

Aberystwyth University

Population genomics of Mediterranean oat (A. sativa) reveals high genetic diversity and three loci for heading date

Canales, Francisco J.; Montilla-Bascón, Gracia; Bekele, Wubishet Abebe; Howarth, Catherine; Langdon, Tim; Risipail, Nicolas; Tinker, Nicholas A.; Prats, Elena

Published in:

Theoretical and Applied Genetics

DOI:

[10.1007/s00122-021-03805-2](https://doi.org/10.1007/s00122-021-03805-2)

Publication date:

2021

Citation for published version (APA):

Canales, F. J., Montilla-Bascón, G., Bekele, W. A., Howarth, C., Langdon, T., Risipail, N., Tinker, N. A., & Prats, E. (2021). Population genomics of Mediterranean oat (*A. sativa*) reveals high genetic diversity and three loci for heading date. *Theoretical and Applied Genetics*. <https://doi.org/10.1007/s00122-021-03805-2>

Document License

CC BY

General rights

Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400

email: is@aber.ac.uk



Population genomics of Mediterranean oat (*A. sativa*) reveals high genetic diversity and three loci for heading date

F. J. Canales¹ · G. Montilla-Bascón¹ · W. A. Bekele² · C. J. Howarth³ · T. Langdon³ · N. Rispaill¹ · N. A. Tinker² · E. Prats¹

Received: 19 November 2020 / Accepted: 24 February 2021
© The Author(s) 2021

Key message Genomic analysis of Mediterranean oats reveals high genetic diversity and three loci for adaptation to this environment. This information together with phenotyping and passport data, gathered in an interactive map, will be a vital resource for oat genetic improvement.

Abstract During the twentieth century, oat landraces have increasingly been replaced by modern cultivars, resulting in loss of genetic diversity. However, landraces have considerable potential to improve disease and abiotic stress tolerance and may outperform cultivars under low input systems. In this work, we assembled a panel of 669 oat landraces from Mediterranean rim and 40 cultivated oat varieties and performed the first large-scale population genetic analysis of both red and white oat types of Mediterranean origin. We created a public database associated with an interactive map to visualize information for each accession. The oat collection was genotyped with 17,288 single-nucleotide polymorphism (SNP) loci to evaluate population structure and linkage disequilibrium (LD); to perform a genome-wide association study (GWAs) for heading date, a key character closely correlated with performance in this drought-prone area. Population genetic analysis using both structure and PCA distinguished two main groups composed of the red and white oats, respectively. The white oat group was further divided into two subgroups. LD decay was slower within white lines in linkage groups Mrg01, 02, 04, 12, 13, 15, 23, 33, whereas it was slower within red lines in Mrg03, 05, 06, 11, 21, 24, and 28. Association analysis showed several significant markers associated with heading date on linkage group Mrg13 in white oats and on Mrg01 and Mrg08 in red oats.

Introduction

Oat is an important crop of Mediterranean origin that is used worldwide as food, feed grain, green fodder and as a winter cover crop in no-till rotations (Stevens et al. 2004). *Avena sativa* L. subsp. *sativa* and subsp. *byzantina* sometimes known as the white and red oat, respectively, are the main cultivated oats. They are grown most widely in temperate areas, with an increasing interest to expand the crop to subtropical areas and Mediterranean countries (Stevens et al.

2004). This is mainly due to its good adaptation to a wide range of soil types and because on marginal soils oats can perform better than other small-grain cereals (Stevens et al. 2004). Currently, the cultivated area of oats in the Mediterranean rim is equal to that of the northern European countries (Rispaill et al. 2018). However, in Mediterranean areas, oat is sensitive to hot and dry weather due, in part, to a high transpiration rate (Ehlers 1989). Indeed, drought is the main limiting factor for oat yield, exceeding losses from all other causes under some conditions (Stevens et al. 2004). In addition, oat fungal diseases such as powdery mildew, caused by *Blumeria graminis* f.sp. *avenae*, and crown rust, caused by *Puccinia coronata*, are major constraints for this crop. The use of resistant varieties is one of the best control alternatives (Stevens et al. 2004). However, most of the current oat cultivars used for winter cropping in the Mediterranean rim are spring cultivars bred in northern countries that are not well-adapted to Mediterranean conditions (Prats et al. 2014). This might explain the poor oat adaptation observed in this area.

Communicated by Andreas Graner.

✉ E. Prats
elena.prats@ias.csic.es

¹ Institute for Sustainable Agriculture, CSIC, Avenida Menéndez Pidal, 14004 Córdoba, Spain

² Ottawa Research and Development Centre, Agriculture and Agri-Food Canada, Ottawa, ON, Canada

³ Institute of Biological, Environmental and Rural Sciences, Aberystwyth Univ, Aberystwyth, UK

In contrast to the poor adaptation observed in many oat cultivars growing under Mediterranean conditions, landraces (for which this area is particularly rich) are regionally adapted to this environment. They provide a tangible gene pool for future genetic breeding programs and food security (Chalak 2015). Transfer of beneficial traits from landraces is relatively straightforward as there is no barrier for crossing. However, during the last century, the replacement of landraces by modern cultivars has led to a significant genetic erosion in many cereal crops (Newton et al. 2010). Studies on maize (Warburton et al. 2008) and wheat (Reif et al. 2005; Roussel et al. 2004) have identified a significant reduction in diversity. These studies also suggest that landraces may be a good source of new allelic diversity for breeding programs. Germplasm banks are facilities designed to provide long-term conservation of genetic resources including landraces. However, more exhaustive characterization of these collections, ranging from morphological or agronomical features to genome information, is necessary to assure maximum utilization of this germplasm. Moreover, association of genetic markers with regions of the genome controlling different traits would enable the development of marker-assisted selection for efficient and precise transfer of useful alleles from landraces to modern cultivars.

Genetic diversity, assessed by various tools including DNA markers, is important information both for genetic conservation and for the efficient breeding of new commercial varieties. DNA markers such as amplified fragment length polymorphisms (AFLPs), random amplified polymorphic DNA (RAPDs) or simple sequence repeats (SSRs or microsatellites) have been used to assess oat genetic diversity and to examine allelic diversity changes over 100 years of oat breeding in both Nordic countries and Canada (e.g., Fu et al. 2007). Recently however, genotyping-by-sequencing (GBS) has emerged as a more robust genomic approach to explore plant genetic diversity of species with complex genomes that lack extensive genomic resources such as oat (Elshire et al. 2011). GBS has been used to identify genomic regions associated with particular traits by bi-parental mapping (e.g., Kebede et al. 2020) or genome-wide association analysis (GWAS, e.g., Esvelt Klos et al. 2017).

In this work, we characterized 709 oat accessions, mainly consisting of Mediterranean red and white oat landraces, from 24 different countries. We created an interactive map to easily access their main morphological and passport information and we dissected their genetic relationships using a detailed genetic diversity study through GBS. Furthermore, we carried out a GWAS of heading date, a key character determining performance in this drought-prone area, that validates the strength of this collection and the genetic information developed.

Materials and methods

Plant material

A collection of 709 accessions of white and red *Avena sativa* L. was assembled, containing a majority of landraces from the Mediterranean area. Seeds were obtained from “Centro de Recursos Fitogenéticos” (INIA, Madrid, Spain) and United States Department of Agriculture (USDA, Washington, USA). Cultivars included within the collection were provided by various institutions as reported by Sánchez-Martín et al. (2014). They were included for comparison since they are widely cultivated in the Mediterranean area. All accessions were purified through single seed descent over three generations prior to genotypic and phenotypic evaluation.

Phenotypic data

The Mediterranean oat collection was evaluated in three different environments in Spain for heading date. Trials were performed in two contrasting locations, Santaella (Spain), with 238 m altitude, and light clay eutric gleysol during growing season 2017–2018 and Córdoba (Spain), with 90 m altitude and light clay calcic cambisol during growing seasons 2016–2017 and 2017–2018. At each location-year, an alpha lattice square design with three replicates was used and the cultivar ‘Patones’ was used as check. The randomization of genotypes was done with SPSS v. 25 software (IBM Corp. Armonk, NY, USA). Each replicate consisted of 27 × 27 1-m-long rows containing the 709 genotypes plus additional checks included in each row and column until completing the lattice square. Rows were separated from each other by 30 cm, at a sowing density of 200 seeds m⁻². Replications were bordered by cv ‘Flegla’. Sowings took place in December, according to local practices. No irrigation was performed in the trials. Hand weeding was carried out when required, and no herbicides or fertilizers were applied. Phenotypic data are available in Online Resource 1.

Interactive map

The interactive map with the 709 oat accessions was created with Google Maps (Google-LTD 2021). Each accession was located according to its site of origin taken from the passport data provided by the germplasm bank database. For each located accession, additional data were stored including (1) general information from its passport data, (2) agronomic information collected from field experiments joined by pictures of spikelets and seeds. In addition, different map layers were added where each variety can be visualized by spot

color depending on subspecies (*byzantina*, *sativa* or admixture), heading date (early, mid or late heading), improvement status (cultivar, landrace or breeding material) and geographic elevation (low or high altitude).

DNA extraction, library construction and sequencing

Leaf tissue was harvested from single plants growing in the field and lyophilized before DNA extraction. DNA was extracted according to the CTAB protocol (Murray and Thompson 1980) and the complexity reduced, multiplexed GBS libraries were constructed by Plateforme d'Analyses Génomiques of the Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval (Quebec City, Canada), based on the *PstI*–*MspI* method described by Huang et al. (2014). Complexity reductions were multiplexed using barcode adapters, with 96 samples per pooled library. Sequencing of each pooled library was performed on a single lane of a HiSeq 2500 platform (Illumina, San Diego, CA, USA) using standard Illumina protocols and kits, producing high-output paired-end 150 bp reads at the Genome Quebec sequencing center (Montreal, Canada).

Tag-level haplotype and SNP analysis

Raw sequence files in FASTQ format were processed using initial steps of the UNEAK-GBS pipeline within TASSEL 3.0 bioinformatics analysis package (Lu et al. 2013) to trim the reads, de-convolute the barcodes and produce a single tag count file for each sample. These files were then used by the Haplotag production pipeline (Tinker et al. 2016) to call marker genotypes (including both tag-level haplotypes and GBS SNPs) based on pre-determined genetic loci identified previously in a set of 4,657 *A. sativa* accessions reported by Bekele et al. (2018). The resulting data matrix was filtered using in-house software, to remove all markers with lower than 80% completeness, less than 5% minor allele frequency, or greater than 10% heterozygosity from the data matrix. Since the production mode was developed in a population that was dominated by North American germplasm, which might introduce ascertainment bias, a second dataset was created by running the Haplotag pipeline in discovery mode together with a representative set of North American landraces from the population reported by Esvelt Klos et al. (2016). These data were filtered using the same parameters as the genotype calls from the production mode. The genetic positions of the markers on the consensus map (Bekele et al. 2018) were identified. Structure and PCA analyses showed no differences in the population structure and clustering of the Mediterranean accessions between the production and discovery dataset. Therefore, all subsequent analyses were performed with

the production dataset to take advantage of the marker map position except for the PCA analysis.

Population structure and kinship

Population structure of the Mediterranean oat GBS dataset was inferred in the production mode matrix with 17,288 polymorphic GBS-SNP markers by the software STRUCTURE 2.3.4 (Pritchard et al. 2000) using the admixture model and the option of correlated allele frequencies between populations, as this configuration is considered best by Falush et al. (2003). Similarly, we let alpha (the degree of admixture) be inferred from the data. Each simulation included 10,000 burn-in iterations, 20,000 iterations and ten independent simulations per k value. The number of subpopulations and the percentages of admixture of each accession (Q matrix) were given by STRUCTURE HARVESTER (Earl and vonHoldt 2012), an online software to visualize STRUCTURE output implementing the Δk method (Evanno et al. 2005). For subsequent analyses, accessions were assigned to a subpopulation when they showed more than 60% membership in a subpopulation. Visualization of STRUCTURE data was done with the online software STRUCTUREPLOT (Ramasamy et al. 2014).

Principal component analysis (PCA) was also performed as an alternative method to infer the structure of the collection with the software package TASSEL 5 (Bradbury et al. 2007). In addition, two UPGMA dendrograms of white and red oat accessions were computed using TASSEL 5 with default settings (Bradbury et al. 2007). The estimated trees were then drawn with the ggplot2 package in R (R-Team, 2017). Accessions were labeled with different colors according to their geographic origin (north Mediterranean, west Mediterranean, south Mediterranean and east Mediterranean).

Linkage disequilibrium

Pairwise measures of LD based on r^2 were computed using TASSEL 5 for the whole collection of Mediterranean oat as well as for each oat subspecies (*sativa* and *byzantina*). LD test was calculated for each marker pair (full matrix) and per chromosome to assess the overall linkage disequilibrium of the collection and the distance of linkage decay per linkage group. The LD decay plots were generated by using locally estimated smoothed lines (LOESS) (Jacoby 2000) in R (R-Team, 2017).

Genome-wide association mapping for heading date

Marker-phenotype associations were estimated with the software package TASSEL 5, according to Rispail et al. (2018).

Associations were tested using a general linear model with PCA covariates (GLM-PCA) and with a mixed linear model (MLM) using the PCA coordinates for Q and the kinship distances for K , as proposed by Yu et al. (2006). Models were run using mean heading dates across the three environments as phenotypes. Mean heading date values for each genotype within an environment were normalized according to their respective checks. Combined quantile–quantile plots drawn to compare the observed distribution of $-\log_{10}(P)$ values for each marker-trait association using the cumulative distribution to assess the correct fitting of the data and the control for type I error. Selection of the optimal model was performed by estimating the genetic inflation factor λ_{median} , as follows: $\lambda_{\text{median}} = \text{median}(Sc + 1, Sc + 2, \dots, Sn)/0.456$. A genetic inflation factor higher than 1.25 ($\lambda > 1.25$) indicates data stratification even after structure corrections. Therefore, p values for all models with $\lambda > 1.25$ were adjusted with their respective λ to avoid false positives (Reed et al. 2015). The critical adjusted p values for assessing the significance of associations were corrected for multiple testing based on the false discovery rate (FDR) criteria (Benjamini and Hochberg 2000). The matrix of p values was used to estimate the corresponding q values with the Q value package (Storey 2002) in R (R-Core-Team, 2017). Manhattan plots of GWAS results were drawn with the qqman package in R (R-Team, 2017).

Sequence homology

We used GBS markers with known genetic position on the oat consensus map (Bekele et al. 2018) to find nearby annotated genes in two sequenced diploid oat genomes (*Avena atlantica* (AA) and *Avena eriantha* (CC) described by Maughan et al. (2019) and in a recently released pseudomolecule assembly of hexaploid oat from the Canadian breeding line ‘OT3098’ (available July 2020 at <https://wheat.pw.usda.gov/GG3/node/922>). GBS marker positions in the diploid genomes were declared based on the best BLAST match of the reference sequences of the GBS markers. In the hexaploid genome, GBS positions were declared when an exact match of a 64-base tag-level haplotype belonging to a given locus was found at one and only one genome position. At the time of submission, the finalization of hexaploid chromosome names and orientations was still before the Nomenclature Committee of the International Oat Conference, thus we refer to chromosome identities and genome positions as “Hexaploid Oat Reference OT3098 V1”. Online Resource 2 shows a working correspondence between Mrg consensus groups and chromosome identities in the above diploid and hexaploid pseudomolecules based on the best match of ten random markers for each linkage group. The function BlastX of the BLAST algorithm (Altschul et al. 1990) was used to validate protein identities in the NCBI non-redundant protein database (database released on 27 Julio 2020), as

implemented in the NCBI webserver (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

Results

Characterization of the oat collection and localization in an interactive map

Data from the accessions of the oat Mediterranean collection including the north-European cultivars are shown on the interactive map (https://www.google.com/maps/d/viewer?mid=1t-O7OUoUPJ_qY5qGQsq66O5zg7k) where the different classification criteria may be used as filters yielding several layers that show a different color code according to the selected criteria. The data underlying the map are also available at Dryad Data (Canales et al. 2021; <https://doi.org/10.5061/dryad.0gb5mkm0g>). Different layers of the map (e.g., Figure 1), showed approximately even representation across most classifications other than improvement status, where the majority of accessions (546 of 709) were landraces (Fig. 1b). The oldest landrace was accession 370, which is a landrace from Evros in Greece collected in 1904. Most landraces were collected between 1940 and 1970, albeit some recently collected landraces were also included, such as accessions 245 and 246, collected in Balearic island, Spain, in 2009. Interestingly, some of the cultivars are even older than the landraces, including cv ‘Selma’ from Germany and ‘Joanette’ from France, both white oats from 1889. A total of 251 red and 435 white oats were included in the collection. As observed in Fig. 1a, both subspecies were collected from all around Mediterranean rim. There were accessions representing low- or high-altitude sites (Fig. 1c) and covering a wide range of heading dates (Fig. 1d). In addition to the main classification criteria, other passport data as well as detailed pictures of the spike morphology and seed morphology, color and size are available by clicking on the corresponding accession (e.g., Online Resource 3, 4).

Genetic relationships among Mediterranean oats

The Haplotag production pipeline provided data for 164,652 tag-level haplotypes and 240,958 SNPs. After the filtering based on MAF, heterozygosity and missing values, a total of 17,288 polymorphic SNP markers corresponding to 12,418 tags were obtained. Of these tags, 4,829 were located on the hexaploid consensus map. Sequence read data from these accessions are available from NCBI SRA archive as BioProject PRJNA693576. The GBS-SNP marker datasets were deposited in the public T3/oat database (<https://oat.triticeatoolbox.org/breed>

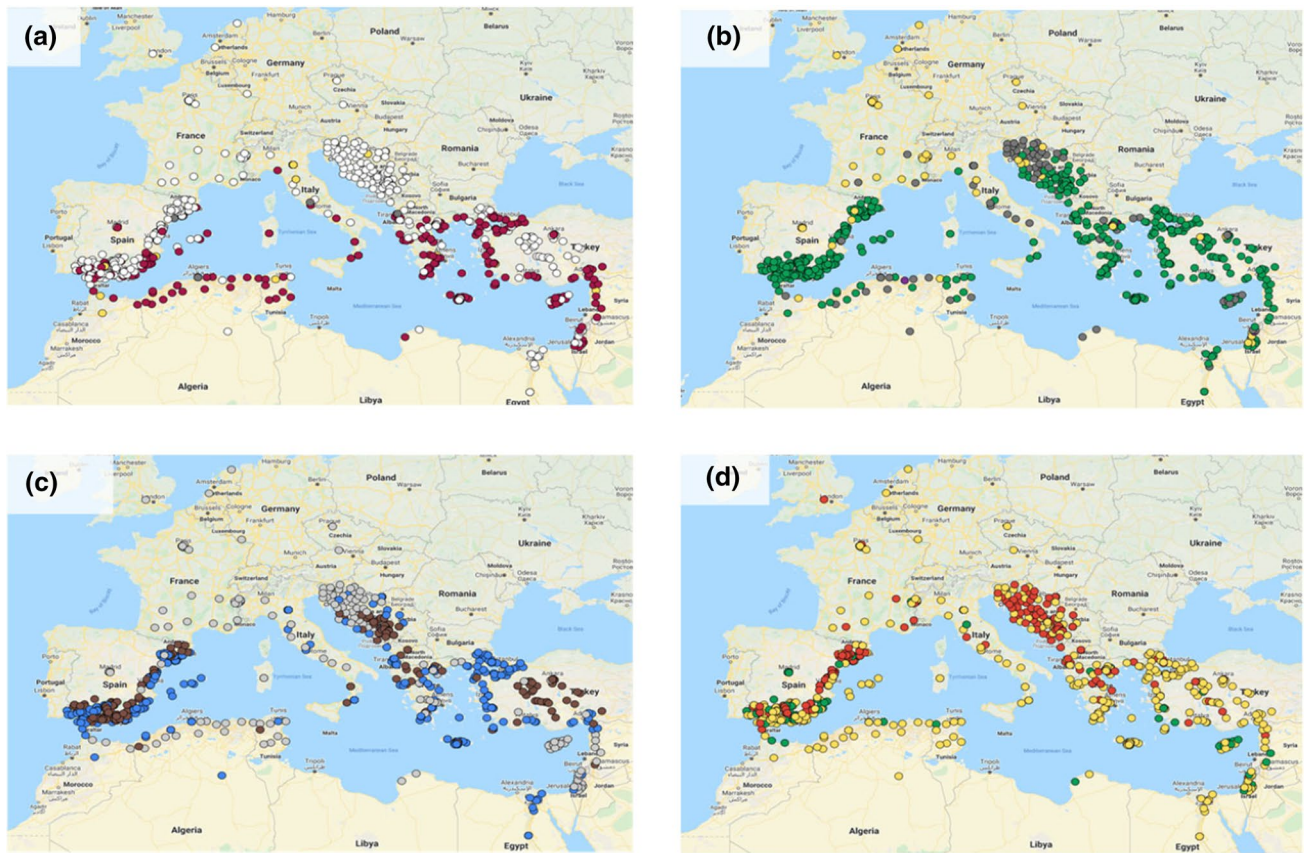


Fig. 1 Example of the interactive oat map as can be seen in google maps. In this example, the accessions were filtered according to **a** subspecies where white, red, yellow and gray dots indicated *sativa*, and *byzantina* subsp, admixture and unknown, respectively, **b** improvement status, where green, yellow, purple and gray dots indicated landraces, cultivars, breeding materials and uncertain materi-

als, respectively, gray dot indicated that information is not available, **c** altitude where brown, blue and gray dots indicated landraces collected in high (>400 m) or low (<400 m) altitude regions, cultivars and other accessions for which altitude was not available were represented by gray dots, **d** heading date, where green, yellow and red dots indicated early, mid or late flowering accessions

ers/trial/4667) and on the Dryad Data site (Canales et al. 2021, <https://doi.org/10.5061/dryad.0gb5mkm0g>).

According to the position of the breaking point in the $L(k)$ curve and the peak in the Δk distribution, accessions were assigned by STRUCTURE to two groups ($K=2$). This binary classification distinguished between the white (subsp. *sativa*) and the red (subsp. *byzantina*) oats (Fig. 2a). Out of the 709 accessions, 435 accessions were clustered in the white group, 251 accessions in the red group, 13 were classified as admixtures, and 10 showed a high level of missing data or heterozygosity and were not classified. Of the 699 classified accessions, only 28 assignments disagreed with prior passport data, and these are most likely due to previous classification errors. Based on passport data, it appears that admixtures are more common among red accessions than among white oats (Fig. 2a).

The high genetic divergence between white and red oats may mask further population structure within groups. Therefore, separate analyses were performed on each group. Both the white and red oat UPGMA tree showed additional

clustering when performed separately (Online Resource 6, 7). STRUCTURE analysis within the white group revealed two subgroups (Fig. 2b and c). Interestingly, most of the west-Mediterranean accessions, the majority of which were Spanish, were in subgroup 1, whereas most of the north-Mediterranean accessions were in subgroup 2. However, half of the east-Mediterranean accessions and a small part of north-Mediterranean accessions were in subgroup 1. Both subgroups contained accessions with a similar range of heading dates. However, further inspection of the UPGMA dendrogram showed that many of the west- and north-Mediterranean accessions tended to cluster based on heading date while those from the east- and south-Mediterranean accessions did not.

To explore the potential biological drivers of subpopulations, we performed PCA analysis (Fig. 3), highlighting different variables that could explain the splitting of these populations. In agreement with STRUCTURE analysis, PCAs separated the white and red oat groups (Fig. 3a). The first two PCs explained approximately 50% of the observed

Fig. 2 Estimated population structure of oat accessions according to STRUCTURE software. **a** Estimated population structure of the whole Mediterranean oat collection. The two colors indicate the two subpopulations in which STRUCTURE software classified the oat accessions, with the admixture barplot showing the admixture proportions of each individual. **b, c** Estimated structure of the white oat subpopulation, according to origin and flowering time

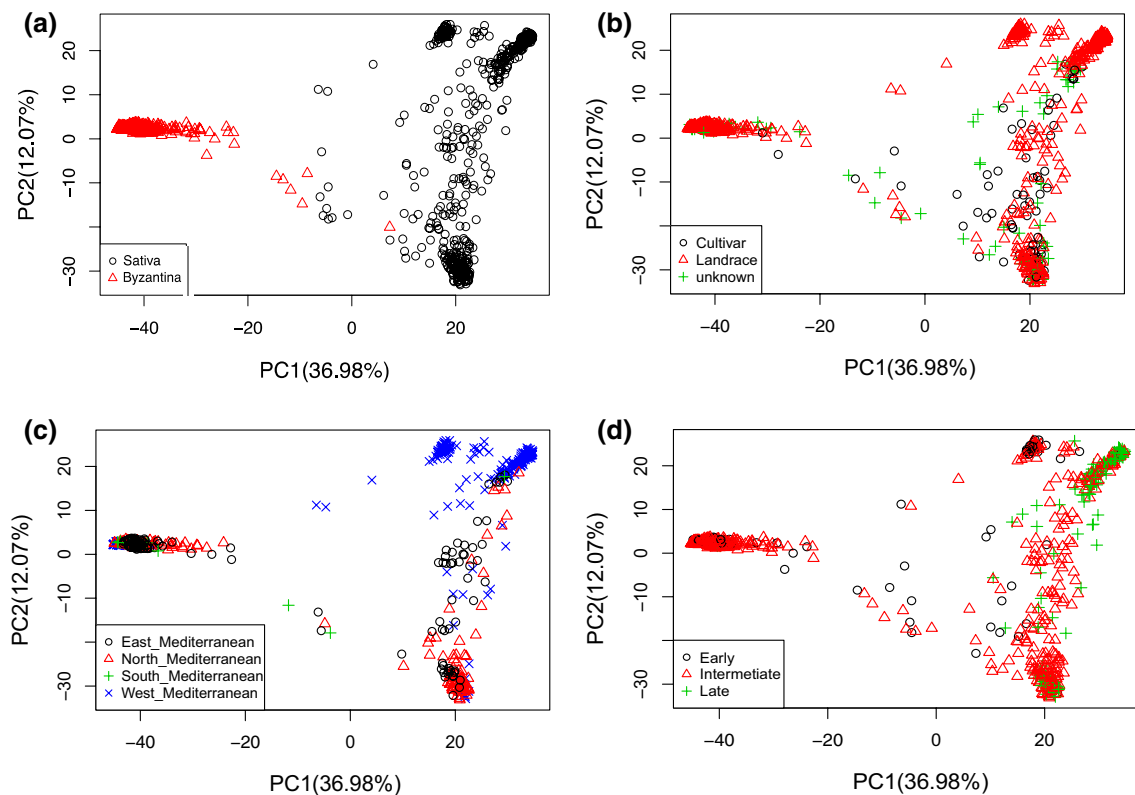
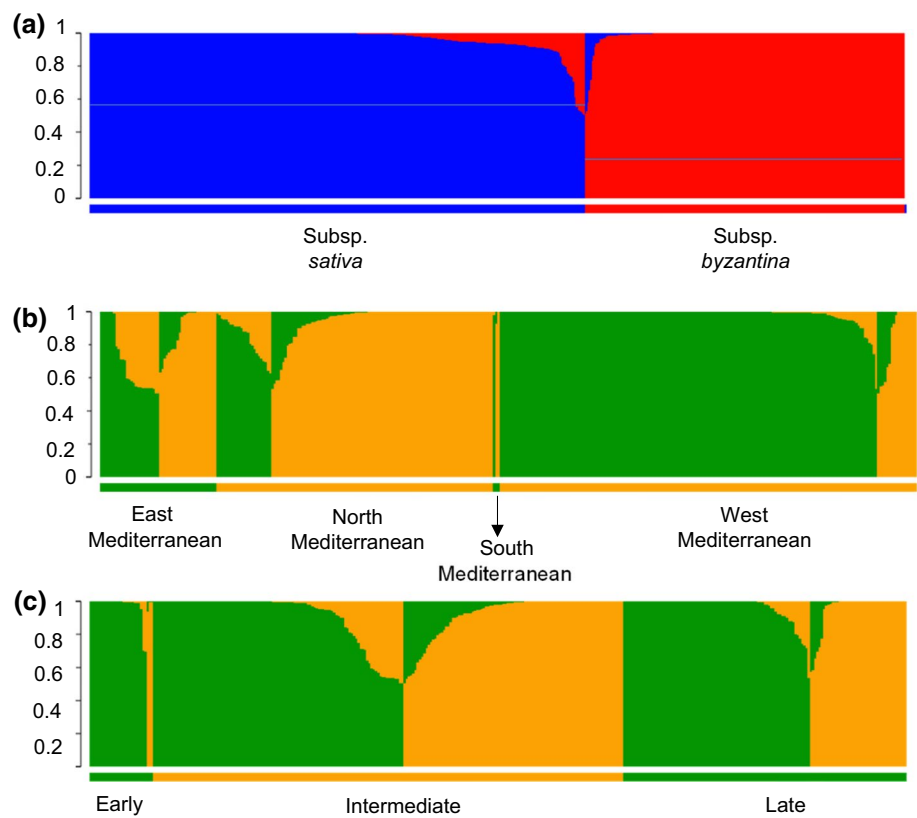


Fig. 3 Scatterplot of principal component analysis scores of components 1 and 2 based on the 17,288 polymorphic SNP markers used in this study. The different PCA graphs depict **a** subspecies, **b** type of breeding material, **c** regions of origin and **d** flowering time

genetic variance. Interestingly, cultivars were distributed among both red and white groups and also within both subpopulations of the white group (Fig. 3b). Most of the west-Mediterranean accessions clustered in the upper part of the PCA plot in three smaller subgroups, whereas most of the north-Mediterranean accessions clustered in the lower part of the figure with a small group of east-Mediterranean accessions (Fig. 3c). Figure 3d shows that the clustering of the three smaller groups of west-Mediterranean accessions could be explained based on heading date, with one group showing most of the early accessions, another with most of the late accessions and the middle group with mixed or intermediate heading dates. In addition, most of the accessions that clustered in the lower part of the plot had late or intermediate heading dates. Interestingly, overall no red accessions were late, confirming the earliness of this subpopulation.

We further compared the genetic relationships of the Mediterranean accessions with a stratified diverse subset of 69 North American landraces previously characterized by Bekele et al. (2018). For this, we used *de-novo* SNP discovery to avoid potential ascertainment bias. After filtering, 20,493 polymorphic SNP markers were available for this analysis. Clustering of Mediterranean accessions obtained with this dataset was identical to the clustering obtained with the dataset from the production pipeline. Based on PCA analysis of these data (Fig. 4), we observed that North American landraces clustered in two groups corresponding approximately with the two subgroups of Mediterranean white oat. The southern USA landraces clustered nearest to the west-Mediterranean accessions, in particular the early heading Spanish group, whereas the remaining North American accessions clustered with the north-Mediterranean accessions (Fig. 4). In an additional analysis, we performed the PCA after downsampling Mediterranean oats to equal

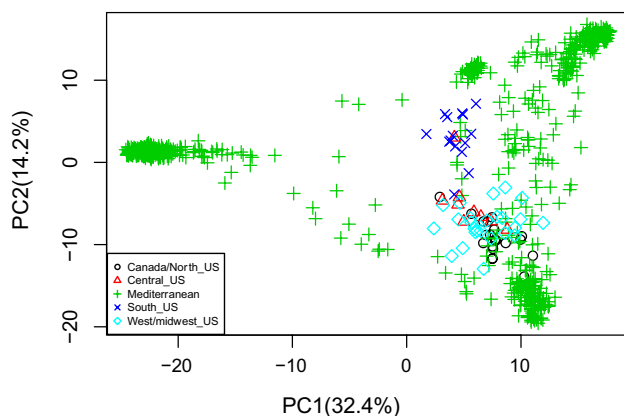


Fig. 4 Scatterplot of principal component analysis scores of components 1 and 2 of Mediterranean and 69 North American oat accessions based on 20,493 polymorphic SNP markers

the number of US oats, to rule out possible artifacts due to group size and obtained similar results (Online Resource 8). Overall, this analysis revealed a high level of diversity in the Mediterranean oat collection when compared to the North American accessions.

Genome-wide association analysis of heading date

Assessment of LD is critical for association analysis since it may determine the power and resolution of associations. The patterns of LD were assessed by pairwise comparison of markers presented as the relationship between r^2 vs. P (Online Resource 9). The rate of LD decay against genetic distance was calculated for total genome data (Fig. 5) as well as for each linkage group (Online Resource 10).

According to the full LD matrix, a total of 9,298,006 marker pairs showed a significant LD, with an average probability of $p = 0.006$. Of these, 5,773,494 marker pairs showed $r^2 < 0.1$, which is considered a nominal value for oats (Montilla-Bascón et al. 2015; Newell et al. 2011). When the two subspecies were assessed separately, 2,869,144 markers pairs showed a significant $r^2 < 0.1$ within white oats, whereas a much lower number of marker pairs (99,327) showed a significant $r^2 < 0.1$ within red oats (Online Resource 9).

Both the red and the white groups showed a similar LD decay, but the decays were faster than that across the full population (Fig. 5). Inspection of the LD by linkage groups showed differential trends. In white oat, several linkage groups showed higher LD decay including Mrg03, 05, 06, 11, 21, 24 and 28, whereas in red oat LD decayed faster on Mrg01, 02, 04, 12, 23 and 33. Strikingly, linkage group

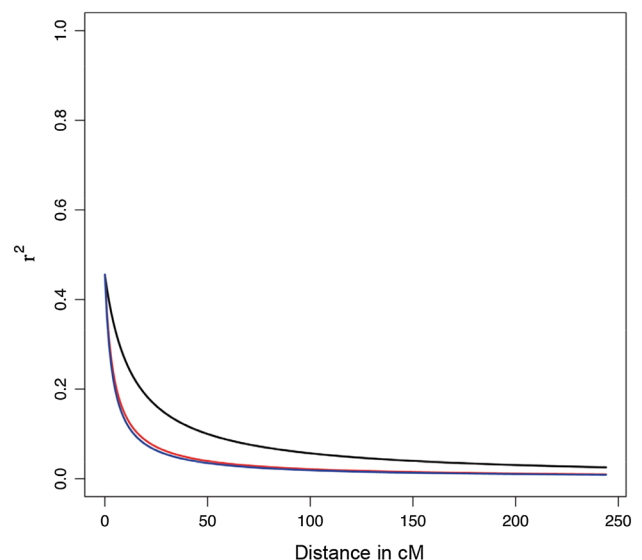


Fig. 5 Linkage disequilibrium (LD) decay by distance across the subspecies *sativa* (blue line), *byzantina* (red line) and both subspecies (black line) (color figure online)

Mrg04 showed a much slower LD decay in white oat than in red oat or in the combined population. The remaining linkage groups showed similar LD decay for white and red oats (Online Resource 10).

Due to the strong population structure, association studies were performed separately for white and red oats. Despite the fact that both models were corrected for population structure, data for white oat did not fit either the GLM-PCA or the MLM models. Increasing the numbers of PCs up to 20 in the GLM-PCA and 10 for MLM data improved the fit of these models, but the genetic inflation factor was still far from the proposed $\lambda = 1.2$ threshold (Fig. 6a, Risipail et al. 2018). To avoid false positives, we adjusted the models based on the genetic inflation factor, which resulted in a much better fit (Fig. 6b). Following this, one significant marker associated with heading date was found within white oat. This marker, avgbs_80864.1, was located in Mrg13 at cM position 60. It was significant in both the GLM-PCA and MLM models

with an adjusted of $P = 7.3e^{-5}$ and an $r^2 = 0.103$. Both GLM-PCA and MLM models revealed two significant associations with marker avgbs_cluster4923.1 located on Mrg01 at cM position 52 (Fig. 7) with an adjusted of $P = 3.18e^{-4}$ and $r^2 = 0.044$ and marker avgbs_cluster_1918.1 located on Mrg08 at cM position 147 (Fig. 7) with an adjusted of $P = 2.56e^{-5}$ and $r^2 = 0.069$.

Identification of candidate genes

To identify potential candidate genes involved in heading date, we located the markers onto the diploid (Maughan et al. 2019) and hexaploid (<https://wheat.pw.usda.gov/GG3/node/922>) oat reference genomes and examined the genomic space surrounding each associated marker using the Gbrowse tool implemented in the EPIC-CoGe Avena genome repository (Table 1, Online Resource 11). The marker avgbs_80864.1 (on Mrg13) was localized to the C genome chromosome AE05 in *A. eriantha* at position 463,706,207 bp and to position 472,093,139 bp on chromosome 2C in the hexaploid. The closest gene in the diploid, named AE025378-RA, was detected at 10 kb downstream of the associated marker. This gene showed similarity with the NAC domain-containing protein 100 (NAC100) of *Arabidopsis thaliana* which was validated by BLASTX (Table 1, Online Resource 11). Additional genes in the vicinity of avgbs_80864.1 marker were located more than 100 kb away. Annotated genes in the hexaploid genome were not located in this region, and although BLAST matches can be located, many of these appear to be repetitive domains, and further gene discovery in the hexaploid will require further annotation.

The avgbs_cluster_4923.1 marker (Mrg01) was localized to chromosome AA2 of *A. atlantica* at position 364,214,519 bp and to chromosome 1D at 249,647,335 bp in the hexaploid. Two potential candidate genes were detected downstream from this marker in the diploid. One, AA010050-RA, located 27 kb upstream and the other, AA010051-RA, at 33 kb downstream. According to BLASTX comparison, AA010050-RA shared more than 70% similarity with a Phosphatidylinositol N-acetylglucosaminyltransferase subunit P-like protein of *Zea mays*. In turn, AA010051-RA encodes a Rop guanine nucleotide exchange factor 3 (ROPGEF3), as confirmed by the BLASTX showing a similarity of 90.13% and 55.41%, respectively, with ROPGEF3 homologues in *Hordeum vulgare* and *Arabidopsis thaliana* (Table 1, Online Resource 11). Three additional genes were also identified between 40 and 50 Kb downstream of the associated marker, one with homology to a pectin acetyltransferase and two unknown proteins. The closest upstream gene, had unknown function and was located more than 140 kb away.

The associated marker avgbs_cluster_1918.1 on Mrg08 was localized on *A. eriantha* genome to chromosome AE06

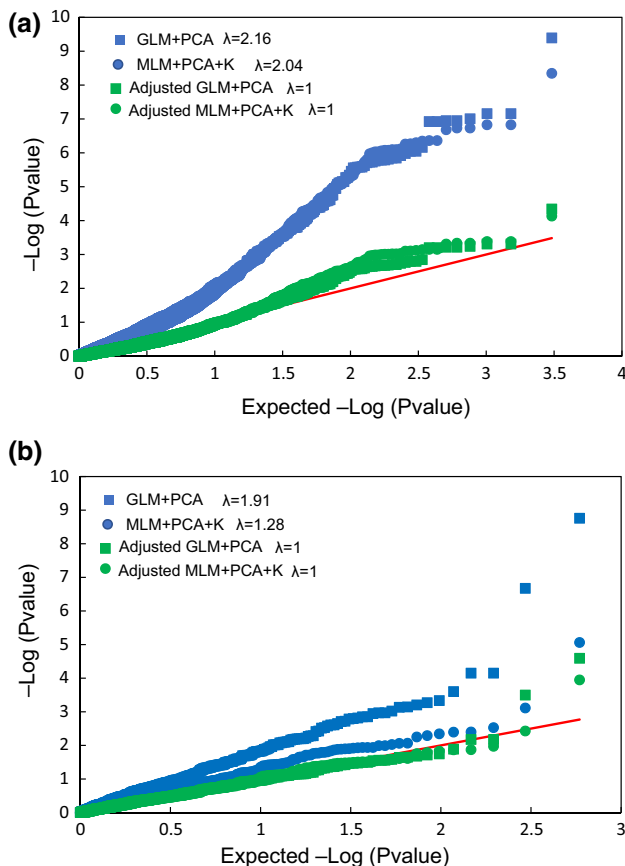
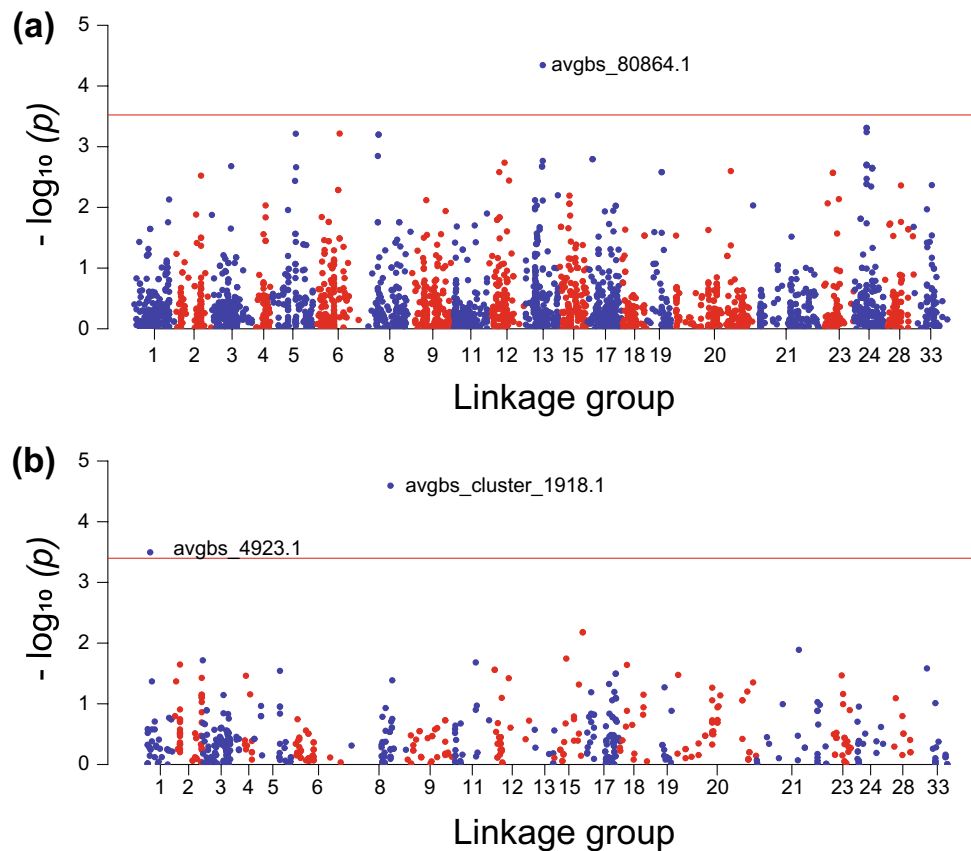


Fig. 6 Distribution of p values for the different models used in this study for the association between markers and heading date in the **a** white oat accessions and **b** red oat accessions. Axes represent the expected p values versus the observed p values in the negative \log_{10} scale, where the solid red line represents the null expectation (absence of type I error). The genetic inflation factor (λ) was used for correction of the models (color figure online)

Fig. 7 Distribution of p values by marker position on chromosomes for tests of association between markers and heading date using the adjusted GLM model in the **a** white oat accessions and **b** red oat accessions. Red and blue colors show divisions between linkage groups. The solid red line represents the genome-wide significance threshold calculated according FDR (color figure online)



at position 326,942,176 bp and to chromosome 1D at 249,647,335 bp in the hexaploid. One potential candidate gene was identified only 0.75 kb downstream of this marker in the diploid. This gene, AE037406-RA, is likely to encode another NAC-related gene since it shares 84.32 and 50.22% similarity with *Hordeum vulgare* and *Arabidopsis thaliana* NAC8, respectively (Table 1).

Discussion

To our knowledge, this is the first large-scale genomic analysis of red and white oat landraces originating from the Mediterranean region. This region is the center of origin for many crop species and their wild relatives, including cultivated oat (Loskutov 2008), and oat has been cultivated in this region for centuries. Hence, this work will provide a valuable resource for future research and crop improvement worldwide.

Over the last two decades, the increasing worldwide cultivation of elite cultivars has led to a significant reduction in genetic diversity of crop species (Warburton et al. 2008; Reif et al. 2005; Roussel et al. 2004) particularly in cereals (Christiansen et al. 2002; Donini et al. 2000; Koebner et al. 2003). This highlights the role of the genetic conservation

of crop germplasm (Frankel and Bennett 1970). Fortunately, genebanks are available as reservoirs of crop genetic diversity. However, genebanks contain duplicated accessions, errors in classification and omissions of key data that would guide plant breeders in the utilization of this germplasm. In particular, many genebank accessions lack data related to phenotype or imagery that could allow correction of classification errors. Thus, this work gathered a wealth of information about the oat genetic and phenotyping diversity of oat along the Mediterranean rim, which could be used to correct the genebanks database errors. The interactive map presented in this work provides an essential resource for future oat breeding for Mediterranean areas, by which oat breeders may select accessions taking into account additional agronomic and geospatial data. The map has been conceived as a living tool in which further information developed on these accessions will be added, including disease and abiotic stress resistance, to facilitate oat breeding under a scenario of climate change. Where possible, we have updated or corrected classifications made by genebanks such that the interactive map includes new information consistent with the genetic analyses presented here. When subspecies classification differed from that of the genebank this was highlighted in the map information, so genebanks can check it if desired.

Table 1 Potential candidate genes of significant markers in two sequenced diploid oat genomes, *Avena atlantica* (AA) and *Avena eriantha* (CC) described by Maughan et al. (2019) using the function BlastX of the BLAST algorithm (Altschul et al. 1990). Blastx 2.2.30

Marker	Linkage group	Chromosome	Nearby gene	Gene/marker distance (kbp)	Position relative to marker	Results BLASTn NCBI				
						Description	Species	E-value	Cov (%)	Ident (%) Accession number
avgbs_80864.1	Mrg13	A. eriantha AE05	AE025378-RA	10	Downstream	NAC domain-containing protein 92-like	<i>A. tauschii</i>	0.0	99	83.38 XM_020312824.1
						30S ribosomal protein 3, chloroplast	<i>B. distachyon</i>	8E-152	17	87.42 XM_003579859.4
						Cyclic pyranophosphate synthase, mitochondrial	<i>B. distachyon</i>	0.0	21	90.88 XM_024456111.1
						Gamma gliadin, omega gliadin-B6, LMW-B2, and LMW-B3 genes	<i>T. aestivum</i>	0.0	68	78.77 MG560141.1
avgbs_cluster_4923.1	Mrg01	A. atlantica AA02	AA010050-RA	27	Downstream	Vacuolar amino acid transporter 1-like	<i>A. tauschii</i>	0.0	72	88.28 XM_020312820.1
						Phosphatidylinositol N-acetylglucosaminyltransferase subunit P-related	<i>Z. mays</i>	0.0	85	74.18 XM_008659956.2
						Rop guanine nucleotide exchange factor 3-like	<i>H. vulgare</i>	0.0	97	90.13 KAE8786664.1
						Pectin acetyltransferase 9	<i>B. distachyon</i>	6E-65	21	81.85 XM_024460397.1
						Proline dehydrogenase 2, mitochondrial-like	<i>A. tauschii</i>	0.0	60	93.93 XM_020328890.1
						GDSL esterase/lipase At5g5050-like	<i>A. tauschii</i>	2E-29	1	83.12 XM_020305331.1
						Branched-chain amino-acid aminotransferase 5,	<i>A. tauschii</i>	0.0	100	91.06 XM_020325288.1

Table 1 (continued)

Marker	Linkage group	Chromosome	Nearby gene	Gene/marker distance (kbp)	Position relative to marker	Results BLASTn NCBI		E-value	Cov (%)	Ident (%)	Accession number
						Description	Species				
avgs_cluster_1918.1	Mrg08	A. eriantha AE06	AE037406-RA	0.75	Downstream	NAC domain-containing protein 8	<i>H. vulgare</i>	0.0	97	84.32	KA8770891.1
			AE037407-RA	133.402	Downstream	Histone H2B.1-like	<i>A. tauschii</i>	0.0	72	95.5	XM_020293063.1
			AE037408-RA	144.59	Downstream	Clone B1 GASR7 gene	<i>T. aestivum</i>	6E-75	17	94.74	KJ000053.1
			AE037414-RA	251.266	Downstream	Transcription factor TGAL3-like	<i>A. tauschii</i>	1E-96	29	90.61	XM_020299922.1
			AE037405-RA	15.178	Upstream	Autophagy-related 2	<i>A. tauschii</i>	0.0	50	88.36	XM_020305375.1

GBS marker positions in the diploid genomes were declared based on the best BLAST match of the reference sequences of the GBS markers

This work has revealed high genetic and phenotypic variability and a complex population structure in a Mediterranean oat collection. The GBS data provided high-density, genome-wide and genetically mapped marker set. Based on the Bayesian STRUCTURE analysis, the Mediterranean oat landraces clustered with high confidence into two subpopulations, corresponding to the *sativa* and *byzantina* subspecies, with higher genetic variability and more subpopulation structure in the *sativa* group. This was also observed in a previous genetic diversity study on Spanish accessions (Montilla-Bascón et al. 2013) suggesting that red oats have been less subjected to population divergence than white oats. As previously reported, we found a significant genetic divergence between white and red oats (Fu et al. 2005; Newell et al. 2012), supporting the independent domestication hypothesis of these two subspecies (Zohary and Hopf 2000). While the Fertile Crescent is considered the center of origin of *A. sativa*, it is argued that white oat evolved during the migration of its ancestors northward, where it was eventually established as a primary crop (Thomas 1995). In contrast, red oats entered cultivation from the western part of the Mediterranean (Loskutov 2008).

Interestingly, the oat cultivars included in the collection for comparison purposes clustered nearer to the north- and east-Mediterranean landraces than to the west-Mediterranean landraces. This could explain the low adaptation of most of the cultivars currently grown in Spain, the majority of which are spring cultivars bred in northern areas (Prats et al. 2014; Risipail et al. 2018; Sánchez-Martín et al. 2017). Our previous study based on 36 north-European cultivars indicated a narrow genetic basis among the tested varieties. Similarly, a lack of genetic diversity has been reported in a collection of Canadian and Chinese oat cultivars, leading to a call for broadening the genetic variation for a sustainable oat improvement (Baohong et al. 2003; Fu et al. 2004). In contrast, the current Mediterranean oat collection showed no tight clustering of cultivars. Although cultivars did not cluster near to Spanish accessions as stated above, they were distributed over the PCA plot near landraces from various regions and distributed between the white and red oat clusters. This pattern suggests that different landraces may have contributed to the improvement of these cultivars and broadened their genetic base.

Adding the data gathered previously on a set of North American oat accessions to our data confirmed the high level of diversity in the Mediterranean collection. It also provided interesting information on the potential breeding history of oat accessions. North American oat landraces clustered primarily with the white oat landraces, suggesting that they had not significantly differentiated after migration to the new world, and that they had integrated little or no red oat germplasm. The fact that southern USA landraces grouped with the western Mediterranean landraces, most of which coming

from Spain, may reflect both human migration patterns, as well as an adaptation to a Mediterranean-type production regime.

Population structure and LD are primary obstacles to the successful identification of linkage-based associations between markers and phenotypic traits (Buckler and Thornsberry 2002). Indeed, the power of association studies depends on the LD between the functional allele responsible for the observed phenotype and the marker. Our data showed extensive LD among markers, in agreement with previous studies (Montilla-Bascón et al. 2015; Rispaill et al. 2018) and a sufficient marker coverage for a GWAS approach (Newell et al. 2011). Furthermore, our results revealed a variable LD pattern in the different linkage groups that was unique to each of the *sativa* and *byzantina* groups. Such patterns have also been found for other crop species such as maize, wheat and sugar beet (Li et al. 2011; Van Inghelandt et al. 2011). Interestingly, Mrg04 showed a much slower LD decay in subs. *sativa* than in *byzantina*. According to Alheit et al. (2012), regions containing strong LD blocks probably harbor QTL responsible for agronomically important traits that reduce LD decay. Thus, our data suggests the presence of QTLs responsible for agronomically important traits in this region, but only in subs. *sativa*. In addition, according to Chao et al. (2010), this divergence in the extent of LD might be attributed to unique breeding histories and selection pressures applied to genes located in different genomes/chromosomes during the process of cultivar development.

Based on LD and structure results, association studies were performed separately for the two subspecies. The GWAS analysis was performed to test the suitability of the GBS approach and the Mediterranean oat collection for association studies by identifying significant associations with heading date. This trait is of fundamental importance for the local adaptation of oats in Mediterranean environments, since oats are particularly susceptible to drought during flowering time and heading date plays an important role for drought avoidance responses (Mahadevan et al. 2016). Our data highlighted markers associated with heading date on three chromosomes. Associations on these chromosomes have been previously reported on Mrg1 at 117.9 cM (Tumino et al. 2016) and at 39.3 cM (Esvelt Klos et al. 2016), on Mrg8 at 71 cM (Bekele et al. 2018) and on Mrg13 at 58.6 cM (Tumino et al. 2016) and at 30.3, 33 and 35.9 cM (Esvelt Klos et al. 2016). Specifically, the associated marker that we identified on Mrg13 was only 1.4 cM apart from the QTL highlighted by Tumino et al. (2016), further validating the presence of a QTL for heading date in this region. By contrast, the markers identified on Mrg1 and Mrg08 were located in different regions than those reported previously suggesting the presence of additional QTLs. Interestingly, (Bekele et al. 2018; Esvelt Klos et al. 2016; Tumino et al. 2016) highlighted markers associated with heading date in

Mrg02, which were not found in our study. This differences might be due to the different photoperiod during flowering for autumn and spring sowed oats.

Interestingly, two of the associated markers were located near genes with high similarity to NAC-domain transcription factors that have been involved in many developmental processes, including flowering. In particular, the closest gene on Mrg13 showed similarity to a NAC domain-containing protein 100 (NAC100), a transcription factor than controls ethylene-regulated cell expansion in flower petals (Pei et al. 2013). The closest gene to Mrg08 marker showed high homology with a suppressor of gamma response 1 (SOG1 / AtNAC8), which is a transcription factor governing multiple responses to DNA damage (Yoshiyama et al. 2009) and showing high expression in *Arabidopsis thaliana* during flowering (Klepikova et al. 2016). Comparison of the Mrg08 region between the *A. byzantina* and *A. sativa* preliminary genome assemblies also identified as potential candidate genes a GASA4-like gibberellin responsive gene and a gene with homology to miR172. Among the genes near the Mrg01 associated marker, one showed similarity with a phosphatidylinositol N-acetylglucosaminyltransferase subunit P, a protein in biosynthesis that is required for pollen germination and pollen tube growth (Koudiri et al. 2018; Lalanne et al. 2004). At this point, the influence of these genes on the time of heading is speculative, requiring further validation. Perfect GBS matches for all three of the high-confidence associations were also located in the recently released hexaploid oat genome. Unfortunately, a fully annotated version of this genome is not yet available so gene matches are tentative, and the reported genome positions may be interpreted at a later date. It is interesting to note, however, that the matches in the diploid genomes may not be located on the corresponding hexaploid chromosomes. For example, *avgbs_cluster_1918.1* is located on a C genome chromosome in *A. eriantha*, while it is on a D genome chromosome in the hexaploid. Oat is known to have undergone extensive genome rearrangement, and forthcoming studies will elucidate the locations of some of these rearrangements.

In conclusion, this work has provided a wealth of genetic and agro-climatic information in Mediterranean landraces and cultivars of oat. These accessions, genetic data and phenotypes will be a vital resource for further discovery-based research as well as for global oat genetic improvement.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00122-021-03805-2>.

Author contribution Statement FJC made most of the experimental work and data analysis. GMB participated in the experimental work. NAT and WAB supervised genetic analysis. CJH and TL contributed to the genetic analysis interpretation of data. NR and EP steered the research, designed experiments, and contributed to the interpretation

of results and writing of the manuscript. All authors also contributed to critical reading and writing.

Funding This work was supported by the Spanish Ministry of Science and Innovation [AGL2016-78965AGR, PID2019-104518RB-I00], (AEI/FEDER, UE) and regional government through the AGR-253 group, the European Regional and Social Development Funds, and Mobility Grant (FPI) 2018 from Ministry of Economy and Competitiveness of the Government of Spain [EEBB-I-18-12824]. FJC is holder of a FPI fellowship from the Spanish Ministry of Economy and Competitiveness [BES-2014-071044]. TL and CJH were supported by BBSRC grant BBS/E/W/0012843B.

Declarations

Conflicts of interest The authors declare that they have no conflict of interests.

Availability of data and material Genotype data were deposited in the public T3/oat database (<https://oat.triticeaetoolbox.org/breeders/trial/4667>). GBS FASTQ raw data have been submitted to the NCBI SRA (Bioproject PRJNA693576). Interactive map with passport and phenotyping data of the accessions is available at (https://www.google.com/maps/d/viewer?mid=1t-O7OUoUPJ_qY5qGQsq66O5zg7k). The data underlying the map and also marker data have been submitted to the DryadData and are accessible as Canales et al. 2021, <https://doi.org/10.5061/dryad.0gb5mkm0g>

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alheit KV, Maurer HP, Reif JC, Tucker MR, Hahn V, Weissmann EA, Wuerschum T (2012) Genome-wide evaluation of genetic diversity and linkage disequilibrium in winter and spring triticales (*x Triticosecale* Wittmack). *BMC Genom* 13:235. <https://doi.org/10.1186/1471-2164-13-235>
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
- Baohong G, Zhou X, Murphy JP (2003) Genetic variation within Chinese and western cultivated oat accessions. *Cereal Res Commun* 31:339–346. <https://doi.org/10.1007/BF03543363>
- Bekele WA, Wight CP, Chao S, Howarth CJ, Tinker NA (2018) Haplotype-based genotyping-by-sequencing in oat genome research. *Plant Biotechnol J* 16:1452–1463
- Benjamini Y, Hochberg Y (2000) On the adaptive control of the false discovery rate in multiple testing with independent statistics. *J Edu Behav Stat* 25:60–83. <https://doi.org/10.3102/10769986025001060>
- Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinform* 23:2633–2635
- Buckler ES, Thornsberry JM (2002) Plant molecular diversity and applications to genomics. *Curr Opin Plant Biol* 5:107–111
- Canales FJ et al (2021) Data set from: Population genomics of Mediterranean oat (*A. sativa*) reveals high genetic diversity and three loci for heading date. Dryad, Dataset. <https://doi.org/10.5061/dryad.0gb5mkm0g>
- Chalak L, Mzid R, Rizk W, Hmedeh H, Kabalan R, Breidy J, Hamadeh B, Machlab H, Rizk H, Elhajj S (2015) Performance of 50 Lebanese barley landraces (*Hordeum vulgare* L. subsp. *vulgare*) in two locations under rainfed conditions. *Annals Agric Sci* 60:325–334. <https://doi.org/10.1016/j.aaoas.2015.11.005>
- Chao S, Dubcovsky J, Dvorak J, Luo M-C, Baenziger SP, Matnyazov R, Clark DR, Talbert LE, Anderson JA, Dreisigacker S, Glover K, Chen J, Campbell K, Bruckner PL, Rudd JC, Haley S, Carver BF, Perry S, Sorrells ME, Akhunov ED (2010) Population- and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (*Triticum aestivum* L.). *BMC Genom* 11:727. <https://doi.org/10.1186/1471-2164-11-727>
- Christiansen MJ, Andersen SB, Ortiz R (2002) Diversity changes in an intensively bred wheat germplasm during the 20th century. *Mol Breeding* 9:1–11. <https://doi.org/10.1023/A:1019234323372>
- Donini P, Law JR, Koebner RMD, Reeves JC, Cooke RJ (2000) Temporal trends in the diversity of UK wheat. *Theor Appl Genet* 100:912–917. <https://doi.org/10.1007/s001220051370>
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Res* 4:359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Ehlers W (1989) Transpiration efficiency of oat. *Agronomy J* 81:810–817. <https://doi.org/10.2134/agronj1989.00021962008100050023x>
- Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-Sequencing (GBS) approach for high diversity species. *PLoS ONE* 6:e19379. <https://doi.org/10.1371/journal.pone.0019379>
- Esvelt Klos K, Huang Y-F, Bekele WA, Obert DE, Babiker E, Beattie AD, Bjornstad A, Bonman JM, Carson ML, Chao S, Gnanesh BN, Griffiths I, Harrison SA, Howarth CJ, Hu G, Ibrahim A, Islamovic E, Jackson EW, Jannink J-L, Kolb FL, McMullen MS, Fetch JM, Murphy JP, Ohm HW, Rines HW, Rossnagel BG, Schlueter JA, Sorrells ME, Wight CP, Yan W, Tinker NA (2016) Population genomics related to adaptation in elite oat germplasm. *Plant Genom*. <https://doi.org/10.3835/plantgenome2015.10.0103>
- Esvelt Klos K, Yimer BA, Babiker EM, Beattie AD, Bonman JM, Carson ML, Chong J, Harrison SA, Ibrahim A, Kolb FL, McCartney CA, McMullen M, Mitchell Fetch J, Mohammadi M, Murphy JP, Tinker NA (2017) Genome-wide association mapping of crown rust resistance in oat elite germplasm. *Plant Genom*. <https://doi.org/10.3835/plantgenome2016.10.0107>
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genet* 164:1567–1587
- Frankel OH (1970) Genetic conservation in perspective. In: Frankel OH, Bennett E (eds) *Genetic Resources in Plants - Their Exploration and Conservation* International Biological Programme Handbook No 11. Blackwell Scientific Publications, Oxford, pp 469–489

- Fu YB, Kibite S, Richards KW (2004) Amplified fragment length polymorphism analysis of 96 Canadian oat cultivars released between 1886 and 2001. *Can J Plant Sci* 84:23–30. <https://doi.org/10.4141/P02-141>
- Fu YB, Peterson GW, Morrison MJ (2007) Genetic diversity of Canadian soybean cultivars and exotic germplasm revealed by simple sequence repeat markers. *Crop Sci* 47:1947–1954. <https://doi.org/10.2135/cropsci2006.12.0843>
- Fu YB, Peterson GW, Williams D, Richards KW, Fetch JM (2005) Patterns of AFLP variation in a core subset of cultivated hexaploid oat germplasm. *Theor Appl Genet* 111:530–539
- Google-LTD (2021) Google Maps of Mediterranean area. Retrieved from <https://www.google.es/maps/@44.6530938,18.9948982,4z>
- Huang Y-F, Poland JA, Wight CP, Jackson EW, Tinker NA (2014) Using genotyping-by-sequencing (GBS) for genomic discovery in cultivated oat. *PLoS ONE* 9(7):e102448. <https://doi.org/10.1371/journal.pone.0102448>
- Jacoby WG (2000) Loess: a nonparametric, graphical tool for depicting relationships between variables. *Elect Stud* 19:577–613. [https://doi.org/10.1016/S0261-3794\(99\)00028-1](https://doi.org/10.1016/S0261-3794(99)00028-1)
- Kebede AZ, Admassu-Yimer B, Bekele WA, Gordon T, Bonman JM, Babiker E, Jin Y, Gale S, Wight CP, Tinker NA, Menzies JG, Beattie AD, Mitchell Fetch J, Fetch TG, Esvelt Klos K, McCartney CA (2020) Mapping of the stem rust resistance gene *Pg13* in cultivated oat. *Theor Appl Genet* 133:259–270
- Klepikova AV, Kasianov AS, Gerasimov ES, Logacheva MD, Penin AA (2016) A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling. *Plant J* 88:1058–1070
- Koebner RMD, Summers RW (2003) 21st century wheat breeding: plot selection or plate detection? *Trends Biotechnol* 21:59–63
- Kouidri A, Baumann U, Okada T, Baes M, Tucker EJ, Whitford R (2018) Wheat TaMs1 is aglycosylphosphatidylinositol-anchored lipid transfer protein necessary for pollen development. *BMC Plant Biol* 18:332
- Lalanne E, Honys D, Johnson A, Borner GHH, Lilley KS, Dupree P, Grossniklaus U, Twell D (2004) SETH1 and SETH2, two components of the glycosylphosphatidylinositol anchor biosynthetic pathway, are required for pollen germination and tube growth in *Arabidopsis*. *Plant Cell* 16:229–240
- Li Y, Haseneyer G, Schoen C-C, Ankerst D, Korzun V, Wilde P, Bauer E (2011) High levels of nucleotide diversity and fast decline of linkage disequilibrium in rye (*Secale cereale* L.) genes involved in frost response. *BMC Plant Biol* 11:6. <https://doi.org/10.1186/1471-2229-11-6>
- Loskutov IG (2008) On evolutionary pathways of *Avena* species. *Genet Res Crop Evol* 55:211–220. <https://doi.org/10.1007/s10722-007-9229-2>
- Lu F, Lipka AE, Glaubitz J, Elshire R, Cherney JH, Casler MD, Buckler ES, Costich DE (2013) Switchgrass genomic diversity, ploidy, and evolution: Novel insights from a network-based SNP discovery Protocol. *Plos Genet* 9(1):e1003215. <https://doi.org/10.1371/journal.pgen.1003215>
- Mahadevan M, Calderini DF, Zwer PK, Sadras VO (2016) The critical period for yield determination in oat (*Avena sativa* L.). *Field Crops Res* 199:109–116. <https://doi.org/10.1016/j.fcr.2016.09.021>
- Maughan PJ, Lee R, Walstead R, Vickerstaff RJ, Fogarty MC, Brouwer CR, Reid RR, Jay JJ, Bekele WA, Jackson EW, Tinker NA, Langdon T, Schlueter JA, Jellen EN (2019) Genomic insights from the first chromosome-scale assemblies of oat (*Avena* spp) diploid species. *BMC Biol* 17:92
- Montilla-Bascón G, Rispail N, Sánchez-Martín J, Rubiales D, Mur LAJ, Langdon T, Howarth C, Prats E (2015) Genome-wide association study for crown rust (*Puccinia coronata* f. sp. avenae) and powdery mildew (*Blumeria graminis* f. sp. avenae) resistance in an oat (*Avena sativa*) collection of commercial varieties and landraces. *Front Plant Sci* 6:103. <https://doi.org/10.3389/fpls.2015.00103>
- Montilla-Bascón G, Sanchez-Martin J, Rispail N, Rubiales D, Mur L, Langdon T, Griffiths I, Howarth C, Prats E (2013) Genetic diversity and population structure among oat cultivars and landraces. *Plant Mol Biol Report* 31:1305–1314. <https://doi.org/10.1007/s11105-013-0598-8>
- Murray MG, Thompson WF (1980) Rapid isolation of high Molecular-Weight plant DNA. *Nucleic Acids Res* 8:4321–4325
- Newell MA, Asoro FG, Scott MP, White PJ, Beavis WD, Jannink J-L (2012) Genome-wide association study for oat (*Avena sativa* L.) beta-glucan concentration using germplasm of worldwide origin. *Theor Appl Genet* 125:1687–1696
- Newell MA, Cook D, Tinker NA, Jannink JL (2011) Population structure and linkage disequilibrium in oat (*Avena sativa* L.): implications for genome-wide association studies. *Theor Appl Genet* 122:623–632
- Newton AC, Akar T, Baresel JP, Bebeli PJ, Bettencourt E, Bladenopoulos KV, Czembor JH, Fasoula DA, Katsiotis A, Koutis K, Koutsika-Sotiriou M, Kovacs G, Larsson H, de Carvalho M, Rubiales D, Russell J, Dos Santos TMM, Pato MCV (2010) Cereal landraces for sustainable agriculture. *Rev Agron Sustain Dev* 30:237–269. <https://doi.org/10.1051/agro/2009032>
- Pei H, Ma N, Tian J, Luo J, Chen J, Li J, Zheng Y, Chen X, Fei Z, Gao J (2013) An NAC transcription factor controls ethylene-regulated cell expansion in flower petals. *Plant Physiol* 163:775–791
- Prats E, Sánchez-Martín J, Montilla-Bascón G, Rubiales D, Rispail N (2014) Overview and prospects of the oat crop in Spain. *Oat Newsllett* 51:9
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959. <https://doi.org/10.1093/genetics/155.2.945>
- Ramasamy RK, Ramasamy S, Bindroo BB, Naik VG (2014) STRU CTURE PLOT: a program for drawing elegant STRUCTURE bar plots in user friendly interface. *SpringerPlus* 3(1):431. <https://doi.org/10.1186/2193-1801-3-431>
- Reed E, Nunez S, Kulp D, Qian J, Reilly MP, Foulkes AS (2015) A guide to genome-wide association analysis and post-analytic interrogation. *Stat Med* 34:3769–3792
- Reif JC, Zhang P, Dreisigacker S, Warburton ML, van Ginkel M, Hoisington D, Bohn M, Melchinger AE (2005) Wheat genetic diversity trends during domestication and breeding. *Theor Appl Genet* 110:859–864
- Rispail N, Montilla-Bascón G, Sanchez-Martin J, Flores F, Howarth C, Langdon T, Rubiales D, Prats E (2018) Multi-environmental trials reveal genetic plasticity of oat agronomic traits associated with climate variable changes. *Front Plant Sci*. <https://doi.org/10.3389/fpls.2018.01358>
- Roussel V, Koenig J, Beckert M, Balfourier F (2004) Molecular diversity in French bread wheat accessions related to temporal trends and breeding programmes. *Theor Appl Genet* 108:920–930
- Sánchez-Martín J, Rispail N, Flores F, Emeran AA, Sillero JC, Rubiales D, Prats E (2017) Higher rust resistance and similar yield of oat landraces versus cultivars under high temperature and drought. *Agron Sustain Dev* 37:3. <https://doi.org/10.1007/s13593-016-0407-5>
- Sánchez-Martín J, Rubiales D, Flores F, Emeran AA, Shtaya MJY, Sillero JC, Allagui MB, Prats E (2014) Adaptation of oat (*Avena sativa*) cultivars to autumn sowings in Mediterranean environments. *Field Crops Res* 156:111–122. <https://doi.org/10.1016/j.fcr.2013.10.018>
- Stevens EJ, Armstrong KW, Bezar HJ, Griffin WB, Hampton JG (2004) Fodder oats: an overview. In: Suttie JM, Reynolds SG (eds) *Fodder oats: A world overview*. FAO, Rome, pp 1–9

- Storey JD (2002) A direct approach to false discovery rates. *J Roy Stat Soc Ser B (Stat Methodol)* 64(3):479–498. <https://doi.org/10.1111/1467-9868.00346>
- Thomas H (1995) Oats. In: Smartt JS, N.W. (ed) *Evolution of crop plants*, Second edn. Longman, London, United Kingdom, pp 133–137
- Tinker NA, Bekele WA, Hattori J (2016) Haplotag: Software for haplotype-based genotyping-by-sequencing analysis. *G3-Genes Genomes Genetics* 6:857–863
- Tumino G, Voorrips RE, Rizza F, Badeck FW, Morcia C, Ghizzoni R, Germeier CU, Paulo M-J, Terzi V, Smulders MJM (2016) Population structure and genome-wide association analysis for frost tolerance in oat using continuous SNP array signal intensity ratios. *Theor Appl Genet* 129:1711–1724
- Van Inghelandt D, Reif JC, Dhillon BS, Flament P, Melchinger AE (2011) Extent and genome-wide distribution of linkage disequilibrium in commercial maize germplasm. *Theor Appl Genet* 123:11–20
- Warburton ML, Reif JC, Frisch M, Bohn M, Bedoya C, Xia XC, Crossa J, Franco J, Hoisington D, Pixley K, Taba S, Melchinger AE (2008) Genetic diversity in CIMMYT nontemperate maize germplasm: Landraces, open pollinated varieties, and inbred lines. *Crop Sci* 48:617–624. <https://doi.org/10.2135/cropsci2007.02.0103>
- Yoshiyama K, Conklin PA, Huefner ND, Britt AB (2009) Suppressor of gamma response 1 (SOG1) encodes a putative transcription factor governing multiple responses to DNA damage. *Proceedings of the National Academy of Sciences of the United States of America* 106:12843–12848
- Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genet* 38:203–208
- Zohary D, Hopf M (2000) *Domestication of plants in the old world*, 3rd edn. Oxford University Press, Oxford-New York

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.