Mizuno, Nobuyuki; Toyoshima, Masami; Fujita, Miki; Fukuda, Shota; Kobayashi, Yasufumi; Ueno, Mariko; Tanaka, Kojiro; ... Mizukoshi, Hiroharu; Yasui, Yasuo; Fujita, Yasunari

RIGHT:

OXFORD

Research Article

The genotype-dependent phenotypic landscape of quinoa in salt tolerance and key growth traits

Nobuyuki Mizuno[1†], Masami Toyoshima[2], Miki Fujita[3], Shota Fukuda[4‡], Yasufumi Kobayashi[2], Mariko Ueno[1], Kojiro Tanaka[5], Tsutomu Tanaka[5], Eiji Nishihara[4], Hiroharu Mizukoshi[5], Yasuo Yasui[1]*, and Yasunari Fujita[2,6]*

[1]Graduate School of Agriculture, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan, [2]Biological Resources and Post-harvest Division, Japan International Research Center for Agricultural Sciences (JIRCAS), Tsukuba, Ibaraki 305-8686, Japan, [3]Gene Discovery Research Group, RIKEN Center for Sustainable Resource Science, Tsukuba, Ibaraki 305-0074, Japan, [4]Faculty of Agriculture, Tottori University, Tottori, Tottori 680-8550, Japan, [5]Technology Development Group, Actree Corporation, Hakusan, Ishikawa 924-0053, Japan, and [6]Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-8577, Japan

*To whom correspondence should be addressed. +81 29 838 6641. Fax. +81 29 838 6643. Email: yasuf@affrc.go.jp (Y.F.); Tel. +81 75 753 6480. Email: yasyas@kais.kyoto-u.ac.jp (Y.Y.)

[†]Present address: Division of Basic Research, Institute of Crop Science, National Agriculture and Food Research Organization (NARO), Tsukuba, Ibaraki 305-8518, Japan.
[‡]Present address: Bioscicence Research Laboratory, Central Research Laboratory, Ishihara Sangyo Kaisha, Ltd., Kusatsu, Shiga 525-0025, Japan.

Abstract

Cultivation of quinoa (*Chenopodium quinoa*), an annual pseudocereal crop that originated in the Andes, is spreading globally. Because quinoa is highly nutritious and resistant to multiple abiotic stresses, it is emerging as a valuable crop to provide food and nutrition security worldwide. However, molecular analyses have been hindered by the genetic heterogeneity resulting from partial outcrossing. In this study, we generated 136 inbred quinoa lines as a basis for the molecular identification and characterization of gene functions in quinoa through genotyping and phenotyping. Following genotyping-by-sequencing analysis of the inbred lines, we selected 5,753 single-nucleotide polymorphisms (SNPs) in the quinoa genome. Based on these SNPs, we show that our quinoa inbred lines fall into three genetic sub-populations. Moreover, we measured phenotypes, such as salt tolerance and key growth traits in the inbred quinoa lines and generated a heatmap that provides a succinct overview of the genotype–phenotype relationship between inbred quinoa lines. We also demonstrate that, in contrast to northern highland lines, most lowland and southern highland lines can germinate even under high salinity conditions. These findings provide a basis for the molecular elucidation and genetic improvement of quinoa and improve our understanding of the evolutionary process underlying quinoa domestication.

Key words: quinoa, genotype, inbred line, phenotype, salt stress

1. Introduction

The decline in global crop diversity constitutes a potential threat to food and nutrition security.[1] Recently, the increasing global vulnerability of the food supply caused by climate change and rapid population growth has made underutilized, orphan, and neglected crops potentially very important for future food and nutrition security.[2] One such potential underutilized crop is quinoa (*Chenopodium quinoa* Willd.),[3] an annual pseudocereal of the Amaranthaceae family that also includes spinach (*Spinacia oleracea* L.) and sugar beet (*Beta vulgaris* L.). Quinoa seeds and leaves have exceptional nutritional profiles, as they contain a constellation of minerals, vitamins, dietary fibre, fats, and high-quality proteins with a balanced amino acid profile.[4–7] Quinoa also thrives in a wide range of marginal environments in terms of altitude, annual precipitation, temperature, and soil conditions and is thus resistant to multiple abiotic stresses such as high salinity, drought, low temperatures, and frost.[8–13]

Quinoa was originally domesticated in the Andean region of South America at least 7,500 years ago[14] and served as a staple food for indigenous people of the Andes in Pre-Columbian times.[15] However, Spanish conquerors in the 1500s forbade the cultivation and consumption of quinoa because it was worshiped as a sacred 'mother grain' used in indigenous ceremonies in non-Christian religion.[15] Therefore, in contrast to potato (*Solanum tuberosum* L.) and tomato (*Solanum lycopersicum* L.), which also originated in the Andes but expanded globally after their discovery by the Conquistadors, quinoa remained a neglected crop in South America for about 500 years until the 20th century.[16] In 1975, the National Academy of Sciences of the United States (NAS) selected quinoa as one of the 36 underexploited plants with promising economic value.[17] In 1993, the US National Aeronautics and Space Administration (NASA) took an interest in quinoa because it possesses desirable food qualities for astronauts on long-duration space missions.[18] The United Nations declared 2013 as the International Year of Quinoa to raise public awareness of the nutritional benefits of this durable plant in terms of food and nutrition security worldwide.[19] Nowadays, although quinoa is cultivated mostly in Bolivia and Peru, quinoa cultivation is being attempted in over 95 countries.[19]

High salinity stress poses a major environmental threat to agriculture, impairing crop production in ~20% of global irrigated land.[20] Seed germination and subsequent seedling establishment are crucial steps in the life cycle of plants, especially under harsh environments such as salt-affected lands.[21] Since quinoa has been recognized as a salt-tolerant crop and a potential candidate for introduction into marginal areas,[22] many studies have focussed on salinity tolerance during germination in quinoa.[21] However, as few studies have made use of genotyped quinoa accessions, little is known about the molecular basis of salt tolerance, a prerequisite for the improvement of crop yield on saline soils.

Quinoa is an allotetraploid ($2n = 4x = 36$) with an estimated genome size of 1.5 Gbp, composed of the two sub-genomes A and B.[23–26] The underlying genome complexity caused by its allotetraploidy have hampered molecular analysis of quinoa genes over the years. However, the quinoa genome was recently sequenced by three independent research groups,[12,27,28] opening the door to revealing the molecular mechanisms governing the development of the mysterious quinoa plant. Moreover, quinoa plants exhibit partial outcrossing due to having both hermaphrodite and female flowers on the same plant,[29,30] resulting in genetic heterogeneity that has also impeded the rigorous molecular analysis of quinoa.

To overcome this substantial limitation, we generated 136 inbred quinoa lines as a basis for molecular studies aimed at identifying genes and characterizing gene function in quinoa through large-scale genotyping and phenotyping efforts. Armed with the inbred lines, we performed a genotyping-by-sequencing (GBS) analysis and thus detected 5,753 single-nucleotide polymorphisms (SNPs) in the quinoa genome. Based on the SNPs, we show that our quinoa inbred lines are genetically divided into three sub-populations and that the highlands of South America near Bolivia and Peru are the centre of diversity of cultivated quinoa. Further, we analysed the phenotypes of the inbred quinoa lines, which we combined with our genetic data into a heatmap that allows a succinct visual overview of the genotype–phenotype relationship among the inbred quinoa lines. We also established that, unlike the northern highland lines, the seeds of the lowland and southern highland lines can germinate under high salinity conditions (600 mM NaCl). These findings will contribute to the molecular basis needed to compensate for the lost breeding opportunities and studies of quinoa over the past 500 years.

2. Materials and methods

2.1. Plant materials, generation of quinoa inbred lines, and growth conditions

Quinoa seeds (except Kd and Iw lines) were obtained from the Germplasm Resources Information Network (GRIN) of the US Department of Agriculture (USDA) (Supplementary Table S1). The sequenced and model inbred quinoa line Kd was established at Japan International Research Center for Agricultural Sciences (JIRCAS) by single-seed propagation in the absence of all other quinoa plants for over 20 years at Kyoto University, Japan.[12] Another quinoa line used for plant virus research, Iw, was kindly provided by Dr Nobuyuki Yoshikawa, Iwate University, Japan.[31] Inbred quinoa lines used in this study were generated by self-crossing through single-seed descent of the quinoa lines collected at JIRCAS in Tsukuba, Japan, as described previously.[12] The seeds of inbred quinoa lines derived from the USDA were named J001–J144, excluding J062 and J070, because a USDA accession often contains seeds with multiple distinct phenotypes in terms of colour and size and none of the J062 and J070 seeds could germinate (Supplementary Table S1). For the bulking of inbred quinoa seeds, all inflorescences of each plant were covered with non-woven pollination bags (Rizo, Tsukuba, Japan) to prevent cross-pollination in phytotrons or a greenhouse.

Quinoa plants were grown on soil in pots as described previously,[12] with minor modifications. Quinoa seeds were sown in a peat moss mixture (Jiffy Mix, Sakata Seeds, Yokohama, Japan) in a cell tray and were allowed to germinate in a growth chamber at 25°C under a 12-h-light/12-h-dark cycle ($75 \pm 25\,\mu$mol photons/m^2/s). After 13–14 days, the seedlings were transferred to a standard potting mix (Tsuchitaro, Sumitomo Forestry, Tokyo, Japan) in 1.2-l plant pots and were grown under ambient light in temperature-controlled phytotrons at JIRCAS with a temperature of 25 ± 5°C and relative humidity of 55 ± 25%. For the bulking of inbred quinoa seeds for genotyping and phenotyping, quinoa plants were grown under ambient light in a greenhouse with a temperature of 27 ± 5°C and relative humidity of 55 ± 40%. For the phenotypic evaluation of the quinoa inbred lines at Tottori University, Japan, quinoa seeds were sown in sandy soil from the Tottori sand dunes in a cell tray and were allowed to germinate under ambient light in a greenhouse with a temperature of 18 ± 13°C. After 7 days, the seedlings were

transferred to sandy soil in 1.5-l plant pots and were grown in the greenhouse.

2.2. GBS and SNP detection

We used 136 of the 142 lines for GBS analysis, since we needed to arrange the total number of samples to be a multiple of 48, including 8 control replicates and in accordance with the combination of the adapters. Total genomic DNA was extracted from fresh quinoa leaves as described previously,[32] using the AP1 buffer from the DNAeasy Plant Mini Kit (Qiagen) instead of the lysis buffer used by Yabe *et al.*[32] GBS reactions were carried out as described previously[33] with minor modifications. *Eco*RI and *Pst*I were used as restriction enzymes in this study. Barcode adaptors are listed in Supplementary Table S1. Barcode-labelled amplicons were size-selected (range $200-500$ bp) with AMPure XP beads (Beckman Coulter) and were sequenced on an Illumina HiSeq2500 System at Macrogen Japan (Kyoto, Japan) as paired-end reads. The reads generated in this study are available from the DDBJ/EMBL/NCBI under the accession numbers listed in Supplementary Table S1. In addition to GBS data, we used published whole genome sequencing (WGS) data of quinoa[12,27] for SNP detection; the relevant accessions are listed in Supplementary Table S2. Low quality reads and adaptors were trimmed using Trimmomatic-0.36[34] with the options 'SLIDINGWINDOW:4:25' and 'MINLEN:40'. The adaptor sequences used were 'CACGACGCTCTTCCGATCT' and 'ACCGCTCTTCCGATCTGTAA'. The trimmed paired-end reads were mapped against the quinoa reference genome (Cq_PI614886_genome_V1_pseudomolecule.fa[27]) using BWA 0.7.15[35] with -L 10 and -B 10 options. Single-end reads were filtered out from BAM files using samtools 1.3.1.[36] We then carried out a local realignment procedure using RealignerTargetCreator and IndelRealigner in GATK 3.7.[37] SNPs were detected using UnifiedGenotyper in GATK 3.7 with the option -out_mode EMIT_ALL_CONFIDENT_SITES. Heterozygous variants and nucleotide sites supported by fewer than three reads were converted to missing data using Tassel 5.2.38[38] and VCFtools 0.1.13,[39] respectively. Nucleotide sites with a minimum mean depth of 10 and with proportion of max-missing data over 0.8 were filtered using VCFtools 0.1.13.[39]

2.3. Construction of the phylogenetic tree and structure analysis

Since the depth of WGS reads[27] was much lower than those of GBS reads obtained in this study, the WGS data were largely eliminated if the uniform filter conditions were used in all analyses. Therefore, we used different filter conditions and plant materials for each analysis (Supplementary Table S3). For STRUCTURE analysis, we used 149 quinoa lines (136 GBS and 13 WGS samples). Heterozygous variants were converted to missing data using Tassel 5.2.38.[38] Minimum coverage depth was set to 3, and nucleotide sites with a proportion of max-missing data over 0.8 were filtered using VCFtools 0.1.13. We performed a population structure analysis using a model-based clustering method implemented in the STRUCTURE 2.3.4 software.[40] The algorithm was run with a burn-in length of 100,000 followed by 100,000 Markov Chain Monte Carlo simulations for estimating parameters. At least 10 runs were performed to estimate the mean likelihood for a number of populations (K) from 1 to 10 using the admixture model and correlated allele frequencies. An average log-likelihood value, LnP(D), and ΔK[41] was calculated across all runs for each K to infer the best fit to the data, as measured by the

likelihood score. Lines with membership probabilities of <80% for all groups were assigned to 'admixed' lines. For construction of the phylogenetic tree and principal component analysis (PCA), we used GBS data from 136 quinoa lines and WGS data from 13 quinoa accessions and 2 accessions of *C. hircinum* (a wild relative of cultivated quinoa). The resulting dataset, consisting of 1.54 Mbp derived from 7,586 SNPs, was used for Neighbour-joining (NJ) analysis. In the NJ analysis, we used SEQBOOT with 100 replicates, followed by DNAdist (with Kimura 2-parameter method), Neighbor and Consensus programmes from the PHYLIP package 3.6 (http://www.evolution.genetics.washington.edu/phylip.htlm (last accessed 21 September, 2020)). NJ tree was rooted using midpoint rooting and visualized by FigTree 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/ (last accessed 21 September, 2020)). PCA based on covariance was performed by Tassel 5.2.38[38] using 7,586 SNPs.

2.4. Genetic diversity analysis and linkage disequilibrium

Nucleotide diversity (π) and Weir and Cockerham mean F-statistics (Fst) were calculated by VCFtools 0.1.13. Sliding window analysis of π, SNP frequency and number of SNPs were represented in a circular plot using the Circos 0.69.[42] For SNP detection, we used 136 GBS samples in addition to Kd, because the sequencing depth of earlier WGS data[27] was insufficient for these analyses (Supplementary Table S3). Pairwise linkage disequilibrium (LD) (r^2) between the SNPs showing >0.05 minor allele frequency (MAF) was calculated using VCFtools 0.1.13. Average r^2 values were calculated in non-overlapping 100-kbp intervals.

2.5. Germination test

Seeds were germinated in 9 cm diameter Petri dishes containing triple sheets of filter paper (Toyo Roshi Kaisha, Ltd, Tokyo, Japan), soaked with 10 ml of water or saline solution (0, 300, and 600 mM NaCl). We supplemented the 600mM NaCl saline solution with 0.2% of plant preservative mixture (Plant Cell Technology, Inc., Washington DC, USA), a broad-spectrum biocide/fungicide for plant tissue culture. Germination assays were carried out at $25 \pm 1°C$ in the dark. We measured total hypocotyl and root length (THRL) of the germinated seedlings in 0, 300, and 600 mM saline solutions after 24 h, 4, and 14 days.

2.6. Phenotypic evaluation

Seeds from 136 lines were harvested at maturity, at which point we phenotyped six traits: 1,000-grain weight (TGW), plant height (PH), stem diameter (SD), leaf dry weight (LDW), seed yield per plant (SYP), and days to flowering (DF). TGW was calculated from the weight and the number of seeds counted by a WAVER IC-VAi seed counter (IDEX, Nagoya, Japan). SD was measured 1.0 cm above ground with a vernier caliper. LDW was measured after drying all leaves from individual plants harvested at maturity at 72°C for at least 72 h. SYP was measured on individual plants after husking seeds.

2.7. Statistical analysis of phenotypic data

A statistical analysis was performed using the R package version 3.3.3. A Shapiro−Wilk normality test assessed whether each phenotype distribution fit within the normal distribution. Statistically significant differences between phenotypic groups were determined by one-way ANOVA and Tukey−Kramer's HSD tests ($P < 0.05$). We

constructed the heatmap with Microsoft Excel for Mac. PCA was performed with the 'prcomp' function in R. The lines with missing data were eliminated prior to PCA.

2.8. Genome-wide association study

We performed genome-wide association study (GWAS) on 12 phenotypic traits (6 traits mentioned above and THRL under different conditions) using TASSEL 5.3.38. SNPs showing ≥ 0.05 MAF and >0.2 missing data across the 137 inbred lines were filtered from the 5,753 SNPs (see Section 2.3). Finally, we used 3,156 SNPs for general linear model (Q, GLM) and mixed linear model (Q + K, MLM) methods. Q-matrix derived from the STRUCTURE program was used for the GLM method. We performed 10,000 permutation runs to obtain the permutation-based significance in GLM analysis. The Q and kinsip (K) matrixes were used for the MLM method. K matrix was calculated based on Centered-IBS method by Tassel 5.2.38. Significant marker-trait associations (MTAs) were identified with a Benjamini−Hochberg false discovery rate (FDR) at the 5% level to correct for multiple testing.

3. Results and discussion

3.1. Generation of 136 inbred quinoa lines

Although quinoa is essentially self-pollinating, it does display genetic heterogeneity due to partial outcrossing due to the fact that individual plants carry both hermaphrodite and female flowers.[29,30] Heterozygous alleles at a given genetic locus within an individual accession were seen in 32% of quinoa accessions examined.[30] Nevertheless, to date, inbred quinoa lines have not been used in most phenotyping and genotyping studies of quinoa. Consequently, we first wished to develop inbred quinoa lines that would be a valuable resource for molecular genetics and biological studies of quinoa. We generated 136 inbred quinoa lines by self-crossing and single-seed descent of Kd,[12] Iw, and all quinoa lines obtained from the USDA-GRIN quinoa germplasm collection and then named these inbred lines J lines (J001–J144, excluding J062 and J070) since one USDA accession often contains seeds with multiple distinct phenotypes in terms of colour and size and none of the J062 and J070 seeds could germinate (Supplementary Table S1). We used pollination bags to prevent cross-pollination when plants set seeds. The ratio of heterozygous sites to all SNPs sites ranged from 0.0024 to 0.019% across all lines, with an average value of 0.0075% (Supplementary Table S1, also see below), indicating that our inbred lines were purified enough for genotyping and phenotyping analyses.

3.2. SNP discovery by GBS

To evaluate genetic diversity in quinoa, we used a GBS approach to identify SNPs in all inbred quinoa lines. We sequenced DNA libraries derived from *Eco*RI/*Pst*I double digests of genomic DNA from our 136 quinoa inbred lines on the HiSeq2500 platform. After primary quality filtering steps, we generated a total of 560 million clean reads (covering 56.6 Gbp), with an average of 8,723,466 reads (corresponding to 880.3 Mbp) per line (Supplementary Table S1). On average, 88.0% of trimmed paired-end reads correctly mapped to the reference genome.[27] The average depth for all nucleotide sites in the GBS samples was 148.1. After filtering the nucleotide sites detected by GBS for all 136 samples and by WGS for Kd, we retained 1,709,689 nucleotide sites, which correspond to 0.14% of the reference genome (Supplementary Table S4). From these 1,709,689

nucleotide sites, we detected 5,753 SNP sites, ranging from 88 (Chr13) to 719 (Chr07) per chromosome. The number of SNPs per chromosome showed a high correlation with the length of each assembled chromosome ($R^2 = 0.93$), as expected. These results suggest that the GBS reads are properly mapped to the reference sequence, since the low mis-mapping rate of GBS reads is presumably due to the high nucleotide divergence between the A and B genomes[12,27] (pair-wise nucleotide divergence at synonymous site = 0.08−0.1). SNPs were distributed over the entire genome (Fig. 1). The average SNP density across all chromosomes was 0.0034 (i.e., one SNP per 297 bp) and ranged from 0.0025 (one SNP per 407 bp, Chr17) to 0.0041 (one SNP per 243 bp, Chr04) depending on the chromosome. These data indicate that the 5,753 SNPs we obtained by GBS can be used as genome-wide markers for population and phylogenetic analyses, as described below.

3.3. Population structure, genetic diversity, and evolutionary history of domestication in quinoa

To investigate the population structure within *C. quinoa*, we conducted multiple analyses (STRUCTURE, NJ phylogenetic, and PCA) using 5,529 SNPs obtained from both GBS data for 136 lines and WGS data for 13 accessions. Assignment analysis using the program STRUCTURE indicated that the highest ΔK value was observed for $K = 2$ (15,466) followed by $K = 3$ (1,490), whereas all larger Ks ($K \geq 4$) had substantially lower ΔK values (Figs. 2A and B). With $K = 2$ (Fig. 2C), one population was mainly composed of lines from Bolivia and Peru, whereas the other population largely contained lines from Chile and a set of US accessions that lack reliable passport data (Supplementary Table S5). These two populations corresponded to those named highland and lowland (Coastal), respectively, in previous reports[27,30,43,44,45] (Fig. 2D). Many of the admixed lines (lines with membership probabilities of <80%) were lowland lines (Fig. 2C). This might indicate low and/or recent gene flow had occurred from highland to lowland population. With $K = 3$ (Fig. 2C), the highland population was further divided into two sub-populations, corresponding to southern and northern highland sub-populations.[30,43,45] The southern highland sub-population was mainly composed of lines from Bolivia, whereas the northern highland sub-population contained lines from Peru. Notably, although we identified a number of admixtures consisting of lowland and southern highland genetic background, we failed to detect any cases of admixtures between lowland and northern highland genetic material (Fig. 2C), suggesting a closer relationship between southern highland and lowland populations than between northern highland and lowland populations, as might be expected due to physical isolation (Fig. 2D). Considering that these data are consistent with previous reports,[30,45,46] it would be more plausible that recent gene flow mainly occurred from the Bolivian population to the Chilean population (Fig. 2D).

We also conducted NJ analysis and PCA to examine the phylogenetic relationships among *C. quinoa* lines, using two *C. hircinum* accessions (a wild relative of domesticated quinoa) as outliers. These two different approaches strengthen our view that quinoa lines are divided into three clusters: Lowland, northern highland, and southern highland (Figs 2D and 3; Supplementary Fig. S1). Four inbred lines (J010, J027, J032, and J040) assigned to southern highland were included in the lowland cluster of the phylogenetic NJ tree (Fig. 3), supporting the notion that the lowland cluster is more closely related to southern highland lines than it is to the northern highland cluster. One quinoa line (O654) was positioned in an
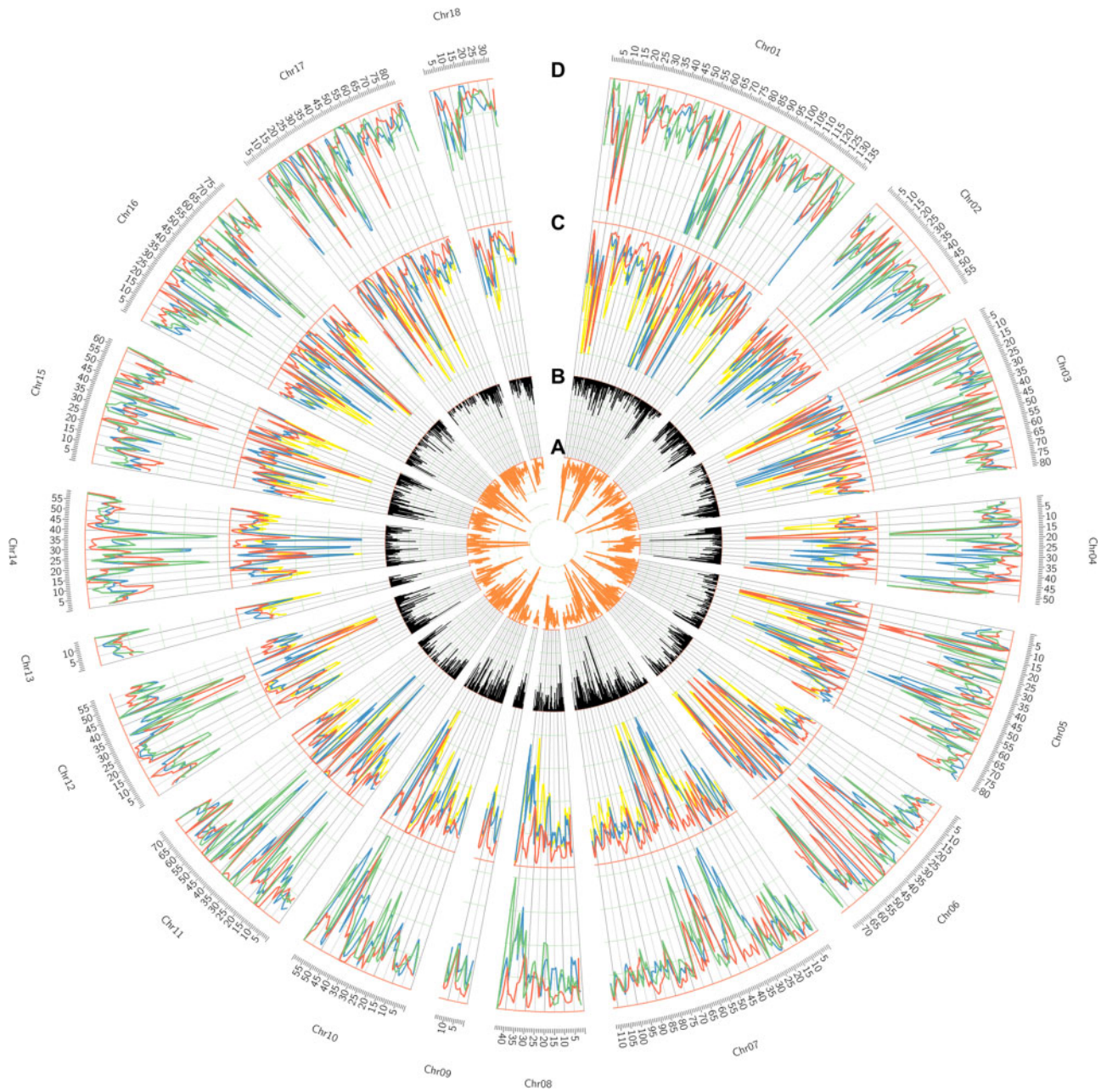
Figure 1. Distribution of SNPs and nucleotide diversity in the quinoa genome. (A) Distribution of SNP frequency, summarized in non-overlapping 1-Mbp windows along the quinoa genome. Lines along the *y*-axis are shown in increments of 0.005. (B) Distribution of SNP number, summarized in non-overlapping 1-Mbp windows. Lines along the *y*-axis are shown in increments of 5. (C) Comparison of nucleotide diversity (π) between highland and lowland lines. Yellow, red, and blue indicate all, lowland, and highland lines, respectively. (D) Comparison of nucleotide diversity (π) between the three populations. Red, green, and blue indicate lowland, northern highland, and southern highland lines, respectively. Mean π was calculated for partially overlapping windows of 2.5 Mbp with a 1-Mbp step. Lines along the *y*-axis are shown in increments of 0.0005.

intermediate position between the *C. hircinum* accessions BYU566 and BYU1101 in the phylogenetic tree (Fig. 3). In addition, line O654 had much higher heterozygosity than the other lines (0.6496 vs 0.0075% for the across-line average; Supplementary Table S2), suggesting that O654, reported in the previous study,[27] may be a hybrid or a hybrid-descendant from a cross with the other species. It is noteworthy that the phylogenetic analysis placed both *C. hircinum* accessions BYU566 and BYU1101 at the base of all quinoa lines analysed in this study (Fig. 3). Neither *C. hircinum* accessions were

included in the three sub-populations revealed by the PCA (Supplementary Fig. S1). These findings suggest that all quinoa lines may have originated from a single domestication event (Fig. 2D).

Next, we evaluated population differentiation using Weir and Cockerham's *F*st value. *F*st between highland and lowland populations was 0.209, compared with 0.111 between northern and southern highland populations (Supplementary Table S6). These numbers indicate that highland and lowland lines are clearly differentiated, corresponding to the highest ΔK ($\Delta K = 15,466$) at $K = 2$ obtained by
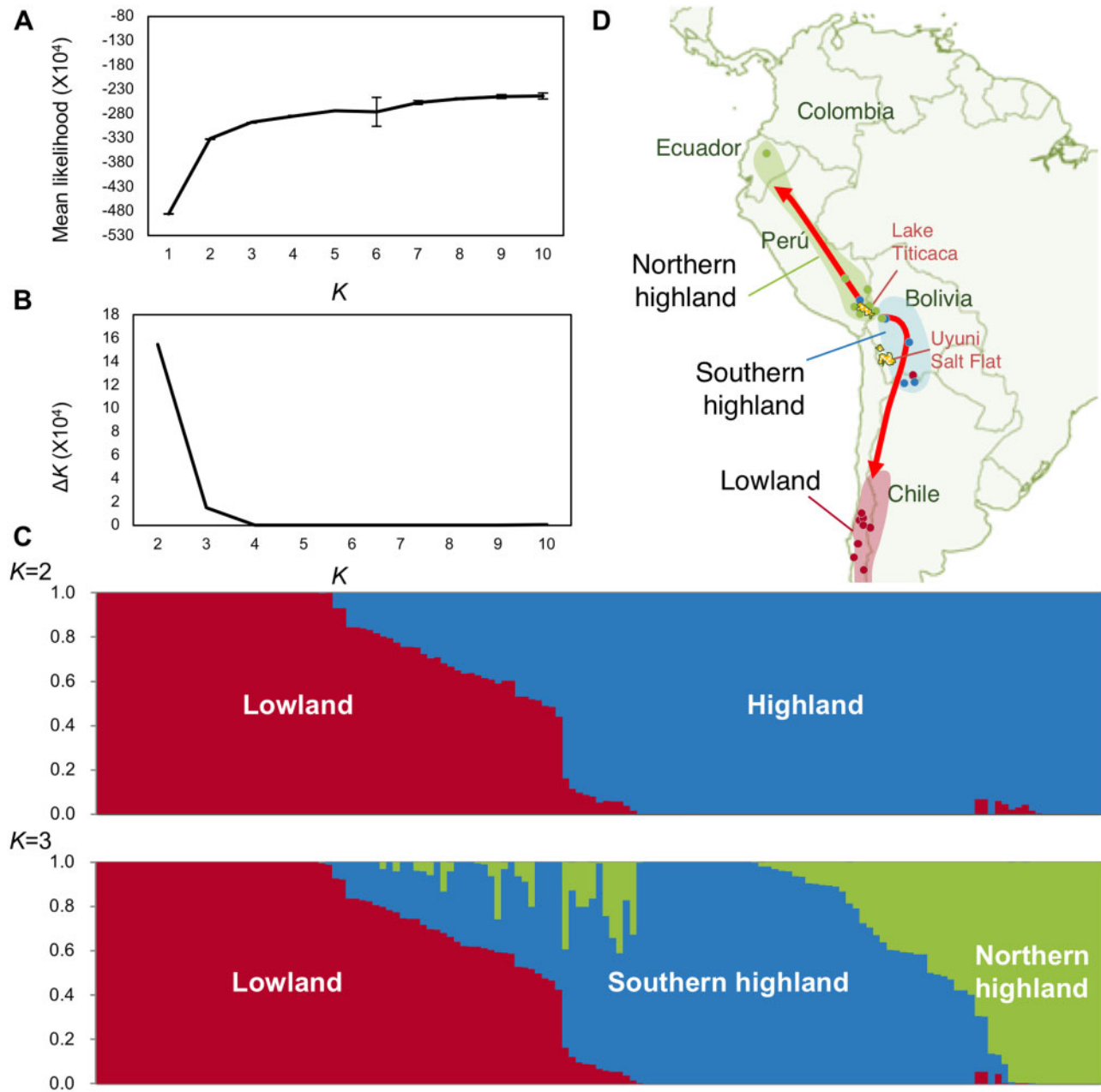
Figure 2. Population structure of 149 quinoa lines. (A) Plots of mean likelihood at each K value. (B) Plots of ΔK at each K value. (C) Model-based ancestry for each quinoa line. Each individual is shown as a vertical bar. The models with $K = 2$ and $K = 3$ show the high ΔK value. At $K = 2$, red and blue indicate lowland and highland lines, respectively. At $K = 3$, red, blue, and green indicate lowland, southern highland, and northern highland lines, respectively. (D) Distribution of quinoa in South America. Estimated areas of northern/southern highland and lowland populations are indicated by green, blue, and red, based on the collection sites (circles) of the accessions with the passport data associated with the inbred lines from this study. Red arrows indicate the estimated evolutionary process of domestication in quinoa (see Section 3 for details).

STRUCTURE. For $K = 3$, Fst between southern highland and lowland populations (0.241) was lower than that between the northern highland and lowland populations (0.291), supporting the view that the lowland population is more closely related to the southern highland population than to the lowland population. We also estimated the genome-wide nucleotide diversity within each population: the nucleotide diversity (π) within the highland population (4.73×10^{-4} for northern highland; 4.86×10^{-4} for southern highland) was higher than across the lowland population (3.90×10^{-4}) (Supplementary Table S7). While previous studies had estimated

higher genetic diversity in the highland quinoa population based on limited sequence comparison[30] and isozyme banding patterns,[45] our genome-wide SNPs data clearly reinforce the notion that South American highland, including Lake Titicaca in the High Andes on the border between Bolivia and Peru, is the centre of diversity from which cultivated quinoa originated.

Together, the results derived from our genetic analyses based on GBS and WGS data support the single domestication hypothesis:[30,45,46] quinoa was most likely first domesticated near Lake Titicaca, which straddles the border between the northern and
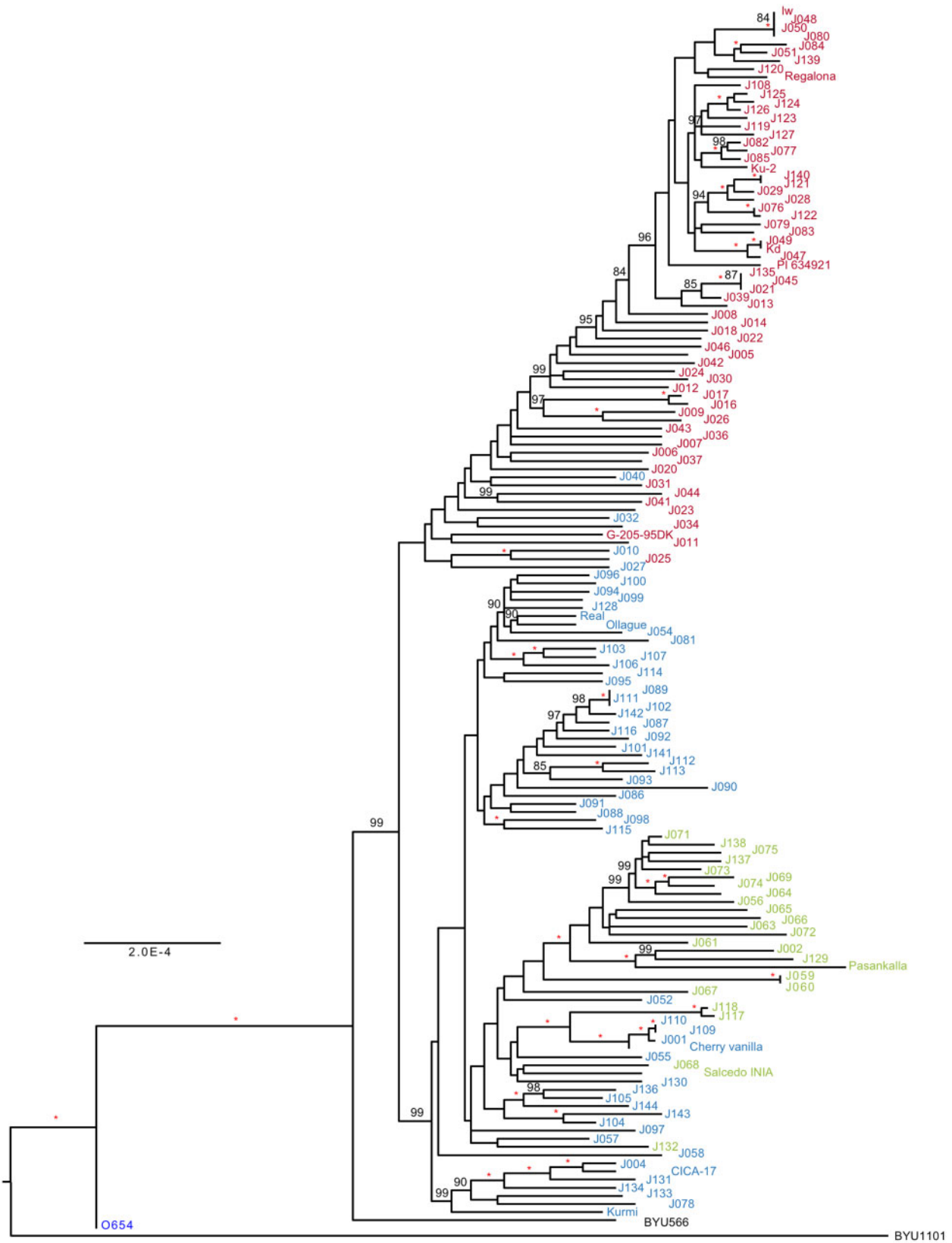
Figure 3. NJ tree of 149 quinoa lines and two *C. hircinum* accessions based on 1.54 Mbp of sequence data corresponding to 7,586 global SNP positions. Red, blue, and green indicate lowland, southern highland, and northern highland lines, respectively. Numbers above branches show bootstrap values based on 100 replicates (not shown if <80%). Red asterisks indicate bootstrap values of 100%. The scale bar corresponds to 2×10^{-4} substitutions per nucleotide site.

southern highland, and then spread to the northern and southern highlands, finally reaching the Chilean lowland region via the southern highland (Fig. 2D). However, a new question arises from our observations: how did the lowland population undergo high genetic differentiation relative to the southern highland population? Our findings may support the possibility that lowland quinoa populations might be derived from crosses between quinoa and wild species, such as *C. hircinum*, in the Chilean lowland region. In rice (*Oryza sativa* L.), two diversified subspecies of cultivated rice (*japonica* and *indica*) arose from a single domestication event (the *japonica* subspecies), and *indica* was later derived from crosses between *japonica* and wild rice species.[47] However, earlier phylogenetic data obtained by WGS indicated that BYU566 clustered with the lowland population, suggesting that quinoa may have independently been domesticated in the highland and lowland.[27] There are differences in both the number of quinoa lines (149 in this study vs 16 in the previous report) and in the methods (GBS + WGS vs WGS) between this study and the previous report,[27] but it remains unclear whether modern-day quinoa originated from a single or two independent domestication events. Therefore, further identification of domestication-related genes such as shattering genes, and/or a genome-wide population analysis using unbiased large sample of quinoa and *C. hircinum* will be required to definitively determine the evolutionary process of domestication in quinoa.

3.4. Phenotypic analyses

3.4.1. Salt tolerance during the germination stage

During the plant life cycle, germination is an early and critical growth stage that determines whether plants will survive and thrive at that location. Here, we therefore focussed on salt tolerance during the germination stage as an initial evaluation of salt tolerance in quinoa lines. In total, we assessed 137 inbred quinoa lines for tolerance to high salt concentrations during germination (Supplementary Tables S8 and S9). We observed that southern highland lines had a much longer mean THRL when compared with lowland and northern highland lines after 24 h under control conditions without added NaCl (Supplementary Fig. S3), indicating that southern highland lines grew more rapidly than all other lines. This may suggest that rapid germination ability is important determinant whether plants can grow in arid regions such as Bolivian Southern Altiplano around Uyuni Salt Flats, where the frequency of rainfall occurrence is very low. However, we failed to detect significant differences in mean THRL between the three population groups after 4 days of growth in the presence of 300 mM NaCl (Supplementary Fig. S3). In contrast, we noticed that most northern highland lines did not germinate after 14 days of treatment with 600 mM NaCl, whereas lowland and southern highland lines did, suggesting that these lines are tolerant to germination under high salinity conditions (Supplementary Fig. S3). Interestingly, we saw no correlation between germination potential at 300 and 600 mM NaCl treatments (R^2=0.01) (Supplementary Fig. S4). These results strongly suggest that germination tests to determine salt tolerance require high salinity conditions (600 mM NaCl) in order to differentiate quinoa lines on the basis of their salt tolerance.

3.4.2. Growth traits

We also assessed 137 inbred quinoa lines for important growth traits such as TGW, PH, SD, LDW, SYP, and DF (Supplementary Tables S8 and S9). The average TGW from the southern highland and lowland lines was higher than for the northern highland lines regardless of the location and season, suggesting that northern highland lines

produced smaller seeds and that TGW is a comparatively conserved trait in quinoa (Supplementary Fig. S3). For the other traits examined in this study, mean values for PH, SD, LDW, SYP, and DF were higher in the lowland lines than in the other lines (Supplementary Fig. S3), consistent with the fact that lowland accessions originated in the Chilean coastal region, which is roughly located on the same latitude as Japan conducted this study.

3.5. Phenotypic diversity

To examine the relationship between population structure and phenotypic diversity, we performed a PCA using data for the phenotypic traits. The individual plotting factors for the first two PCs are listed in Supplementary Table S10. PC1 and PC2 accounted for 34.0 and 28.8% of the total phenotypic variance, respectively, and the following variables showed substantial loadings (loading coefficients ≥0.3): THRL (600 mM NaCl at 14 days) and TGW for PC1; four variables of plant size, namely PH, SD, LDW, and DF, for PC2. The PCA plot exhibited partially overlapping distributions for the three sub-groups (Fig. 4): PC1 separated the northern highland lines from the others, whereas PC2 separated the lines into highland and lowland lines, therefore following altitude as a discriminating factor.

The PCA plot and associated loading coefficients (Supplementary Table S10) clearly suggest that northern highland lines are characterized by lower salt tolerance and smaller grain size, whereas most southern highland and lowland lines are highly salt-tolerant during germination and produce larger grains (Fig. 4). In addition, most lowland lines displayed larger plant size and late flowering phenotypes, whereas southern highland lines had smaller plant size and early flowering phenotypes under our experimental conditions in Japan (Fig. 4). These data therefore demonstrate that quinoa lines should be classified into three, rather than two populations in terms of both genotype and phenotype. Notably, combining the NJ tree and the heatmap illustrating the phenotypes examined here offered a visual confirmation of the separation of quinoa lines and associated phenotypes into the three phylogenetic sub-groups indicated by the PCA (Fig. 5). Furthermore, this figure provides quick access to the genotype–phenotype state of each quinoa line (Fig. 5).

Based on our genotype data, and assuming a single domestication event[45,46] (Fig. 2D), the phenotypic landscape of quinoa lines in terms of THRL in the presence of 600 mM NaCl indicates that cultivated quinoa originated near Lake Titicaca, and then gained high salt tolerance during germination in the saline soils around the Uyuni
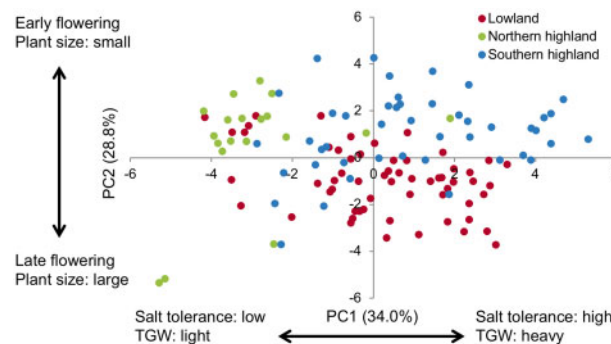

Figure 4. PCA of quinoa lines based on data for the phenotypic traits. Scatterplot of the first two components (PC1 and PC2) from the PCA using 123 quinoa lines with complete phenotypic data. The proportion of variance explained by each component is given in parentheses along each axis. TGW, thousand-grain weight.
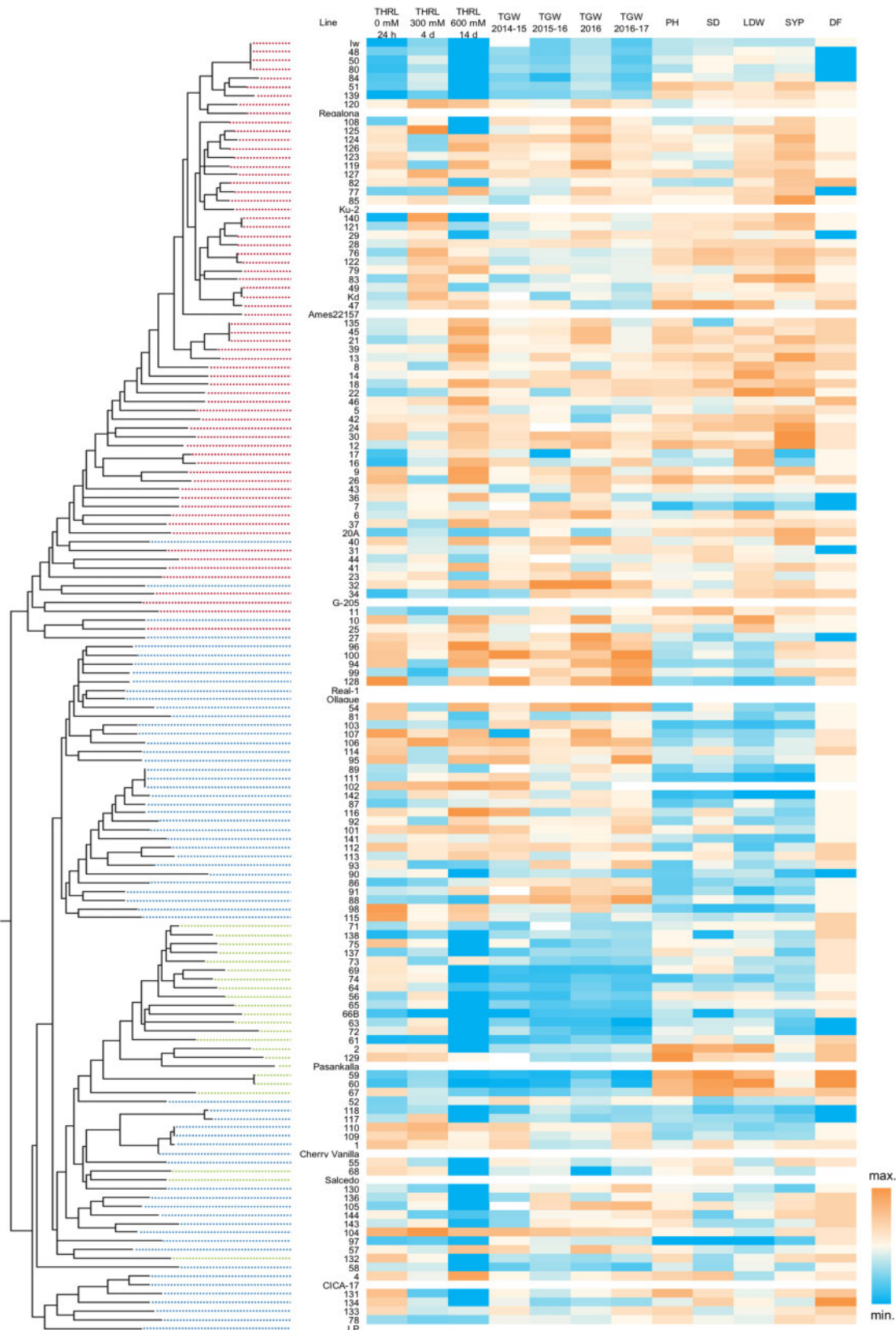
Figure 5. Heatmap of the phenotypic traits in quinoa lines. The columns of the heatmap are ordered based on the NJ tree, which is shown on the left with its branch lengths transformed to be proportional. Missing values in the heatmap are shown in white. Red, green, and blue dotted lines represent lowland, northern highland, and southern highland lines, respectively. THRL, the average total hypocotyl and root length (mm) of seeds treated with 0, 300, or 600 mM of NaCl for each indicated time; TGW, thousand grain weight (g); PH, plant height (cm); SD, stem diameter (mm); LDW, leaf dry weight (g); SYP, seed yield per plant (g); DF, days to flowering (see Supplementary Tables S8 and S9 for details).

Salt Flats in Bolivia, and then spread to lowland regions in Chile, where it retained its high salt tolerance during germination (Fig. 2D). Therefore, the combination of genotype and phenotype data allows us to easily imagine the evolutionary history of quinoa domestication. However, since the sample size for experiments performed under strict and uniform conditions was small due to limitations of our facilities, future analyses with larger sample sizes and fewer lines are needed to confirm the results obtained here and to elucidate the mechanisms underlying salt tolerance and other key traits in quinoa. Overall, the visual data of the heatmap, combined with the phylogenetic tree, aids our understanding of the phenotypic landscape in quinoa for important growth traits (Fig. 5).

3.6. Association mapping of the phenotypic traits

Although association mapping, which is based on the concept of LD, has been used for detecting quantitative trait loci (QTL) in a wide range of organisms, the rate of LD decay is unknown for quinoa. Therefore, before investigating the association between phenotype and genotype by GWAS, we first evaluated LD decay/extension. We calculated the pairwise r^2 correlation value between SNPs with an MAF of at least 0.05. Lowland lines showed a much greater LD decay rate than highland lines (Supplementary Fig. S5). Indeed, the distance between two SNPs around which LD dropped below 0.1 was 1 Mbp for lowland lines and 5.5–8 Mbp in highland populations. Such LD decay distance was greater than in other self-pollinating crops such as rice (*indica*, ca. 65–75kb; *japonica*, ca. 150–200 kb)[48,49] and sorghum (*Sorghum bicolor*, ca. 150 kb),[50] indicating lower resolution in association mapping for identifying QTLs of quinoa.

Despite this slow LD decay rate, we performed an association mapping analysis with 3,156 SNPs and 12 phenotypic traits in our quinoa population. We identified 3,091 and 4 significant MTAs ($P < 0.01$ for GLM, FDR <0.05 for MLM), respectively, during association mapping using the GLM and MLM methods (Supplementary Table S11 and S12). The MLM method detected far fewer MTAs than the GLM method. This discrepancy suggests that the MLM may yield fewer false-positive results than the GLM, because the MLM used the kinship as well as the population structure.[51] The strong population structure and well-differentiated phenotypes among populations might underlie the much smaller number of MTAs identified by MLM. Taking into account the low LD decay rate as well as the strong population structure, these observations suggest that GWAS may not be a suitable technique for identifying QTLs in quinoa.

4. Future perspectives

Quinoa has been recognized as a valued crop with the potential to contribute to global food security due to its exceptional nutritional profile and great adaptability to harsh environments. There is thus an increasing effort to expand quinoa production into non-native areas worldwide.[3,19,22] To date, however, low heat tolerance, stem lodging, negative effects of long photoperiods, and high susceptibility to biotic stresses such as diseases and pests have been major limiting factors in recent quinoa expansion to areas where the pseudocereal is not adapted and for effective large-scale cultivation.[52,53] Conversely, the High Andes is a major production area that provides the world with high-quality quinoa, but they are also considered to be marginal arable land (with yearly rainfall below $200-500$ mm; over 3,700 m above sea level) and are in the front line of climate change. Indeed, the High Andes faces severe problems in quinoa production, mainly caused by extreme weather events induced by climate change, such

as drought and frost.[13] Thus, quinoa breeding is now necessary to yield maximization for potential production and yield optimization for production stability inside and outside its countries of origin.[54] In this light, our current findings provide a useful basis for the molecular elucidation and genetic improvement of quinoa. One hundred and thirty-six newly developed inbred quinoa lines, which consist of three clusters grouping lines into northern highland, southern highland, and lowland sub-populations, will be powerful resources not only for quinoa breeding but also for the elucidation of the molecular mechanisms underlying its great tolerance to abiotic stresses and high nutritional value. Our genotype-informed phenotypic analysis of quinoa under high salt conditions and other important growth traits will greatly contribute to the molecular dissection of traits and breeding in quinoa. We hope that our work will contribute to fill in the gap of the past 500 years in quinoa breeding and studies spanning the 16th to the 20th century, in sharp contrast to other crops that originated from the Andes, including potato, tomato, chili pepper (*Capsicum annuum* L.), cotton (*Gossipium hirsutum* L.), and tobacco (*Nicotiana tabacum* L.), which have all expanded worldwide via the Spaniards since the 16th Century.

Recently, quinoa has been studied as a model plant to understand the molecular mechanisms of abiotic stress tolerance in plants. In the case of salt tolerance, the roles and functions of inorganic ions for osmotic adjustment,[11] tonoplast channels,[55] epidermal bladder cells,[28,56–60] and xylem ion loading[61] have been reported. Recently, the role of peroxisome proliferation in heat and drought stress[62] and heat stress responses in quinoa shoots[63] have been also demonstrated. Apart from these, other molecular analyses concerning evolution of quinoa genome,[64–66] the decision to flower,[67] inflorescence patterning,[68] and abscisic acid (ABA)-mediated seed dormancy and germination[69] have also been reported.

To date, though, the quinoa research community lacks the resources of gain- and loss-of-function technologies to study gene function *in situ*, which constitutes a significant limiting factor for molecular analysis in quinoa. There is therefore an urgent need to develop transformation and genome editing techniques in quinoa to allow the dissection of gene function. In contrast to rice and maize (*Zea mays* L.), which require high fluence rates to keep plants healthy, quinoa plants grow very well in growth chambers normally used for the model plant Arabidopsis (*Arabidopsis thaliana* L.), thereby enabling researchers to shift their focus from the weedy model plant *Arabidopsis* to the new model plant quinoa with great practical value and appropriate inbred lines. Further technical development for the study and breeding of quinoa should facilitate breeding of other staple crops based on the mechanisms of tolerance to abiotic stresses described in quinoa.

Data availability

The raw Illumina reads used in this study are available from DDBJ/EMBL/NCBI under the accession numbers listed in Supplementary Table S1.

Acknowledgements

manuscript. Computations were partially performed on the NIG supercomputer at ROIS National Institute of Genetics.

Accession numbers

DRX089213, DRX089214, DRX089215, DRX089216, DRX089217, DRX089218, DRX089219, DRX089220, DRX089221, DRX089222, DRX089223, DRX089224, DRX089225, DRX089226, DRX089227, DRX089228, DRX089229, DRX089230, DRX089231, DRX089232, DRX089233, DRX089234, DRX089235, DRX089236, DRX089237, DRX089238, DRX089239, DRX089240, DRX089241, DRX089242, DRX089243, DRX089244, DRX089245, DRX089246, DRX089247, DRX089248, DRX089249, DRX089250, DRX089251, DRX089252, DRX089253, DRX089254, DRX089255, DRX089256, DRX089257, DRX089258, DRX089259, DRX089260, DRX089261, DRX089262, DRX089263, DRX089264, DRX089265, DRX089266, DRX089267, DRX089268, DRX089269, DRX089347, DRX089270, DRX089271, DRX089272, DRX089273, DRX089274, DRX089275, DRX089276, DRX089277, DRX089278, DRX089279, DRX089280, DRX089281, DRX089282, DRX089283, DRX089284, DRX089285, DRX089286, DRX089287, DRX089288, DRX089289, DRX089290, DRX089291, DRX089292, DRX089293, DRX089294, DRX089295, DRX089296, DRX089297, DRX089298, DRX089299, DRX089300, DRX089301, DRX089302, DRX089303, DRX089304, DRX089305, DRX089306, DRX089307, DRX089308, DRX089309, DRX089310, DRX089311, DRX089312, DRX089313, DRX089314, DRX089315, DRX089316, DRX089317, DRX089318, DRX089319, DRX089320, DRX089321, DRX089322, DRX089323, DRX089324, DRX089325, DRX089326, DRX089327, DRX089328, DRX089329, DRX089330, DRX089331, DRX089332, DRX089333, DRX089334, DRX089335, DRX089336, DRX089337, DRX089338, DRX089339, DRX089340, DRX089341, DRX089342, DRX089343, DRX089344, DRX089345, DRX089346, DRX089348.

Funding

Conflict of interest

None declared.

Supplementary data

Supplementary data are available at DNARES online.

References

1. Khoury, C.K., Bjorkman, A.D., Dempewolf, H., et al. 2014, Increasing homogeneity in global food supplies and the implications for food security, *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 4001–6.
2. Mayes, S., Massawe, F.J., Alderson, P.G., Roberts, J.A., Azam-Ali, S.N. and Hermann, M. 2012, The potential for underutilized crops to improve security of food production, *J. Exp. Bot.*, **63**, 1075–9.
3. Bazile, D., Jacobsen, S.-E. and Verniau, A. 2016, The global expansion of quinoa: trends and limits, *Front. Plant Sci.*, **7**, 622.
4. Vega-Galvez, A., Miranda, M., Vergara, J., Uribe, E., Puente, L. and Martinez, E.A. 2010, Nutrition facts and functional potential of quinoa (*Chenopodium quinoa* willd.), an ancient Andean grain: a review, *J. Sci. Food Agric.*, **90**, 2541–7.
5. Nowak, V., Du, J. and Charrondiere, U.R. 2016, Assessment of the nutritional composition of quinoa (*Chenopodium quinoa* Willd), *Food Chem.*, **193**, 47–54.
6. Pathan, S., Eivazi, F., Valliyodan, B., Paul, K., Ndunguru, G. and Clark, K. 2019, Nutritional composition of the green leaves of quinoa (*Chenopodium quinoa* Willd), *J. Food Res.*, **8**, 55–65.
7. Rodríguez, J.P., Rahman, H., Thushar, S. and Singh, R.K. 2020, Healthy and resilient cereals and pseudo-cereals for marginal agriculture: molecular advances for improving nutrient bioavailability, *Front. Genet.*, **11**, 49.
8. Jacobsen, S.E. and Mujica, A. 2003, Quinoa: an alternative crop for saline soils, *J. Exp. Bot.*, **54**, i25.
9. Jacobsen, S.E., Monteros, C., Christiansen, J.L., Bravo, L.A., Corcuera, L.J. and Mujica, A. 2005, Plant responses of quinoa (*Chenopodium quinoa* Willd.) to frost at various phenological stages, *Eur. J. Agron.*, **22**, 131–9.
10. Razzaghi, F., Ahmadi, S.H., Adolf, V.I., Jensen, C.R., Jacobsen, S.E. and Andersen, M.N. 2011, Water relations and transpiration of quinoa (*Chenopodium quinoa* Willd.) under salinity and soil drying, *J. Agron. Crop Sci.*, **197**, 348–60.
11. Hariadi, Y., Marandon, K., Tian, Y., Jacobsen, S.E. and Shabala, S. 2011, Ionic and osmotic relations in quinoa (*Chenopodium quinoa* Willd.) plants grown at various salinity levels, *J. Exp. Bot.*, **62**, 185–93.
12. Yasui, Y., Hirakawa, H., Oikawa, T., et al. 2016, Draft genome sequence of an inbred line of *Chenopodium quinoa*, an allotetraploid crop with great environmental adaptability and outstanding nutritional properties, *DNA Res.*, **23**, 535–46.
13. Bonifacio, A. 2019, Improvement of quinoa (*Chenopodium quinoa* Willd.) and Qañawa (*Chenopodium pallidicaule* Aellen) in the context of climate change in the high Andes, *Cien. Inv. Agric.*, **46**, 113–24.
14. Dillehay, T.D., Rossen, J., Andres, T.C. and Williams, D.E. 2007, Preceramic adoption of peanut, squash, and cotton in northern Peru, *Science*, **316**, 1890–3.
15. González, J.A., Eisa, S.S.S., Hussin, S.A.E.S. and Prado, F.E. 2015, Quinoa: an Incan crop to face global changes in agriculture, In: Murphy, K., Matanguiban, J., eds. *Quinoa: Improvement and Sustainable Production*, pp. 1–18. Wiley Blackwell: New Jersey.
16. Gomez-Pando, L. 2015, Quinoa breeding, In: Murphy, K., Matanguiban, J. eds. *Quinoa: Improvement and Sustainable Production*, pp. 87–107. Wiley Blackwell: New Jersey.
17. National Academy of Sciences. 1975, *Underexploited Tropical Plants with Promising Economic Value*. National Academy of Sciences: Washington, USA.
18. Schlick, G. and Bubenheim, D.L. 1993, Quinoa: an emerging "new" crop with potential for CELLS, *NASA Tech. Pap.*, **3422**, 1–9.
19. Bazile, D., Pulvento, C., Verniau, A. et al. 2016, Worldwide evaluations of quinoa: preliminary results from post International Year of Quinoa FAO projects in nine countries, *Front. Plant Sci.*, **7**, 850.
20. Morton, M.J.L., Awlia, M., Al-Tamimi, N., Saade, S., Pailles, Y., Negrão, S. and Tester, M. 2019, Salt stress under the scalpel - dissecting the genetics of salt tolerance, *Plant J.*, **97**, 148–63.
21. Adolf, V.I., Jacobsen, S.-E. and Shabala, S. 2013, Salt tolerance mechanisms in quinoa (*Chenopodium quinoa*), *Environ. Exp. Bot.*, **92**, 43–54.
22. Choukr-Allah, R., Rao, N., Hirich, A., et al. 2016, Quinoa for marginal environments: toward future food and nutritional security in MENA and central Asian regions, *Front. Plant Sci.*, **7**, 346.
23. Palomino, G., Hernández, L.T., de la, C. and Torres, E. 2008, Nuclear genome size and chromosome analysis in *Chenopodium quinoa* and *C. berlandieri* subsp., *Nuttalliae, Euphytica*, **164**, 221–30.
24. Yangquanwei, Z., Neethirajan, S. and Karunakaran, C. 2013, Cytogenetic analysis of quinoa chromosomes using nanoscale imaging and spectroscopy techniques, *Nanoscale Res. Lett.*, **8**, 463.
25. Walsh, B.M., Adhikary, D., Maughan, P.J., Emshwiller, E. and Jellen, E.N. 2015, Chenopodium polyploidy inferences from Salt Overly Sensitive 1 (SOS1) data, *Am. J. Bot.*, **102**, 533–43.

26. Kolano, B., McCann, J., Orzechowska, M., Siwinska, D., Temsch, E. and Weiss-Schneeweiss, H. 2016, Molecular and cytogenetic evidence for an allotetraploid origin of *Chenopodium quinoa* and *C. berlandieri* (Amaranthaceae), *Mol. Phylogenet. Evol.*, **100**, 109–23.

27. Jarvis, D.E., Ho, Y.S., Lightfoot, D.J., et al. 2017, The genome of *Chenopodium quinoa*, *Nature*, **542**, 307–12.

28. Zou, C., Chen, A., Xiao, L., et al. 2017, A high-quality genome assembly of quinoa provides insights into the molecular basis of salt bladder-based salinity tolerance and the exceptional nutritional value, *Cell Res.*, **27**, 1327–40.

29. Maughan, P.J., Bonifacio, A., Jellen, E.N., et al. 2004, A genetic linkage map of quinoa (*Chenopodium quinoa*) based on AFLP, RAPD, and SSR markers, *Theor. Appl. Genet.*, **109**, 1188–95.

30. Christensen, S.A., Pratt, D.B., Pratt, C., et al. 2007, Assessment of genetic diversity in the USDA and CIP-FAO international nursery collections of quinoa (*Chenopodium quinoa* Willd.) using microsatellite markers, *Plant Genet. Res.*, **5**, 82–95.

31. Tamura, A., Kato, T., Taki, A., et al. 2013, Preventive and curative effects of apple latent spherical virus vectors harboring part of the target virus genome against potyvirus and cucumovirus infections, *Virology*, **446**, 314–24.

32. Yabe, S., Hara, T., Ueno, M., et al. 2014, Rapid genotyping with DNA micro-arrays for high-density linkage mapping and QTL mapping in common buckwheat (*Fagopyrum esculentum* Moench), *Breed. Sci.*, **64**, 291–9.

33. Elshire, R.J., Glaubitz, J.C., Sun, Q., et al. 2011, A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species, *PLoS One*, **6**, e19379.

34. Bolger, A.M., Lohse, M. and Usadel, B. 2014, Trimmomatic: a flexible trimmer for Illumina sequence data, *Bioinformatics*, **30**, 2114–20.

35. Li, H. and Durbin, R. 2009, Fast and accurate short read alignment with Burrows–Wheeler transform, *Bioinformatics*, **25**, 1754–60.

36. Li, H., Handsaker, B., Wysoker, A., et al.; 1000 Genome Project Data Processing Subgroup. 2009, The Sequence Alignment/Map format and SAMtools, *Bioinformatics*, **25**, 2078–9.

37. DePristo, M.A., Banks, E., Poplin, R.E., et al. 2011, A framework for variation discovery and genotyping using next-generation DNA sequencing data, *Nat. Genet.*, **43**, 491–8.

38. Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y. and Buckler, E.S. 2007, TASSEL: software for association mapping of complex traits in diverse samples, *Bioinformatics*, **23**, 2633–5.

39. Danecek, P., Auton, A., Abecasis, G., et al.; 1000 Genomes Project Analysis Group. 2011, The variant call format and VCFtools, *Bioinformatics*, **27**, 2156–8.

40. Pritchard, J.K., Stephens, M. and Donnelly, P. 2000, Inference of population structure using multilocus genotype data, *Genetics*, **155**, 945–59.

41. Evanno, G., Regnaut, S. and Goudet, J. 2005, Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study, *Mol. Ecol.*, **14**, 2611–20.

42. Krzywinski, M., Schein, J., Birol, I., et al. 2009, Circos: an information aesthetic for comparative genomics, *Genome Res.*, **19**, 1639–45.

43. Zhang, T., Gu, M., Liu, Y., et al. 2017, Development of novel InDel markers and genetic diversity in *Chenopodium quinoa* through whole-genome re-sequencing, *BMC Genomics*, **18**, 685.

44. Maughan, P.J., Smith, S.M., Rojas-Beltrán, J.A., et al. 2012, Single nucleotide polymorphism identification, characterization, and linkage mapping in quinoa, *Plant Genome*, **5**, 114–25.

45. Wilson, H.D. 1988, Quinoa biosystematics I: domesticated populations, *Econ. Bot.*, **42**, 461–77.

46. Gandarillas, H. 1979, Genética y orígen, In: Tapia, M. E., ed. *Quinua y Kañiwa: Cultivos Andinos*, pp. 45–64. Instituto Interamericano de Ciencias Agrícolas: Bogota.

47. Huang, X., Kurata, N., Wei, X. et al. 2012, A map of rice genome variation reveals the origin of cultivated rice, *Nature*, **490**, 497–501.

48. Xu, X., Liu, X., Ge, S. et al. 2012, Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes, *Nat. Biotechnol.*, **30**, 105–11.

49. Mather, K.A., Caicedo, A.L., Polato, N.R., Olsen, K.M., McCouch, S. and Purugganan, M.D. 2007, The extent of linkage disequilibrium in rice (*Oryza sativa* L.), *Genetics*, **177**, 2223–32.

50. Morris, G.P., Ramu, P., Deshpande, S.P., et al. 2013, Population genomic and genome-wide association studies of agroclimatic traits in sorghum, *Proc. Nat. Acad. Sci. U.S.A.*, **110**, 453–8.

51. Yu, J., Pressoir, G., Briggs, W.H., et al. 2006, A unified mixed-model method for association mapping that accounts for multiple levels of relatedness, *Nat. Genet.*, **38**, 203–8.

52. Schmöckel, S.M., Lightfoot, D.J., Razali, R., Tester, M. and Jarvis, D.E. 2017, Identification of putative transmembrane proteins involved in salinity tolerance in *Chenopodium quinoa* by integrating physiological data, RNAseq, and SNP analyses, *Front. Plant Sci.*, **8**, 1023.

53. Gomez-Pando, L.R., Aguilar-Castellanos, E., Ibañez-Tremolada, M., et al. 2019, Quinoa (*Chenopodium quinoa* Willd.) breeding, In: Al-Khayri, J.M., Jain, S.M. and Johnson, D.V. eds. *Advances in Plant Breeding Strategies: Cereals*, *vol.* 5, pp. 259–313. Springer International Publishing, New York.

54. Tilman, D., Cassman, K.G., Matson, P.A., Naylor, R. and Polasky, S. 2002, Agricultural sustainability and intensive production practices, *Nature*, **418**, 671–7.

55. Bonales-Alatorre, E., Shabala, S., Chen, Z.H. and Pottosin, I. 2013, Reduced tonoplast fast-activating and slow-activating channel activity is essential for conferring salinity tolerance in a facultative halophyte, quinoa, *Plant Physiol.*, **162**, 940–52.

56. Shabala, S., Bose, J. and Hedrich, R. 2014, Salt bladders: do they matter? *Trends Plant Sci.*, **19**, 687–91.

57. Böhm, J., Messerer, M., Müller, H.M., et al. 2018, Understanding the molecular basis of salt sequestration in epidermal bladder cells of *Chenopodium quinoa*, *Curr Biol.*, **28**, 3075–85.e7.

58. Kiani-Pouya, A., Roessner, U., Jayasinghe, N.S., et al. 2017, Epidermal bladder cells confer salinity stress tolerance in the halophyte quinoa and *Atriplex* species, *Plant Cell Environ.*, **40**, 1900–15.

59. Kiani-Pouya, A., Rasouli, F., Bazihizina, N., Zhang, H., Hedrich, R. and Shabala, S. 2019, A large-scale screening of quinoa accessions reveals an important role of epidermal bladder cells and stomatal patterning in salinity tolerance, *Environ. Exp. Bot.*, **168**, 103885.

60. Kiani-Pouya, A., Rasouli, F., Shabala, L., Tahir, A.T., Zhou, M. and Shabala, S. 2020, Understanding the role of root-related traits in salinity tolerance of quinoa accessions with contrasting epidermal bladder cell patterning, *Planta*, **251**, 103.

61. Zarei, M., Shabala, S., Zeng, F., et al. 2020, Comparing kinetics of xylem ion loading and its regulation in halophytes and glycophytes, *Plant Cell Physiol.*, **61**, 403–15.

62. Hinojosa, L., Sanad, M.N.M., Jarvis, D.E., Steel, P., Murphy, K. and Smertenko, A. 2019, Impact of heat and drought stress on peroxisome proliferation in quinoa, *Plant J.*, **99**, 1144–58.

63. Tovar, J.C., Quillatupa, C., Callen, S.T., et al. 2020, Heating quinoa shoots results in yield loss by inhibiting fruit production and delaying maturity, *Plant J.*, **102**, 1058–73.

64. Maughan, P.J., Chaney, L., Lightfoot, D.J., Cox, B.J., Tester, M., Jellen, E.N. and Jarvis, D.E. 2019, Mitochondrial and chloroplast genomes provide insights into the evolutionary origins of quinoa, *Sci. Rep.*, **9**, 185.

65. Winkel, T., Aguirre, M.G., Arizio, C.M., et al. 2018, Discontinuities in quinoa biodiversity in the dry Andes: an 18-century perspective based on allelic genotyping, *PLoS One*, **13**, e0207519.

66. Heitkam, T., Weber, B., Walter, I., Liedtke, S., Ost, C. and Schmidt, T. 2020, Satellite DNA landscapes after allotetraploidization of quinoa (*Chenopodium quinoa*) reveal unique A and B subgenomes, *Plant J.*, **103**, 32–52.

67. Golicz, A.A., Steinfort, U., Arya, H., Singh, M.B. and Bhalla, P.L. 2020, Analysis of the quinoa genome reveals conservation and divergence of the flowering pathways, *Funct. Integr. Genomics*, **20**, 245–58.

68. Wu, Q., Bai, X., Zhao, W., et al. 2019, Investigation into the underlying regulatory mechanisms shaping inflorescence architecture in *Chenopodium quinoa*, *BMC Genomics*, **20**, 658.

69. Wu, Q., Bai, X., Wu, X., et al. 2020, Transcriptome profiling identifies transcription factors and key homologs involved in seed dormancy and germination regulation of *Chenopodium quinoa*, *Plant Physiol. Biochem.*, **151**, 443–56.