Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A multi-reference parent nested-association mapping population to dissect the genetics of quantitative traits in durum wheat

Samir Alahmad (s.alahmad@uq.edu.au)
The University of Queensland https://orcid.org/0000-0003-0136-1740
Yichen Kang
The University of Queensland
Eric Dinglasan
The University of Queensland
Dilani Jambuthenne
The University of Queensland
Hannah Robinson
The University of Queensland
Yongfu Tao
The University of Queensland
Jason A. Able
The University of Adelaide School of Agriculture Food & Wine
Jack Christopher
The University of Queensland
Kai P. Voss-Fels
The University of Queensland
Filippo M. Bassi
ICARDA: International Center for Agricultural Research in the Dry Areas
Lee T. Hickey
The University of Queensland

Research Article

Keywords: Flowering time, GWAS, LD decay, NAM, plant height, structured population

Posted Date: August 12th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1869638/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

Durum wheat (*Triticum turgidum L.*) breeding programs face many challenges surrounding the development of stable varieties with high quality and yield. Therefore, researchers and breeders are focused on deciphering the genetic architecture of biotic and abiotic traits with the aim of pyramiding desirable traits. These efforts require access to diverse genetic resources, including wild relatives, germplasm collections, and mapping populations. Advances in accelerated generation technologies have enabled the rapid development of mapping populations with significant genetic diversity. Here, we describe the development of a durum Nested Association Mapping (dNAM) population, which represents a valuable genetic resource for mapping the effects of different alleles on trait performance. We created this population to understand the quantitative nature of drought-adaptive traits in durum wheat. We developed 920 F_6 lines in only 18 months using speed breeding technology, including the F_4 generation in the field. Large variation in above- and belowground traits was observed, which could be harnessed using genetic mapping and breeding approaches. We genotyped the population using 13,393 DArTseq markers. Quality control resulted in 6,785 high-quality polymorphic markers used for structure analysis, linkage disequilibrium decay, and marker-trait association analyses. To demonstrate the effectiveness of dNAM as a resource for elucidating the genetic control of quantitative traits, we took a genome-wide mapping approach using the FarmCPU method for plant height and days to flowering. These results highlight the power of using dNAM as a tool to dissect the genetics of durum wheat traits, supporting the development of varieties with improved adaptation and yield.

Introduction

Durum wheat (*Triticum turgidum* L. subsp. *durum* Desf.) contains two sub-genomes, A and B ($2n = 4x = 28$), and its high protein content makes it an ideal ingredient for pasta. Durum wheat is a major staple food in Mediterranean regions, including Italy, Spain, Morocco, Algeria, Tunisia, Turkey, the Middle East, and Ethiopia, and it is widely produced as a cash crop in North and South America and India. Recent evidence indicates that the origins of all cereals including durum wheat trace back to the Fertile Crescent in southeastern Turkey and northern Syria (Lev-Yadun et al. 2000). Durum wheat is one of the earliest cultivated cereals, resulting in a genetic pool rich in phenotypic and genotypic variation; such variation is essential for breeding (Mengistu et al. 2016). The global demand for durum wheat has increased dramatically in recent years due to its high sensitivity to changes in climate, including severe drought and heat, in key durum production regions of western Canada and north Africa (Glasser et al. 2022), the global pandemic, and the recent war in Ukraine. The challenge in meeting this demand is complicated by climate variability, with unprecedented high temperatures and the increased frequency and duration of drought in Europe already negatively affecting crop production (Hochman et al. 2017). To overcome this challenge, researchers seek to better understand the genetic architecture of biotic and abiotic traits and to empower breeding programs to assemble these traits for the development of resilient durum wheat varieties.

Recent advances in sequencing technology have driven down costs, enabling researchers to access a wealth of resources and tools to better dissect the architecture of complex traits, such as drought adaptation. A key example is the sequencing of the reference genome of the modern durum wheat variety Svevo (Maccaferri et al. 2019). However, to effectively harness the potential benefits of these new genomic tools, rich allelic diversity is required. Yet, constraints associated with the genetic diversity and resolution of available durum wheat mapping populations could limit the potential benefits of deploying these new genomic tools (Kidane et al. 2019). Diversity panels are typically deployed for marker-trait association studies due to a high level of historical recombination. Unfortunately, however, such panels lack the statistical power that classical pedigree-based mapping provides (Mackay et al. 2009). Novel structured mapping populations such as Multi-parent Advanced Generation Inter-Cross (MAGIC; Mackay et al. 2014; Milner et al. 2016) and Nested Association Mapping (NAM) can provide a high degree of genetic diversity in addition to a dramatic increase in allelic recombination frequencies (Ladejobi et al. 2016).

A NAM structure is based on crossing a panel of donor lines with a number of reference lines (often varieties or an elite lines) and then producing recombinant inbred lines (RILs) from each donor and reference line cross (Guo et al. 2013). The design of such structure introduces chromosomal diversity through recombination in the form of ‘mosaic’ segments from each donor line sharing the same reference background. The NAM structure combines the advantages of well-described bi-parental and association mapping approaches while minimizing the disadvantages of each approach (Yu et al. 2008). The advantages of a NAM

population include (but are not limited to) high allelic richness, high statistical power, a low number of markers required for whole-genome scanning, and high mapping resolution (Buckler et al. 2009).

NAM was first applied by Yu et al. (2006) to study the effects of many quantitative trait loci (QTLs) controlling flowering time in maize. Since then, NAM populations have been developed for many strategic crops, including sorghum (Mace et al. 2013; Bouchet et al. 2017; Perumal et al. 2021), wheat (Richard et al. 2015; Chidzanga et al. 2021), barley (Ziems et al. 2017), rice (Fragoso et al. 2017; Kitony et al. 2021), soybean (Guo et al. 2013; Maranna et al. 2019), and Ethiopian durum wheat (Kidane et al. 2019). The Ethiopian durum NAM population (EtNAM) combines 49 varieties developed by Ethiopian farmers with one internationally representative modern variety to bridge the gap between farmer-developed and modern varieties while generating rich allelic diversity for biotic traits (Kidane et al. 2019). Although EtNAM is an invaluable resource, it was created to explore the genetics of disease and adaptation to Ethiopian production regions. This created the need for a more globally representative NAM population targeting Mediterranean and similar water-limited growing environments.

Here, we describe the development of a durum NAM (dNAM) population that combines the genetics of ICARDA elite donor lines carrying desirable drought-adaptive traits with two Australian varieties. The population segregated for plant height (PH), days to flowering (DTF), and several drought-adaptive traits, including above- and -below-ground water use traits, which can be explored using mapping approaches. We genotyped the entire dNAM population using 13,395 DArTseq markers on the DArTseq genotyping-by-sequencing platform, allowing the characterisation of population diversity, population structure, and linkage disequilibrium (LD). To highlight the usefulness of dNAM, we performed a genome-wide association study (GWAS) for two commonly phenotyped agronomic traits, PH and DTF, and detected several marker-trait associations. The novel genomic resource developed in this study will be highly valuable for dissecting complex traits, especially those that contribute to improved drought adaptation, which should facilitate the breeding of climate-resilient durum wheat varieties.

Materials And Methods

Selection of durum NAM parental lines

The dNAM population was generated by crossing ICARDA elite donor lines with two Australian durum varieties, Jandaroi and DBA Aurora, which served as reference varieties (Fig. 1a). A description of the parental lines, including pedigree information, is provided in Table 1. Selection of genotypes from ICARDA's germplasm pool in Morocco was based on three decades of detailed characterisation of grain yield, agronomic traits, and biotic and abiotic traits in Mediterranean water-limited environments (Table 2). The elite durum donor lines that were imported into Australia include Fastoz2, Fastoz3, Fastoz6, Fastoz7, Fastoz8, Fastoz10, Outrob4, and Fadda98. Furthermore, the ICARDA lines have been used as parents in breeding programs targeting marginal rainfall regions of West Asia and North Africa. The two Australian durum commercial reference varieties DBA Aurora and Jandaroi are widely grown and are preferred varieties of Australian farmers. This is largely due to their high yield potential and protein content and their grain qualities, which are preferred by the industry.

Table 1
List of parental lines used to develop the durum Nested Association Mapping population, including their pedigrees and origins.

Genotype name	Ploidy	Origin	Pedigree
DBA-Aurora	Tetraploid	Australia	Tamaroi*2/Kalka//RH920318/Kalka///Kalka*2/Tamaroi
Jandaroi	Tetraploid	Australia	110780/111587
Outrob4	Tetraploid	ICARDA	Ouassel1/4/GdoVZ512/Cit//Ruff/Fg/3/Pin/Gre//Trob
Fadda98	Tetraploid	ICARDA	Awl2/Bit
Fastoz2	Tetraploid	ICARDA	T.polonicumTurkeyIG45272/6/ICAMORTA0463/5/Mra1/4/Aus1/3/Scar/GdoVZ579//Bit
Fastoz3	Tetraploid	ICARDA	Msbl1//Awl2/Bit/3/T.dicoccoidesSYRIG117887
Fastoz6	Tetraploid	ICARDA	Azeghar1/6/Zna1/5/Awl1/4/Ruff//Jo/
Fastoz7	Tetraploid	ICARDA	CandocrossH25/Ysf1//CM829/CandocrossH25
Fastoz8	Tetraploid	ICARDA	MorlF38//Bcrch1/Kund1149/3/Bicrederaa1/Miki
Fastoz10	Tetraploid	ICARDA	Younes/TdicoAlpCol//Korifla

Table 2
List of the 10 families and their corresponding cross-IDs, the number of available recombinant inbred lines, and the segregating traits for each family.

Family	Cross-ID	No. of RILs	Traits of interest based on information about the donor lines
1	DBA Aurora × Fastoz7	92	Drought-adaptive traits, Hessian fly, Root rot
2	DBA Aurora × Outrob4	92	Drought-adaptive traits, Soil-borne disease (root rot, fusarium), Leaf rust
3	DBA Aurora × Fastoz 8	92	Drought-adaptive traits, Yield potential, Leaf rust
4	DBA Aurora × Fadda98	92	Drought-adaptive traits, Tan spot, Soil-borne disease (root rot, fusarium)
5	DBA Aurora × Fastoz3	92	Drought-adaptive traits, Vitreousness, Protein content, Tan spot
6	Jandaroi × Fastoz8	92	Drought-adaptive traits, Yield potential, Leaf rust
7	Jandaroi × Fastoz10	92	Drought-adaptive traits, Yield, Leaf rust, Tan spot
8	Jandaroi × Fastoz6	92	Drought-adaptive traits, Gluten strength, Stem rust and root rot
9	Jandaroi × Fastoz2	92	Drought-adaptive traits, Yellow pigments, Leaf rust, Tan spot
10	Jandaroi × Outrob4	92	Drought-adaptive traits, Soil-borne disease (root rot, fusarium), Leaf rust

Development of the durum NAM population

The dNAM population was generated in the speed breeding (SB) facility at The University of Queensland, St Lucia, Brisbane, QLD, Australia, with one generation progressed in the field (Fig. 1b). SB allows six plant generations of spring durum wheat to be grown in a single year (Watson et al. 2018, Alahmad et al. 2022). The parental lines were initially grown under SB conditions, and at flowering stage, eight ICARDA donor lines were crossed with the two Australian reference varieties. The crossing scheme, which is shown in Fig. 1a, resulted in the production of 10 families. The first five families shared DBA Aurora as the common parent, and the remaining five families shared Jandaroi as the common parent. Two donor lines (Outrob4 and Fastoz8) were crossed to both reference varieties. The number of recombinant inbred lines (RILs) per family, their cross-IDs, and the traits carried by the donor lines are listed in Table 2. Following crossing, F_1 seeds were planted under SB conditions, and 500 F_2 seeds per family were

harvested. The F_2 segregating populations were re-sown and progressed under SB using the single seed descent (SSD) method (Fig. 1b). The F_4 plants were grown in single 2-m-long rows under field conditions at The University of Queensland, Gatton Research Unit (27°32’34”S; 152°19’59”E) to generate 50–100 g of bulk F_4-derived seed. Plant height (PH) and days to flowering (DTF) were recorded on F_4-derived rows grown in the field. Ninety-two lines per family were subsequently selected based on agronomic appearance and a broad representation of genetic diversity within families. A single spike per field row was selected, harvested, and progressed in the SB system for two more generations to generate F_6 homozygous lines.

Phenotypic data used in the association analysis included the arithmetic mean values for the two PH measurements/genotype/row and the single value/genotype for DTF. An individual was considered flowering when 50% of the spikes on the main stem reached anthesis in the row of plants. The F_4 generation, which was evaluated in the field at Gatton, displayed large variation for PH and flowering time in both reference variety backgrounds (Fig. 2a).

Genotyping method and marker quality control

Leaf tissue samples for the 920 introgression lines and the 10 parental lines were sampled from F_6 generation plants grown under SB conditions. Leaf samples were freeze-dried, and DNA was extracted from the samples using Diversity Arrays Technology (DArT) protocols. The 10 families were then genotyped using the DArTseq platform. Genotyping resulted in 13,395 DArTseq markers, which were ordered according to their genetic positions in the wheat consensus map (version 4.0) provided by DArT, Canberra, Australia. High-quality markers were filtered using the R package “SelectionTools” (http://cran-r.uni-giessen.de/~user/). Quality control and filtering thresholds were applied. For example, markers with a minimum gene diversity frequency of $\leq 10\%$ and genotypes with $\leq 20\%$ missing marker information were removed, and the remaining markers were imputed for missing marker calls using Beagle v4.1 (Browning and Browning, 2016). A total of 6,785 high-quality DArTseq markers and 926 individuals were used in subsequent analyses.

Population structure and linkage disequilibrium decay

All population structure analyses were conducted using the R package “SelectionTools.” To investigate population structure in the dNAM panel, genetic distance was calculated with pairwise Roger’s distances between all genotypes using the 6,785 high-quality markers. Based on the resulting distance matrix, a principal coordinate analysis-based k-means clustering approach was performed, assuming k = 15 subgroups as a starting point. To determine the best number of clusters, the R package “NbClust version 3” was used. This package relies on 30 inbuilt indices as well as multiple factors including distance measures, the number of clusters, and the number of indices used in the package (Charrad et al. 2014). The appropriate number of clusters was confirmed based on the highest number of indices, for which the first three principal components were considered. Based on the crossing scheme, the dNAM populations were divided into two main hierarchical clusters representing the genetic backgrounds of the two Australian reference varieties (DBA Aurora and Jandaroi). To visualise the results, a genetic distance matrix was constructed by plotting principal coordinates (Fig. 3), and a heat map of genetic relatedness and NJ trees were constructed based on the clustering results (Fig. 4a, 4b).

Linkage disequilibrium (LD) decay analysis was performed for the entire dNAM population, as well as analysis of decay within each of the 10 dNAM families. Pearson correlation coefficients (r^2) for pairs of adjacent markers along the durum genome were calculated for 6,785 markers to determine the level of decay due to recombination. Locally estimated scatterplot smoothing (LOESS) curves were constructed to display the level of LD decay within each family and across the entire dNAM population, with the genetic distance shown in megabase pairs. Marker pairs with a genetic distance exceeding 50 Mbp were considered unlinked and eliminated from the analysis (Roncallo et al. 2021).

Marker–trait associations for plant height and days to flowering

Marker–trait associations (MTAs) between individual polymorphic markers and each phenotype, PH, and DTF were identified using the 6,785 quality-controlled DArTseq markers. GWAS was conducted in GAPIT Version 3 (Wang et al. 2021) using the Fixed and Random model Circulating Probability Unification (FarmCPU) model. This model, which was selected due to its high statistical power in detecting significant MTAs, enables efficient computation, eliminates confounding factors, and is significantly superior in avoiding false positives compared to mixed linear model approaches. FarmCPU implements two separate models that run and update iteratively. The fixed model handles false positive associations by testing one marker at a time and considering the rest of

the associated markers as covariates. The random model uses the associated markers to obtain the kinship, which is included as a covariate in the model to account for the relationship between the tested lines in the model. This eliminates the risk of overfitting while controlling for false positives (Liu et al. 2016). The fixed effect component of the FarmCPU model can be displayed as

$$y = M + QTN1 + QTN2 + \ldots QTNn + [PCs] + e$$

The random effect component of the FarmCPU model can be displayed as

$$y = Q[PCs] + K + e$$

Where Y is the vector of a phenotype, M is the matrix of marker information, QTN is a quantitative trait nucleotide, Q is the population structure using PCA, K is the kinship matrix, and e is the residual error.

Three principal components were used as covariates to control for population structure in the model. Model fitness was evaluated using quantile-quantile plots between expected and observed −log10(p) values. The significance p-value threshold for the marker-trait association was set to (−log10(p) ≥ 3), corresponding to a significance level of $p \leq 0.001$.

Results

Trait variation among durum NAM lines under field conditions

Plant height

Following crossing and the generation of F_1 seeds, we observed segregation for several traits during SSD under SB conditions from F_1 through to F_4, which was expected due to the occurrence of multiple recombination events during self-pollination. The F_4:F_5 dNAM lines grown under field conditions displayed a high degree of variation for PH among the lines within each family, across families with a common reference parent, and across families which did not share the reference parent. For instance, PH ranged from 49–106 cm for families derived from DBA Aurora as a common reference variety and 49–96.5 cm for families derived from Jandaroi. The highest degree of within-family variation was observed among Family 4 lines derived from the cross between DBA Aurora (PH = 68 cm) and Fadda98 (PH = 79 cm). Plant height for this family ranged from 55.5–106 cm, suggesting transgressive segregation for the PH trait, with several lines showing taller or shorter PH than the respective parents (Fig. 2b). The lowest degree of segregation was observed in Family 9, as the individuals derived from the cross between Jandaroi (56 cm) and Fastoz2 (67.3 cm) displayed PH values ranging from 55.5–75.5 cm, with several plants slightly taller than the donor parent (Fig. 2b).

Days to flowering

We obtained phenotypic data for DTF by subtracting the flowering date from the sowing date for the F_4:F_5 generation in the field. A reasonable degree of variation in DTF was observed among the dNAM lines across all families, ranging from 79–106 days. In families derived from DBA Aurora (DTF = 93 days), DTF ranged from 80–106 days, and in families derived from Jandaroi (DTF = 88 days), DTF ranged from 79–94 days. Family 3, derived from a cross between DBA Aurora and Fastoz8 (DTF = 85 days), and Family 4, derived from a cross between DBA Aurora and Fadda98 (DTF = 92 days), displayed high degrees of transgressive segregation, with DTF values of 80–98 days and 81–106 days, respectively. The least within-family variation for DTF was observed in Families 1, 6, and 9, with DTF values of 85–95, 80–91, and 79–90 days, respectively. The across-family differences for DTF (i.e., least variation and most variation for DTF) for the entire dNAM population was 15 days.

The overall correlation between PH and DTF among individuals of the entire dNAM under field conditions was $r = 0.2$ ($p < 0.01$). However, the correlation between the two traits varied within each family. For instance, Families 2, 3, 4, 5, and 7 displayed no association or a very weak association, with Pearson's correlation coefficients ranging from $r = 0.01$ to 0.16. A significant negative correlation for PH and DTF was observed for individuals in Family 1 ($r = -0.2$; $p < 0.01$). Finally, significant positive correlations were observed for both traits between individuals within Families 10 and 8 ($r = 0.27$; $p < 0.01$ and $r = 0.44$; $p < 0.001$, respectively; Supplementary Table 1).

Genetic relatedness and linkage decay

We performed quality control on the 13,393 DArTseq markers and retained 6,785 high-quality polymorphic markers. Thereafter, markers were physically mapped to both sub-genomes of Svevo durum wheat (Maccaferri et al. 2019). Of these DArTseq markers, 3,171 were distributed in the A genome and 3,708 were distributed in the B genome, with the number of markers per chromosome ranging from 225 (1A) to 803 (7A) for sub-genome A and from 295 (4B) to 708 (7B) for sub-genome B. We used these markers for subsequent genetic analyses, including calculating genetic distance using the pairwise Roger's distance approach. We determined that additive genetic variance was explained by the first 10 principal components (PC1 = 31.8%, PC2 = 9.6%, PC3 = 7.8%, and PC4 to PC10 ranging from 6.0–1.49%). The large variation explained by the first three PCs accounted for 49.2% of the genetic variability.

The optimum number of clusters for the 926 genotypes was 3, as determined by k-means clustering based on seven indices (Fig. 3). The first cluster consisted of individuals that shared Jandaroi and DBA Aurora as reference parents, as well as individuals with the same ICARDA elite donor parents (Outrob4 and Fadda98; n = 183 individuals). The second cluster included individuals that shared Jandaroi as a reference parent (n = 282 individuals), and the third cluster included individuals that shared DBA Aurora as a reference parent (n = 462 individuals). We calculated the genetic relationships between individuals across dNAM, which allowed us to discriminate between individuals from different genetic backgrounds in the form of a heatmap (Fig. 4a) and a neighbour-joining unrooted dendrogram (Fig. 4b). This resulted in the clustering of five families in the background of each reference variety. The LD of the entire dNAM population quickly decayed and reached a threshold of r^2 = 0.2 at a distance of ~ 3 Mbp and r^2 = 0.1 at a distance of ~ 11 Mbp. Variation in LD decay was evident among dNAM families and reached a threshold of r^2 = 0.2 at ~ 8 Mbp (Family 1) to ~ 13 Mbp (Family 2) (Fig. 5).

GWAS for plant height and anthesis time for the durum NAM lines

Based on quantile-quantile plots between expected and observed −log10(p) values with a cut-off threshold (−log10(p) ≥ 3), we identified 35 significant MTAs for PH and 20 for DTF (Fig. 6). The identified MTAs corresponded to 22 and 13 unique QTLs for PH and DTF, respectively (Supplementary Table 2).

These QTLs were distributed across the durum genome, with QTLs for PH distributed on 11 chromosomes and QTLs for DTF distributed on 10 chromosomes. Notably, for PH, six QTLs detected in this GWAS were also identified in previous studies, including the *RhtB1* gene (Mazucotelli et al. 2020; Wang et al. 2020; Zhang et al. 2021). In particular, marker '2371505' had the most significant association (−log10(p) = 30.9) with PH and was aligned with *RhtB1*(Mazucotelli et al. 2020). However, GWAS for PH using the dNAM populations revealed 16 novel QTLs underpinning variation in PH across dNAM lines. Similarly, six QTLs detected for DTF in this study were also identified in previous studies, including the *Ppd-A1*, *TaGI*, *VRN1*, and *VRN3* genes (Shi et al. 2019; Gupta et al. 2020; Semagn et al. 2021; Mangini et al. 2021). Interestingly, marker '982956' had the most significant association (−log10(p) = 25.8) with DTF and was precisely aligned with *Ppd-A1* (Mangini et al. 2021). However, six novel QTLs for DTF were also detected in durum wheat.

Discussion

Here, we rapidly developed dNAM, the first durum wheat NAM population, via SB. We developed this population to dissect the genetics of target traits and enhance adaptation to Mediterranean environments. We genotyped this population using 13,393 DArTseq markers. dNAM exhibits significant trait variation and diversity, with huge variation in abiotic stress traits for durum wheat. We observed a positive correlation between PH and DTF, which was also reported by Ruan et al. (2020). This observation indicates that partially shared genetic regions underpin both traits, which we detected on chromosomes 1B, 7B, 2A, 4A, and 5A; this conclusion was also described in previous reports (Ruan et al. 2020; Supplementary Table 2). The structure and statistical power of this population make it ideal for mapping the underlying genetic architecture of important traits. This notion was confirmed by our identification of previously well-described QTLs and genes for PH and DTF, as well as the discovery of novel QTLs for both traits.

We provided insights into the rapid development of the population, the genetic makeup of dNAM, its population structure, LD decay, and the relatedness among individuals. To demonstrate the power of dNAM to detect significant MTAs, we mapped two important agronomic traits using FarmCPU. The genetic architecture of countless numbers of traits could be explored using this

technique. Such analysis would enable the efficient identification and discovery of novel QTLs and the rapid introgression of these genes into adapted materials using trait pyramiding backcrossing schemes, thereby enhancing the productivity and end-use grain quality of modern durum varieties. The dNAM population represents a valuable resource for durum wheat breeders and geneticists.

Durum NAM population development and diversity

Genetic variation allows plant breeders to increase production more sustainably with crops that have enough plasticity to withstand climate change and evolving pathogens. Plant breeding requires ongoing cycles of selection and the progression of the best performing lines to increase production and help meet the demands of several billion people for high-quality food (Tester and Langridge, 2010). A powerful pre-breeding approach is to develop new sources of genetic variation such as NAM populations. The concept of creating variation using a nested crossing structure has been used for many winter cereals (Richard et al. 2015; Ziems et al. 2017; Kitony et al. 2021; Kidane et al. 2019), pulses (Maranna et al. 2019), and oilseed crops (Hu et al. 2018) to investigate the genetic effects of complex traits that are desired by breeders. However, developing such useful populations is challenging due to the time required to complete multiple generation cycles to create inbred lines (Alahmad et al. 2022). Consequently, in the current study, we used SB to develop the first durum wheat NAM population. SB permits rapid generation cycling using a prolonged photoperiod and controlled temperatures. The significantly reduced time required for a single generation allows RILs to be developed in a reasonably short period of time. For instance, six generations of durum wheat can be produced in a single year using protocols detailed previously (Ghosh et al. 2018; Watson et al. 2018).

In the current study, a significant level of variation was explained by the first two PCs, likely due to the genetic background shared by the two reference parents. The remaining PCs explain the variability across different families in the background of each reference variety, DBA Aurora and Jandaroi. The population structure and genetic relatedness between dNAM lines were as expected, with clear population stratification. The genomic relationship matrix perfectly matched the crossing scheme, which involved five families within each reference parent, as shown in Fig. 4. Indeed, previous NAM populations developed for durum and bread wheat showed a clear structure and hierarchical relationship matching the designed crossing schemes (Kidane et al. 2019; Chidzanga et al. 2021). Based on genetic linkage information, we calculated the genome-wide LD decay across the 10 dNAM families. The LD of the dNAM population quickly decayed and reached a threshold of r^2 = 0.2 at an ~3 Mbp distance compared to the average LD decay distance of approximately 7.4 Mbp across the EtNAM (Kidane et al. 2019). Analysing global LD decay provides insight into the genetic diversity of a population due to the recombination of LD blocks for key drought-adaptive and developmental traits, thereby assisting plant breeders apply selection within each population (Voss-Fels et al. 2015).

Durum NAM is a valuable genetic resource

The value of using dNAM populations lies in the significant range of genetic variation generated in a sufficiently large population, which is likely due to the reshuffling of the donor genes in the reference variety backgrounds. This allows for the identification of gene variants with the greatest potential for improving complex traits such as yield and quality. NAM populations are especially useful when paired with high-throughput phenotyping and screening techniques (Perumal et al. 2021). For instance, we previously identified a QTL for seminal root angle (*qSRA-6A*) with major effects under controlled conditions (Alahmad et al. 2019) using the high-throughput 'clear pot' method, which is suitable for screening small grains such as winter cereals (Richard et al. 2015*;* Robinson et al. 2016; Alahmad et al. 2018). We also successfully screened another subset of the dNAM population for crown rot disease under field conditions, resulting in the identification of QTLs underpinning tolerance to crown rot (*qCR-6B*), staygreen traits (*qSG-integral-6B*, *qSG-0.1-6B*), and PH (*qPH-6B*), as reported by Alahmad et al. (2020).

In the current study, we performed a GWAS to identify genomic regions controlling two phenological traits, PH and DTF. To highlight the power of mapping dNAM populations to dissect the genetic basis of complex traits, we used the FarmCPU model to detect significant QTLs associated with PH and DTF. We discovered 13 and 22 QTLs that were significantly associated with DTF and PH, respectively. The GWAS revealed the complex genetic architecture and quantitative nature of DTF and PH, as the detected QTLs were distributed across the durum wheat genome. Interestingly, the QTLs identified using dNAM for PH on chromosomes 1B, 2B, 4A, 5B, 6A, 6B, 7A, and 7B were also reported in previous studies (Milner et al. 2016). Also, the most significant QTL associated

with DTF in this study is aligned with the recently reported flowering genes *ppd-1B*, *Vrn1*, and *Vrn2*, which were previously mapped to chromosomes 2A, 2B, and 5A, respectively (Gupta et al. 2020). In addition to the previously reported QTLs for PH and DTF, six novel QTLs controlling DTF and six for PH were detected using the dNAM population developed in our study. Notably, dNAM has also been successfully used to map key drought-adaptive traits such as canopy dynamics and root system architecture (Alahmad et al. 2019; Kang et al. 2021). Perhaps the QTLs reported for above- and belowground traits using the dNAM populations developed in this study could be evaluated in multiple environments and used as markers for assisted breeding and introgression into elite germplasm. This would provide insights into the genetic architecture of a trait and GxE interactions, empowering plant breeders with knowledge about how to strategically select for these traits in durum breeding programs.

Future applications and material availability

The dNAM developed in this study is a structured population based on 10 donor ICARDA elite lines that were crossed to two Australian reference parents. The progenies were forwarded in a series of six generations of selfing until homozygosity. In this population, the entire genomes of the donor lines were re-shuffled in the form of chromosomal blocks into the background of the shared reference parent. While mapping traits by analysing the whole dNAM population would be desirable, it would be more cost effective to analyse the 10 parental lines for a trait of interest. Based on the results, sub-populations derived from donors carrying the desirable trait could then be evaluated, enabling a more targeted mapping approach (Alahmad et al. 2019; Alahmad et al. 2020; Kang et al. 2021). This approach would be particularly suitable for time-consuming assays or field experiments that are expensive and require large numbers of genotypes. The evaluated subset for the trait of interest could subsequently serve as a useful resource and be used as a training population for other breeding tools such as genomic selection. This would enable the fast-tracking and pyramiding of traits (including quantitative traits such as grain quality) into future varieties using genome-wide marker information (Hayes et al. 2009).

While we described the development of 920 dNAM lines, the population could be rapidly expanded by performing additional crosses to the reference parents via speed breeding, allowing additional genetic variation to be harnessed. The dNAM population could also be expanded by targeting traits of interest such as a disease. Individual donor lines could be selected from gene banks using the Focused Identification of Germplasm Strategy (FIGS) method (Endresen et al. 2012). We would like to use the dNAM population to establish collaborations with researchers and pre-breeders targeting adaptive traits and seeking to mine unique genetic diversity that combines Australian and ICARDA genetic pools. Seeds from the 920 dNAM lines and their parental lines, as well as the GBS DArTseq marker data, can be obtained upon request from the corresponding authors. Please direct requests to The University of Queensland, St Lucia, QLD, Australia or ICARDA, Rabat, Morocco using a Standard Material Transfer Agreement (SMTA).

Declarations

Ethics approval and consent to participate

Not Applicable.

Consent for publication

The authors give the consent for publication.

Competing interests

The authors declare no conflicts of interest.

Availability of data and materials

Data and materials are available upon request

Funding

The development of dNAM was made possible by funds provided by Monsanto's Beachell-Borlaug International Scholar Program (MBBISP) and Research Scholarship funds for a PhD program provided by The University of Queensland (UQRS) 2015 - 2019. Genotyping of the NAM population was Funded by Dr Filippo Bassi from ICARDA in Morocco and Ayed Al-Abdallat from Jordon University through a Food and Agriculture Organization grant (FAO; grant LoA/TF/W3B-PR-02/JORDAN/2016/AGDT).

Author Contributions:

SA, LH, and FMB conceived and designed the scheme for dNAM population development; SA and LH performed the crosses and developed the dNAM population. SA collected phenotypic data. LH and FMB genotyped the dNAM population. ED provided advice about structure and diversity analyses; SA, YK, ED, DJ, and HR drafted the original manuscript; all authors revised the manuscript. All authors read and approved the final manuscript.

References

1. Alahmad S, Dinglasan E, Leung KM, Riaz A, Derbal N et al. (2018) Speed breeding for multiple quantitative traits in durum wheat. Plant Methods. 1:1-15, https://doi.org/10.1186/s13007-018-0302-y
2. Alahmad S, El Hassouni K, Bassi FM, Dinglasan E, Youssef C et al. (2019) A major root architecture QTL responding to water limitation in durum wheat. Front Plant Sci. 10:10-436. https://doi.org/10.3389/fpls.2019.00436
3. Alahmad S, Kang Y, Dinglasan E, Mazzucotelli E, Voss-Fels KP et al. (2020) Adaptive traits to improve durum wheat yield in drought and crown rot environments. Int J Mol Sci. 15:5260. http://doi.org/10.3390/ijms21155260
4. Alahmad S, Rambla C, Voss-Fels KP, Hickey LT (2022) Accelerating breeding cycles. In Reynolds MP, Braun H (ed) Wheat improvement - Food security in a changing climate. Springer International Publishing, Cham, 2022. pp. 557-571.
5. Bouchet S, Olatoye MO, Marla SR, Perumal R, Tesso T et al. (2017) Increased power to dissect adaptive traits in global sorghum diversity using a nested association mapping population. Genet. 206:573-585. https://doi.org/10.1534/genetics.116.198499
6. Browning BL, Browning SR (2016) Genotype imputation with millions of reference samples. Am J Hum Genet. 1:116-126. https://doi.org/10.1016/j.ajhg.2015.11.020
7. Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ et al. (2009) The genetic architecture of maize flowering time. Sci. 325:714-718. https://doi.org/10.1126/science.1174276
8. Charrad M, Ghazzali N, Boiteau V, Niknafs A (2014) NbClust: An R package for determining the relevant number of clusters in a data set. J Stat Softw. 6:1-36. https://doi.org/10.18637/jss.v061.i06
9. Chidzanga C, Fleury D, Baumann U, Mullan D, Watanabe S et al. (2021) Development of an Australian bread wheat nested association mapping population a new genetic diversity resource for breeding under dry and hot climates. Int J Mol Sci. 9:4348. https://doi.org/10.3390/ijms22094348
10. Endresen DTF, Street K, Mackay M, Bari A, Amri A et al. (2012) Sources of resistance to stem rust (Ug99) in bread wheat and durum wheat identified using focused identification of germplasm strategy. Crop Sci. 2:764-773. https://doi.org/10.2135/cropsci2011.08.0427
11. Fragoso CA, Moreno M, Wang ZH, Heffelfinger C, Arbelaez LJ et al. (2017) Genetic architecture of a rice nested association mapping population. G3-Genes Genomes Genet. 7:1913-1926. https://doi.org/10.1534/g3.117.041608
12. Ghosh S, Watson A, Gonzalez-Navarro OE, Ramirez-Gonzalez RH, Yanes L et al. (2018) Speed breeding in growth chambers and glasshouses for crop breeding and model plant research. Nat Protoc. 12:2944-2963. https://doi.org/10.1038/s41596-018-0072-z
13. Glasser R, Johnstone C, Kapetas A (2022) The geopolitics of climate and security in the Indo-Pacific. Australian Strategic Policy Inst. https://ad-aspi.s3.ap-southeast-2.amazonaws.com/2022-02/Climate%20and%20security%20in%20the%20IndoPacific.pdf?VersionId=1TtZ51hxj7WoOCMxGqRCk8GvpviUdq4L. Accessed 19 06 2022.
14. Guo ZG, Tucker DM, Wang DL, Basten CJ, Ersoz E et al. (2013) Accuracy of across-environment genome-wide prediction in maize nested association mapping populations. G3-Genes Genomes Genet. 3:263-272. https://doi.org/10.1534/g3.112.005066

15. Gupta P, Kabbaj H, El Hassouni K, Maccaferri M, Sanchez-Garcia M et al. (2020) Genomic regions associated with the control of flowering time in durum wheat. Plants. 12:1628. https://doi.org/10.3390/plants9121628
16. Hayes BJ, Bowman PJ, Chamberlain A, Goddard ME (2009) Invited review: Genomic selection in dairy cattle: Progress and challenges. Int J Dairy Sci. 92:433-443. https://doi.org/10.3168/jds.2008-1646
17. Hochman Z, Gobbett DL, Horan H (2017) Climate trends account for stalled wheat yields in Australia since 1990. Glob. Change Biol. 23:2071–2081. https://doi.org/10.1111/gcb.13604
18. Hu J, Guo C, Wang B, Ye J, Liu M et al. (2018) Genetic properties of a nested association mapping population constructed with semi-winter and spring oilseed rapes. Front Plant Sci. 9:1740. https://doi.org/10.3389/fpls.2018.01740
19. Kang Y, Haeften SV, Bustos-Korts D, Vukasovic S, Khan SU et al. (2021) Physiological and genetic drivers underpinning canopy development are associated with durum wheat yield in rainfed environments. bioRxiv. https://doi.org/10.1101/2021.08.21.457180 (Preprint posted August 20, 2021)
20. Kidane YG, Gesesse CA, Hailemariam BN, Desta EA, Mengistu DK et al. (2019) A large nested association mapping population for breeding and quantitative trait locus mapping in ethiopian durum wheat. Plant Biotechnol J. 17:1380-1393. https://doi.org/10.1111/pbi.13062
21. Kitony JK, Sunohara H, Tasaki M, Mori JI, Shimazu A et al. (2021) Development of an aus-derived nested association mapping (Aus-NAM) population in rice. Plants. 6:1255. https://doi.org/10.3390/plants10061255
22. Ladejobi O, Elderfield J, Gardner KA, Gaynor RC, Hickey LT et al. (2016) Maximizing the potential of multi-parental crop populations. Appl Transl Genom. 11:9-17. https://doi.org/10.1016/j.atg.2016.10.002
23. Liu X, Huang M, Fan B, Buckler ES, Zhang Z (2016) Iterative usage of fxed and random efect models for powerful and efficient genomewide association studies. PLoS Genet. 12:e1005767. https://doi.org/10.1371/journal.pgen.1005767
24. Lev-Yadun S, Gopher A, Abbo S (2000) The cradle of agriculture. Science. 288:1602. https://doi.org/10.1126/science.288.5471.1602
25. Maccaferri M, Harris NS, Twardziok SO, Pasam RK, Gundlach H et al. (2019) Durum wheat genome highlights past domestication signatures and future improvement targets. Nat Genet. 51:885-895. https://doi.org/10.1038/s41588-019-0381-3
26. Mace E, Hunt C, Jordan D (2013) Supermodels: Sorghum and maize provide mutual insight into the genetics of flowering time. Theor Appl Genet. 126:1377-1395. https://doi.org/10.1007/s00122-013-2059-z
27. Mackay TF, Stone EA, Ayroles JF (2009) The genetics of quantitative traits: Challenges and prospects. Nat Rev Genet. 10:565-577. https://doi.org/10.1038/nrg2612
28. Mackay IJ, Bansept-Basler P, Barber T, Bentley AR, Cockram J et al. (2014) An eight-parent multiparent advanced generation inter-cross population for winter-sown wheat: Creation, properties, and validation. G3-Genes Genomes Genet. 4:1603-1610. https://doi.org/10.1534/g3.114.012963
29. Mangini G, Blanco A, Nigro D, Signorile MA, Simeone R (2021) Candidate genes and quantitative trait loci for grain yield and seed size in durum wheat. Plants. 10:312. https://doi.org/10.3390/plants10020312
30. Maranna S, Kumawat G, Nataraj V, Gireesh C, Gupta S et al. (2019) NAM population – a novel genetic resource for soybean improvement: Development and characterization for yield and attributing traits. Plant Genet Resour. 6:545-553. https://doi.org/10.1017/S1479262119000352
31. Mazzucotelli E, Sciara G, Mastrangelo AM, Desiderio F, Xu SS (2020) The global durum wheat panel (GDP): An international platform to identify and exchange beneficial lleles. Front. Plant Sci. 11:569905. https://doi.org/10.3389/fpls.2020.569905
32. Mengistu DK, Kidane YG, Catellani M, Frascaroli E, Fadda C et al. (2016) High-density molecular characterization and association mapping in ethiopian durum wheat landraces reveals high diversity and potential for wheat breeding. Plant Biotechnol J. 14:1800-1812. https://doi.org/10.1111/pbi.12538
33. Milner SG, Maccaferri M, Huang BE, Mantovani P, Massi A et al. (2016) A multiparental cross population for mapping qtl for agronomic traits in durum wheat (Triticum turgidum ssp. Durum). Plant Biotechnol J. 1:735-748. https://doi.org/10.1111/pbi.12424

34. Perumal R, Tesso TT, Morris GP, Jagadish SK, Little CR et al. (2021) Registration of the sorghum nested association mapping (NAM) population in RTx430 background. J Plant Regist 2:395-402, https://doi.org/10.1002/plr2.20110
35. Richard CA, Hickey LT, Fletcher S, Jennings R, Chenu K et al. (2015) High-throughput phenotyping of seminal root traits in wheat. Plant Methods. 1:1-11. https://doi.org/ 10.1186/s13007-015-0055-9
36. Robinson H, Hickey LT, Richard C, Mace E, Kelly A et al. (2016) Genomic regions influencing seminal root traits in barley. Plant Genome. 9:1. https://doi.org/10.3835/plantgenome2015.03.0012
37. Roncallo PF, Larsen AO, Achilli AL, Saint Pierre C, Gallo CA et al. (2021) Linkage disequilibrium patterns, population structure and diversity analysis in a worldwide durum wheat collection including Argentinian genotypes. BMC Genomics. 1:1-17. https://doi.org/10.1186/s12864-021-07519-z
38. Ruan Y, Zhang W, Knox RE, Berraies S, Campbell HL et al. (2020) Characterization of the genetic architecture for fusarium head blight resistance in durum wheat: The complex association of resistance, flowering time, and height genes. Front Plant Sci. 11:592064. https://doi.org/10.3389/fpls.2020.592064
39. Semagn K, Iqbal M, Chen H, Perez-Lara E, Bemister DH et al. (2021) Physical mapping of QTL in four spring wheat populations under conventional and organic management systems. I. earliness. Plants. 5:853. https://doi.org/10.3390/plants10050853
40. Shi C, Zhao L, Zhang X, Lv G, Pan Y et al. (2019) Gene regulatory network and abundant genetic variation play critical roles in heading stage of polyploidy wheat. BMC Plant Biol. 19:6. https://doi.org/10.1186/s12870-018-1591-z
41. Tester M, Langridge P (2010) Breeding technologies to increase crop production in a changing world. Sci. 327:818-822, https://doi.org/10.1126/science.1183700
42. Voss-Fels KP, Frisch M, Qian L, Kontowski S, Friedt W et al. (2015) Subgenomic diversity patterns caused by directional selection in bread wheat gene pools. The Plant Genome. 8:2. https://doi.org/10.3835/plantgenome 2015.03.0013
43. Wang Z, Hu H, Jiang X, Tao Y, Lin Y (2020) Identification and validation of a novel major quantitative trait locus for plant height in common wheat (Triticum aestivum L.). Front. Genet. 11:602495. https://doi.org/10.3389/fgene.2020.602495
44. Wang J, Zhang Z (2021) GAPIT version 3: boosting power and accuracy for genomic association and prediction. Genomics, proteomics & bioinformatics. 4:629-640. https://doi.org/10.1016/j.gpb.2021.08.005
45. Watson A, Ghosh S, Williams M, Cuddy WS, Simmonds J et al. (2018) Speed breeding is a powerful tool to accelerate crop research and breeding. Nat Plants. 4:23-29. https://doi.org/10.1038/s41477-017-0083-8
46. Yu J, Holland JB, Mcmullen MD, Buckler ES (2008) Genetic design and statistical power of nested association mapping in maize. Genet. 178:539-551. https://doi.org/10.1534/genetics.107.074245
47. Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M et al. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet. 38:203-208. https://doi.org/10.1038/ng1702
48. Zhang J, She M, Yang R, Jiang Y, Qin Y (2021) Yield-related QTL clusters and the potential candidate genes in two wheat DH populations. Int J Mol Sci. 21:11934. https://doi.org/10.3390/ijms222111934
49. Ziems LA, Hickey LT, Platz GJ, Franckowiak JD, Dracatos PM et al. (2017) Characterization of Rph24: A gene conferring adult plant resistance to puccinia hordei in barley. Phytopathology. 107:834-841. https://doi.org/10.1094/phyto-08-16-0295-r

Figures

Figure 1

Development of the durum wheat Nested Association Mapping (dNAM) population. (a) Visual representation of the nested structure of the 10 NAM bi-parental families created by crossing eight elite ICARDA donor lines with two Australian varieties (Jandaroi and DBA Aurora). The 10 families are as follows: Family 1 = DBA Aurora × Fastoz7; Family 2 = DBA Aurora × Outrob4; Family 3 = DBA Aurora × Fastoz8; Family 4 = DBA Aurora × Fadda98; Family 5 = DBA Aurora × Fastoz3; Family 6 = Jandaroi × Fastoz8; Family 7 = Jandaroi × Fastoz10; Family 8 = Jandaroi × Fastoz6; Family 9 = Jandaroi × Fastoz2; and Family 10 =

Jandaroi × Outrob4. (b) Crossing and most single seed descent (SSD) generations were produced under speed breeding conditions (brown), and one generation was produced in the field (green).

Figure 2

Phenotypic variation expressed by the 10 dNAM populations (*n* = 92 individuals per population) grown under field conditions in the F_4:F_5 generation. a) Seed increase at Gatton research station in a 2-m-row design during the grain filling stage. b) Boxplots showing plant height in the 10 dNAM population. Each boxplot represents one family coloured according to the key at the bottom right. c) Boxplots showing variation in the dNAM population for days to flowering. Each boxplot represents one family (coloured using the same key).

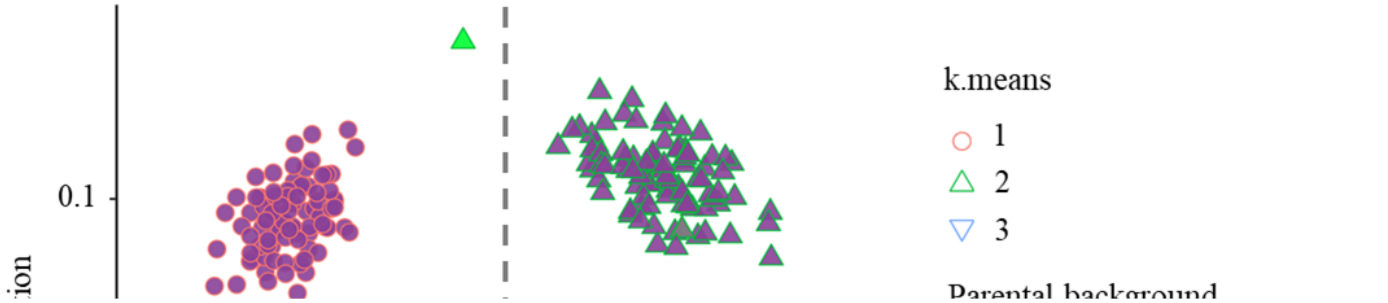


Figure 3

Population structure of the dNAM population. Principal component analysis based on pairwise Roger's distances calculated from 6,785 high-quality polymorphic SNP markers in 926 dNAM lines representing the 10 families. Genetic distance was calculated, and the variance explained by principal components was PC1 = 31.8% and PC2 = 9.6%. The vertical dashed line separates the five families with the DBA Aurora background from families with the Jandaroi background. The optimum number of clusters (3) is displayed as shapes, and the entire dNAM population is clustered by family and coloured according to the donor parents.

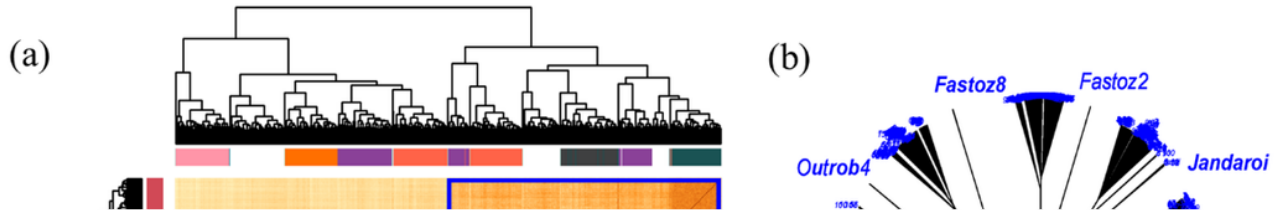


Figure 4

Calculation of the genetic relatedness of 926 F_6 dNAM lines using 6,785 DArTseq high-quality markers. a) Heatmap displaying two major clusters representing the Australian reference parents Jandaroi (blue box) and DBA Aurora (red box). Lines with high levels of genetic similarity clustered together within each reference parent, representing 10 sub-dNAM populations (dark brown). The dendrogram at the top is color-coded according to the dNAM families, and the dendrogram on the left corresponds to the three distinct clusters displayed based on the calculated k-means values. b) Neighbour-Joining unrooted dendrogram analysis discriminated individuals across and within genetic backgrounds. Individuals sharing Jandaroi as a reference clustered into five families (in blue), as did individuals sharing DBA Aurora as a reference (in red).

Figure 5

Comparison of linkage disequilibrium (LD) decay of the entire dNAM population and the 10 families. LD decay is shown as the intersection of the LOESS curves at a threshold of r^2 = 0.1, 0.2. Each LOESS curve represents one family (coloured using the same key). The LD decay of the entire dNAM is referred to as ALL.



Figure 6

Genome-wide association mapping for plant height and flowering time for 926 dNAM lines and 10 parental lines using 6,785 high-quality DArTseq markers. a) The Manhattan plots highlight marker associations for days to flowering in blue and plant height in orange. The blue dotted horizontal line represents the significance level of the marker-trait association at the family-wise *p*-value (threshold ≥ 3) and the red line shows the significance level using Bonferroni threshold of *p*-value (threshold ≥ 5.14). b) QQ plots for days to flowering in blue and plant height in orange are displayed.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SamirAlahmadSuptables.docx