Chromosome-level genome assemblies for highland quinoa cultivated in the Andean highlands

Quinoa is emerging as a key seed crop for global food security due to its ability to grow in marginal environments and its excellent nutritional properties. Our comprehensive genomic analyses revealed that the quinoa inbred lines can be categorized into three genetic sub-populations: the northern highland, southern highland, and lowland groups (2020 Research Highlights: Genetic and phenotypic variation of agronomic traits and salt tolerance among quinoa inbred lines). However, unlike lowland quinoa lines, no useful genomic information is available for highland quinoa lines. The high-quality genome assemblies obtained in this study will provide the basis for advancing functional genomics in quinoa to facilitate the development of climate-adapted highland quinoa breeding materials and contribute to a better understanding of the domestication process of quinoa, including its adaptation to harsh environments and its origin.

The whole genome sequences of J075 as a representative line of northern highland quinoa, and J100 as a representative line of southern highland quinoa were decoded using PacBio high-fidelity (HiFi) sequencing. Furthermore, we obtained their chromosome-level genome assemblies by integrating the HiFi sequencing reads and aligning them using the linkage map generated by dpMig-Seq. The assembled genomes of J075 and J100, each comprising 18 pseudochromosomes, exhibited sizes of 1.29 Gb and 1.32 Gb, with contig N50 values of 66.3 Mb and 12.6 Mb and scaffold N50 values of 71.2 Mb and 70.6 Mb, respectively, and were predicted to harbor 65,303 and 64,945 genes. As an example of comparative analysis of genome structure and diversity among quinoa genotypes, we selected genes homologous to the reported betalain biosynthetic gene sequences from each line and assessed genomic structural variation. Regions on chromosomes containing clusters of betalain synthesis genes *CqDODA1* and *CqCYP76AD1* were less conserved in genome structure compared to adjacent regions in each line, suggesting their association with the phenotypic differences between the dark-red highland and the light-colored lowland quinoa.

Detailed comparative analysis of genomes among representative quinoa genotypes is expected to facilitate the selection of crop design materials for developing quinoa varieties with enhanced stress tolerance, yield stability, and high nutritional value, which are critical for food security. These genomic data are publicly available via the Kazusa DNA Research Institute’s "Plant GARDEN" portal.

Authors: Kobayashi, Y., Fujita, Y., Nagatoshi, Y., Fujii, K. [JIRCAS],
Hirakawa, H., Shirasawa, K. [Kazusa DNA Res. Inst.], Nishimura, K. [Okayama Univ.],
Oros, R. [PROINPA], Almanza, G.R. [Universidad Mayor de San Andres], Yasui, Y. [Kyoto Univ.]

Table 1. Genome assembly statistics of Northern and Southern highland quinoa lines

	Northern highland J075	Southern highland J100
Genome size (Gb)	1.29	1.32
Number of Genes	65,303	64,945
N50 (Mb) *	71.2	70.6
BUSCO (%)[†]	99.2	99.1
LAI score[‡]	17.40	17.75

*The N50 value is a statistical measure that reflects the contiguity and overall quality of an assembly. The higher the N50, the longer the average contiguous sequences in the assembly.

[†] The BUSCO measures genome assembly completeness by checking for the presence of essential, single-copy genes from the embryophyta dataset.

[‡] The LAI measures how effectively a genome assembly reconstructs LTR retrotransposons—higher scores reflect more complete and continuous assemblies of these repetitive regions.

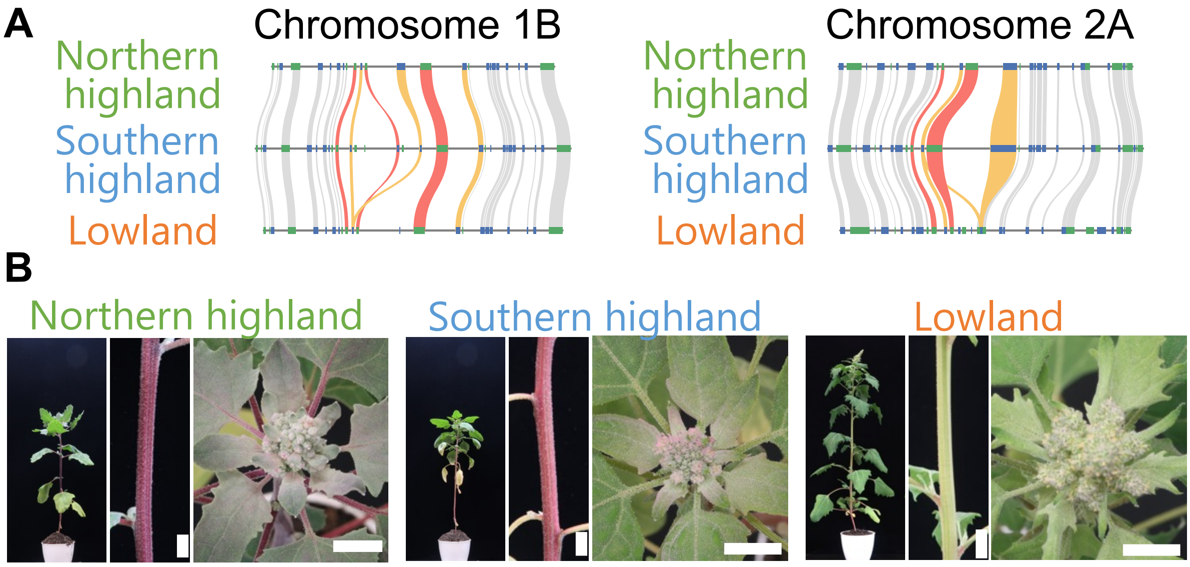

Fig. 1. Genomic structure of betalain biosynthesis gene clusters and plant color phenotypes in representative quinoa lines

(A) Structural conservation of the regions encompassing the betacyanin biosynthesis gene cluster on chromosomes 1B and 2A. The betalain biosynthesis genes, *CqDODA1* and *CqCYP76AD1*, are connected by red and orange lines, respectively, across quinoa genotypes.

(B) Phenotypes of stem and shoot apex colors in northern and southern highland and lowland inbred lines after 56 days of growth. Bars indicate 1 cm.

Reference: Kobayashi et al. (2024) *Frontiers in Plant Science* 15: 1434388. © The Author(s) 2024
The table and figure are modified from Kobayashi et al. (2024).