 10 linkage groups of the sorghum-sudangrass hybrid. From chromosomes 1 to 10, ten haplotype maps represent an individual’s genotype. Each row of numbers on the left represents a sample number. The horizontal axis represents the tag name. There are two chromatids from each parent in green and blue; missing data are shown in red and heterozygosity in purple. (TIF)

**S2 Fig.** Heat maps for 10 linkage groups of the density genetic map for the sorghum-sudangrass hybrid. Ten heat maps are shown from chromosome 1 to chromosome 10, in which markers are listed alphabetically by row and column. Different colors indicate the strength of linkage: yellow represents weak links, whereas red represents strong links. (TIF)

**S1 Table.** Statistics of population resequencing in sorghum-sudangrass hybrid. (XLS)

**S2 Table.** Population resequencing and coverage statistics. (XLS)

**S3 Table.** Statistics of parental SNP detection results. (XLS)

**S4 Table.** Statistics of parents and F₂ population aa × bb marker with 9× 3×. (XLS)

**S5 Table.** Statistics of parents and F₂ population aa × bb marker with 9× 4×. (XLS)

**S6 Table.** High density genetic map of sorghum-sudangrass hybrid. (XLS)

**S1 File.** Construction of ultra-high-density genetic linkage map of a sorghum-sudangrass hybrid. Also available on protocols.io. http://dx.doi.org/10.17504/protocols.io.14egn2226g5d/v1. (DOC)

**Author Contributions**

**Conceptualization:** Zhuo Yu.

**Data curation:** Qianqian Lu.

**Funding acquisition:** Xiaoxia Yu.

**Investigation:** Qianqian Lu, Huiting Wang, Xia Zhang, Yaqi Zhao.

**Methodology:** Qianqian Lu, Xiaoxia Yu.

**Software:** Qianqian Lu.

**Writing – original draft:** Qianqian Lu.
Writing – review & editing: Xiaoxia Yu.

References

1. Yu Z, Yamada T. Comparative research on main characteristics of pacetter varieties. Chinese Journal of Grassland. 2006; 28(6):1–6. https://doi.org/10.3321/j.issn.1673–5021.2006.06.001
2. Saballos A, Vermeris W, Rivera L, Ejeta G. Allelic association, chemical characterization and saccharification properties of brown midrib mutants of sorghum (Sorghum bicolor (L.) Moench). Bioenerg Res. 2008; 1(3):193–204. https://doi.org/10.1007/s12155-008-9025-7
3. Mahmoudzadeh VM, Oad R. Sorghum-sudangrass water productivity under subsurface drip irrigation. Irrig and Drain. 2018; 67(5):702–712. https://doi.org/10.1002/ird.2278
4. Han LY, Li J, Na RS, Yu Z, Zhou H. Effect of two additives on the fermentation, in vitro digestibility and aerobic security of sorghum-sudangrass hybrid silages. Grass Forage Sci. 2015; 70(1):185–194. https://doi.org/10.1111/1365-2495.12092
5. Peng JL, Quinby AP, Corbin MD, Bates GE. Forage mass, nutritive value, and economic viability of cowpea overseeded in tall fescue and sorghum-sudangrass swaards. Crop, Forage & Turfgrass Management. 2020; 6(1):20003. https://doi.org/10.1002/cft2.20003
6. Wells MS, Martinson KL, Sheander DJ, Sheaffer CC. A survey investigating alfalfa winter injury in Minnesota and Wisconsin from the winter of 2012–2013. Forage & Grazinglands. 2014; 12(1):1–7. https://doi.org/10.2134/FG-2013-0051-RS
7. Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. Euphytica. 2005; 142(1):169–196. https://doi.org/10.1007/s10681-005-1681-5
8. Nave RLG, Quinby MP, Griffith AP, Corbin MD, Bates GE. Forage mass, nutritive value, and economic viability of cowpea overseeded in tall fescue and sorghum-sudangrass swaards. Crop, Forage & Turfgrass Management. 2020; 6(1):20003. https://doi.org/10.1002/cft2.20003
9. Wells MS, Martinson KL, Undersander DJ, Sheaffer CC. A survey investigating alfalfa winter injury in Minnesota and Wisconsin from the winter of 2012–2013. Forage & Grazinglands. 2014; 12(1):1–7. https://doi.org/10.2134/FG-2013-0051-RS
10. Fang JY, Jing HC, Zhang WH. Welcome the era when grass husbandry becomes half of China’s modern agriculture. Chinese Science Bulletin. 2018; 63(17):1615–1618. CNKI:SUN:KXTB.0.2018-17-002.
11. Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. Euphytica. 2005; 142(1):169–196. https://doi.org/10.1007/s10681-005-1681-5
12. Zhu JC, Guo YS, Su K, Liu ZD, Ren ZH, Li K, et al. Construction of a highly saturated genetic map for Vitis by next-generation restriction site-associated DNA sequencing. BMC Plant Biol. 2018; 18(1):347. https://doi.org/10.1186/s12870-018-1575-z PMID: 30541441
13. Yu C, Wan HH, Peter MB, Cheng BX, Luo L, Pan HT, et al. High density genetic map and quantitative trait loci (QTLs) associated with petal number and flower diameter identified in tetraploid rose. J Integr Plant Biol. 2021; 20(5):1287–1301. https://doi.org/10.1007/s10215-021-00039
14. Kim M, Kim JY, Sung K. The optimum seeding and harvesting dates of sorghum-sudangrass hybrid (Sorghum bicolor (L.) Moench) via optimum moving response surface methodology. Grassl Sci. 2021; 67(1):3–11. https://doi.org/10.1111/grs.12279
15. Iptas S, Brohi AR. Effect of nitrogen rate and stubble height on dry matter yield, crude protein content and crude protein yield of a sorghum-sudangrass hybrid [Sorghum bicolor (L.) Moench × Sorghum sudanense (Piper) Stapf.] in the three-cutting system. J Agron Crop Sci. 2003; 189(4):227–232. https://doi.org/10.1046/j.1439-037X.2003.00001.x
16. Ni YQQ, Yu XX, Yu Z, Wu GF, Yang DS, Li JQ, et al. SSR analysis of varieties and new strains of the Mengnong series sorghum-sudangrass hybrid. Chinese Journal of Grassland. 2019; 9(5):1–6. https://doi.org/10.1111/grs.12357
17. Liang YS, Yan C, Qin XJ, Nan WB, Zhang HM. Construction of three half-sib SSR linkage maps derived from overwintering cultivated rice and segregation distortion loci mapping. Genome. 2020; 63(4):239–251. https://doi.org/10.1139/gen-2019-0160 PMID: 32053497
18. Cheng LR, Chen XC, Jiang CH, Ma B, Ren M, Cheng YZ, et al. High-density SNP genetic linkage map construction and quantitative trait locus mapping for resistance to cucumber mosaic virus in tobacco (Nicotiana tabacum L.). Crop J. 2019; 7(4):539–547. https://doi.org/10.1016/j.cj.2018.11.010
19. Guo F, Yu HW, Tang Z, Jiang XL, Wang L, Wang X, et al. Construction of a SNP-based high-density genetic map for pummelo using RAD sequencing. Tree Genetics & Genomes. 2015; 11(1):1–11. https://doi.org/10.1007/s11295-014-0831-0
20. Lu XP, Yun JF. Genetic mapping and gene localization of main agronomic characters of sorghum–sudangrass hybrid. Acta Agrestia Sinica. 2005; 3:262–263.

21. Li JQ, Wang LH, Zhan QW, Liu YL, Zhang Q, Li JF, et al. Mapping quantitative trait loci for five forage quality traits in a sorghum–sudangrass hybrid. Genet. Mol. Res. 2015; 14(4):13266–13273. https://doi.org/10.4238/2015.October.26.23 PMID: 26535640

22. Yu XX, Nan ZB, Yu Z, Yang DS, Xie R, Shi Y, et al. Construction of genetic linkage map based on SSR markers and identification of QTLs for main traits of sorghum–sudangrass hybrid. Molecular Plant Breeding. 2018; 16(15):5347–5358. https://doi.org/10.1007/s10018-018-1169-6 PMID: 31218798

23. Shi Y, Yu XX, Yu Z, Yang DS, Jiang C, Zhang MF. Construction of AFLP molecular genetic linkage map of sorghum–sudangrass hybrid. Journal of Northwest A & F University(Natural Science Edition). 2017; 45(10):7–14. https://doi.org/10.13207/j.carolcarrollnkijnwafu.2017.10.002

24. Yu XX, Nan ZB, Yu Z, Yang DS, Jiang C, Zhang MF. Construction of AFLP molecular genetic linkage map of sorghum–sudangrass hybrid. Journal of Northwest A & F University(Natural Science Edition). 2017; 45(10):7–14. https://doi.org/10.13207/j.carolcarrollnkijnwafu.2017.10.002

25. Wang N, Fang LC, Xin HP, Wang LJ, Li SH. Construction of a high-density genetic map for grape using next generation restriction-site associated DNA sequencing. BMC Plant Biol. 2012; 12(1):1–15. https://doi.org/10.1186/1471-2229-12-148 PMID: 22908993

26. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One. 2011; 6(5):e19379. https://doi.org/10.1371/journal.pone.0019379 PMID: 21573248

27. Xu LY, Wang LY, Wei K, Tan LQ, Su JJ, Cheng H. High-density SNP linkage map construction and QTL mapping for flavonoid-related traits in a tea plant (Camellia sinensis) using 2b-RAD sequencing. BMC Genomics. 2018; 19(1):95. https://doi.org/10.1186/s12864-018-5291-8 PMID: 30577813

28. Sun XW, Liu DY, Zhang XF, Li WB, Liu H, Hong WG, et al. SLAF-seq: an efficient method of large-scale De novo SNP discovery and genotyping using high-throughput sequencing. PLoS One. 2013; 8(3):e58700. https://doi.org/10.1371/journal.pone.0058700 PMID: 23527008

29. Jin P, Wang LH, Zhao WJ, Zheng J, Wang YH, Liu YL, et al. Construction of high density genetic map and QTL mapping in sorghum × sudangrass. Euphytica. 2021; 217(8):162. https://doi.org/10.1007/s10681-021-02895-9

30. Malmberg MM, Barbulescu DM, Drayton MC, Shinozuka M, Thakur P, Ogaji YO, et al. Evaluation and recommendations for routine genotyping using whole genome resequencing in Canola. Front Plant Sci. 2018; 9:1809. https://doi.org/10.3389/fpls.2018.01809 PMID: 30581450

31. Peng J, Chen LB, Tao LL, Liu SR, Wei CL. QTL mapping for leaf area of tea plants (Camellia sinensis) based on a high-quality genetic map constructed by whole genome resequencing. Chinese Journal of Grassland. 2018; 40(5):11–17. https://doi.org/10.16742/j.zgcdxb.2018-05-02

32. Peng J, Chen LB, Tao LL, Liu SR, Wei CL. QTL mapping for leaf area of tea plants (Camellia sinensis) based on a high-quality genetic map constructed by whole genome resequencing. Chinese Journal of Grassland. 2018; 40(5):11–17. https://doi.org/10.16742/j.zgcdxb.2018-05-02

33. Wang P, Zhou GL, Yu HH, Yu SB. Fine mapping a major QTL for flag leaf size and yield-related traits in rice. Theor Appl Genet. 2011; 123(8):1319–1330. https://doi.org/10.1007/s00122-011-1669-6 PMID: 21830109

34. Yamakawa H, Haque E, Tanaka M, Takagi H, Asano K, Shimosaka E, et al. Polyploid QTL-seq towards rapid development of tightly linked DNA markers for potato and sweetpotato breeding through whole-genome resequencing. Plant Biotechnol J. 2021; 19(10):2040–2051. https://doi.org/10.1111/pbi.13633 PMID: 34003333

35. Yu Z, Song CH, Yu XX, Shi Y, Jiang C, Fang YY. Breeding of new cultivar of sorghum bicolor × sorghum sudanese cv. Mengnong No.9 with super-low HCN content. Chinese Journal of Grassland. 2015; 37(6):14–21.

36. Kosambi DD. The estimation of map distances from recombination values. Ann Eugen. 1944; 12:172–175.

37. Mun JH, Chung H, Chung WH, Oh M, Jeong YM, Kim N, et al. Construction of a reference genetic map of Raphanus sativus based on genotyping by whole-genome resequencing. Theor Appl Genet. 2015; 128(2):259–272. https://doi.org/10.1007/s00122-014-2426-4 PMID: 25403353

38. Hoffmann AA, Rieseberg LH. Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? Annu Rev Ecol Evol Syst. 2008; 39(1):21–42. https://doi.org/10.1146/annurev.ecolsys.39.110707.173532 PMID: 20419035

39. Luo XB, Xu L, Wang Y, Dong JH, Chen YL, Tang MJ, et al. An ultra-high-density genetic map provides insights into genome synteny, recombination landscape and taproot skin colour in radish (Raphanus sativus L.). Plant Biotechnol J. 2020; 18(1):274–286. https://doi.org/10.1111/pbi.13195 PMID: 31218798
39. Yu LA, Ma XK, Deng B, Yue JJ, Ming R. Construction of high-density genetic maps defined sex determination region of the Y chromosome in spinach. Mol Genet Genomics. 2021; 296(1):41–53. https://doi.org/10.1007/s00438-020-01723-4 PMID: 32955620

40. Ji GS, Zhang QJ, Du RH, Lv P, Ma X, Fan S, et al. Construction of a high-density genetic map using specific-locus amplified fragments in sorghum. BMC Genomics. 2017; 18(1):51. https://doi.org/10.1186/s12864-016-3430-7 PMID: 28061813

41. Yang XP, Sood S, Glynn N, Islam MS, Comstock J, Wang JP. Constructing high-density genetic maps for polyploid sugarcane (Saccharum spp.) and identifying quantitative trait loci controlling brown rust resistance. Mol. Breed. 2017; 37(10):1–12. https://doi.org/10.1007/s11032-017-0716-7

42. Wang JH, Su K, Guo YS, Xing HY, Zhao YH, Liu ZD, et al. Construction of a high-density genetic map for grape using specific length amplified fragment (SLAF) sequencing. PLoS One. 2017; 12(7):e0181728. https://doi.org/10.1371/journal.pone.0181728 PMID: 28746364

43. Gu QS, Ke HF, Liu ZW, Lv X, Sun ZW, Zhang M, et al. A high-density genetic map and multiple environmental tests reveal novel quantitative trait loci and candidate genes for fibre quality and yield in cotton. Theor Appl Genet. 2020; 133(12):3395–3408. https://doi.org/10.1007/s00122-020-03676-z PMID: 32894321

44. Liang JC, Sun J, Ye YY, Yan XW, Yan TX, Rao YL, et al. QTL mapping of PEG-induced drought tolerance at the early seedling stage in sesame using whole genome re-sequencing. PLoS One. 2021; 16(2):e0247681. https://doi.org/10.1371/journal.pone.0247681 PMID: 33626101

45. Dong ZX, Alam MK, Xie ML, Yang L, Liu J, Helal MMU, et al. Mapping of a major QTL controlling plant height using a high-density genetic map and QTL-seq methods based on whole-genome resequencing in Brassica napus. G3 (Bethesda). 2021; 11(7):jkb118. https://doi.org/10.1093/g3journal/jkb118 PMID: 33836054

46. Sun ZQ, Qi FY, Liu H, Qin L, Xu J, Shi L, et al. QTL mapping of quality traits in peanut using whole-genome resequencing. The Crop Journal. 2022; 10(1):177–184. https://doi.org/10.1016/j.cj.2021.04.006

47. Tian Y, Yang L, Lu HF, Zhang B, Li YF, Liu C, et al. QTL analysis for plant height and fine mapping of two environmentally stable QTLs with major effects in soybean. J Integr Agr. 2022; 21(4):933–946. https://doi.org/10.1016/S2096-3119(21)63693-6

48. Lin Y, Jiang XJ, Hu HY, Zhou KY, Wang Q, Yu SF, et al. QTL mapping for grain number per spikelet in wheat using a high-density genetic map. The Crop Journal, 2021; 9(5):1108–1114. https://doi.org/10.1016/j.cj.2020.12.006

49. Semagn K, Bjørnstad A, Ndjonjdop MN. Principles, requirements and prospects of genetic mapping in plants. Afr J Biotechnol, 2006; 5(25):2569–2587.

50. Gong HG, Rehman F, Yang TS, Li Z, Zeng SH, Pan LZ, et al. Construction of the first high-density genetic map and QTL mapping for photosynthetic traits in Lycium barbarum L. Mol. Breed. 2019; 39(7):106. https://doi.org/10.1007/s11032-019-1000-9

51. Li F, Han ZY, Qiao WH, Wang JR, Song Y, Cui YX, et al. High-quality genomes and high-density genetic map facilitate the identification of genes from a weedy rice. Front Plant Sci. 2021; 12:775051. https://doi.org/10.3389/fpls.2021.775051 PMID: 34868173

52. Guo Y, Zhang GZ, Guo BJ, Qu CY, Zhang MX, Kong FM, et al. QTL mapping for quality traits using a high-density genetic map of wheat. PLoS One, 2020; 15(3):e0230601. https://doi.org/10.1371/journal.pone.0230601 PMID: 32208463

53. Ma LL, Guan ZR, Zhang ZT, Zhang XX, Zou CY, et al. Identification of quantitative trait loci for leaf-related traits in an IBM Syn10 DH maize population across three environments. Plant Breeding, 2018; 137(2):127–138. https://doi.org/10.1111/pbr.12566

54. Zhao X, Yu K, Pang C, Wu X, Shi R, Sun C, et al. (2021). QTL analysis of Five silique-related traits in Brassica napus L. across multiple environments. Front Plant Sci, 2021; 12:766271. https://doi.org/10.3389/fpls.2021.766271 PMID: 34887891

55. Han WL, Zhao JY, Deng XJ, Gu AX, Li DL, Wang YX, et al. Quantitative trait locus mapping and identification of candidate genes for resistance to fusarium wilt race 7 using a resequencing-based high-density genetic bin map in a recombinant inbred line population of Gossypium barbadense. Front Plant Sci, 2022; 13:815643. https://doi.org/10.3389/fpls.2022.815643 PMID: 35371113

56. Huang LB, Yan XH. Construction of a genetic linkage map in Pyropia yezoensis (Bangiales, Rhodophyta) and QTL analysis of several economic traits of blades. PLoS One. 2019; 14(3):e0209128. https://doi.org/10.1371/journal.pone.0209128 PMID: 30849086