

Aus dem Institut für Pflanzenbau und Pflanzenzüchtung
der Christian-Albrechts-Universität zu Kiel

**Quantitative Trait Loci (QTL) mapping and genome-wide
transcriptome analysis to identify agronomically important genes
in quinoa (*Chenopodium quinoa*)**

Dissertation
zur Erlangung des Doktorgrades
der Agrar- und Ernährungswissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Nathaly Fernanda Maldonado Taipe

aus Quito, Ecuador

Kiel, 2022

Dekan: Prof. Dr. Karl H. Mühling

1. Berichterstatter: Prof. Dr. Christian Jung

2. Berichterstatter: Prof. Dr. Daguang Cai

Tag der mündlichen Prüfung: 29.06.2022

Table of Contents

1. General introduction	1
1.1. Quinoa botany, diversity, and importance	1
1.2. Worldwide quinoa cultivation and breeding	2
1.3. Flowering time regulation and photoperiod response in the model plant <i>Arabidopsis thaliana</i>	2
1.4. Flowering time regulation and photoperiod response in quinoa and related species ..	4
1.5. QTL mapping as a tool for unraveling the genetic basis of agronomically important traits.....	5
1.6. Next-generation sequencing enables the generation of high-quality genetic maps and identification of QTL with high precision.....	6
1.7. Whole transcriptome sequencing as a tool to unravel flowering time regulation in crops.....	7
1.8. Increasing genetic variation for quinoa breeding	9
1.9. Hypotheses, objectives, and scientific questions	10
2. High-density mapping of QTL controlling agronomically important traits in quinoa (<i>Chenopodium quinoa</i> Willd.).....	11
2.1. Abstract.....	11
2.2. Introduction	11
2.3. Methods	12
2.3.1. Plant material and growth conditions.....	12
2.3.2. Phenotypic evaluation	12
2.3.4. Whole-genome sequencing and bioinformatics	13
2.3.5. Linkage map construction.....	14
2.3.6. QTL mapping and pleiotropy analysis	14
2.3.7. Epistasis analysis	14
2.3.8. Candidate gene identification and haplotype analyses	15
2.3.9. Heritability estimates and statistical analysis	15
2.4. Results	15
2.4.1. Segregation and phenotypic analysis of F ₂ and F ₃ populations	15
2.4.2. Sequencing the F ₂ population revealed millions of SNPs	17
2.4.3. Construction of a high-density linkage map	19
2.4.4. QTL mapping, pleiotropic loci identification, and epistasis calculation	19
2.4.5. Identification of putative candidate genes controlling agronomically important traits.....	26
2.5. Discussion.....	27

2.6.	Acknowledgements	30
2.7.	References	30
2.8.	Supplementary Data	34
3.	Leaf and shoot apical meristem transcriptomes of quinoa (<i>Chenopodium quinoa</i>) in response to photoperiod and plant development	37
3.1.	Abstract.....	37
3.2.	Introduction	37
3.3.	Materials and Methods	38
3.4.	Results	41
3.4.1.	The transition of the shoot apical meristem into a floral meristem.....	41
3.4.2.	Transcriptome sequencing.....	41
3.4.3.	Identification of differentially expressed genes regulating flowering time	42
3.4.4.	Functional annotation and classification of the DEGs	46
3.4.5.	Co-expression analysis of the DEGs	47
3.4.6.	Validation of RNA-seq pipeline by RT-qPCR.....	49
3.5.	Discussion.....	51
3.6.	Acknowledgments	54
3.7.	References	54
3.8.	Supplementary data	60
4.	<i>Agrobacterium</i> -mediated transformation experiments	63
4.1.	Introduction	63
4.2.	Materials and methods.....	63
4.2.1.	Plant material.....	63
4.2.2.	Plant transformation and <i>in vitro</i> cultivation.....	64
4.2.4.	DNA isolation and PCR	66
4.3.	Results	66
4.3.1.	Regeneration of whole plants from quinoa callus	66
4.3.2.	<i>Agrobacterium</i> -mediated transformation	70
4.4.	Discussion.....	73
4.5.	References	74
4.6.	Supplementary Data	76
5.	Closing discussion	77
5.1.	Two pleiotropic QTL were found in two different environments	78
5.2.	Histologic and transcriptomic changes towards elucidating quinoa flowering regulation.....	78
5.3.	Quinoa appears recalcitrant to <i>Agrobacterium</i> -mediated transformation	80

5.4. QTL mapping and transcriptome analysis provided breeding perspectives for quinoa	81
6. Summary.....	83
7. Zusammenfassung	84
8. Appendix.....	85
8.1. Supplementary Tables	85
8.2. Supplementary Figures	102
9. References.....	130
10. Supplementary data on CD/DVD	138
11. Curriculum Vitae and Publications.....	139
11.1. Curriculum vitae	139
11.2. Publications	139
12. Publications and declaration of own contribution	141
13. Acknowledgements.....	144

List of Abbreviations

°C	Celsius
%	Percent
AFLPs	Amplified fragment length polymorphisms
<i>AGL15</i>	<i>AGAMOUS-LIKE 15</i>
<i>AP1</i>	<i>APETALA1</i>
<i>AP2-1</i>	<i>APETALA2-1</i>
<i>B. Vulgaris</i>	<i>Beta vulgaris</i>
BAP	6-Benzylaminopurine
<i>BBM</i>	<i>BABY BOOM</i>
<i>BBX19</i>	<i>DOUBLE B-BOX TYPE ZINC FINGER 19</i>
bHLH	basic helix–loop–helix
<i>BIC1</i>	<i>BLUE-LIGHT INHIBITOR OF CRYPTOCHROMES 1</i>
BLAST	Basic Local Alignment Search Tool
<i>BTC1</i>	<i>BOLTING TIME CONTROL 1</i>
<i>Bv</i>	<i>Beta vulgaris</i>
<i>C. quinoa</i>	<i>Chenopodium quinoa</i>
<i>C. rubrum</i>	<i>Chenopodium rubrum</i>
ca.	Circa
cm	Centimeters
cM	centimorgans
CO	<i>CONSTANS</i>
COL	<i>CO</i> -like genes
<i>Cq</i>	<i>Chenopodium quinoa</i>
CRISPR	Clustered regularly interspaced short palindromic repeats
<i>CRY1</i>	<i>CRYPTOCHROME 1</i>
<i>CRY2</i>	<i>CRYPTOCHROME 2</i>
das	Days after sowing
DEG	Differentially expressed gene
DNA	Deoxyribonucleic acid
DTF	Days to flowering
DTM	Days to maturity
ECV	Environmental coefficient of variation
EF	Early-flowering phenotype
ELF3	<i>EARLY FLOWERING 3</i>
ELF5	<i>EARLY FLOWERING 5</i>
ERF113	<i>ETHYLENE-RESPONSIVE TRANSCRIPTION FACTOR 113</i>
ET	Eigentrail
FAO	Food and Agriculture Organization
FC	Fold change
<i>FIP1</i>	<i>FRIGIDA INTERACTING PROTEIN</i>
<i>FLC</i>	<i>FLOWERING LOCUS C</i>
<i>FLD</i>	<i>FLOWERING LOCUS D</i>
<i>FLM</i>	<i>FLOWERING LOCUS M</i>
<i>FRI</i>	<i>FRIGIDA</i>

<i>FRL</i>	<i>FRI</i> -like genes
<i>FT</i>	<i>FLOWERING LOCUS T</i>
<i>FTIP3</i>	<i>FT INTERACTING PROTEIN 3</i>
<i>FTL</i>	<i>FT</i> -like genes
<i>FUL</i>	<i>FRUITFULL</i>
<i>FWA</i>	<i>FLOWERING WAGENINGEN</i>
g	grams
GA	Genetic advance
GA	Gibberellic acids
Gb	Giga basepair
GBS	Genotyping by sequencing
GCV	Genotypic coefficients of variation
<i>GI</i>	<i>GIGANTEA</i>
GO	Gene Ontology
GWAS	Genome-Wide Association Study
h	hours
h ²	Heritability
<i>HD3</i>	<i>HEADING DATE 3</i>
<i>HY5</i>	<i>ELONGATED HYPOCOTYL 5</i>
IBA	Indole-3-butyric acid
InDel	Insertion and deletions
ISSRs	Inter-simple sequence repeats
kb	Kilobase pair
KEGG	Kyoto Encyclopedia of Genes and Genomes
l	Liters
LD	Long day
LF	Late-flowering
<i>LFY</i>	<i>LEAFY</i>
LG	Linkage group
<i>LHY</i>	<i>LATE ELONGATED HYPOCOTYL</i>
lncRNA	Long non-coding RNA
LOD	Logarithm of odds
m	Meters
maf	Minor allele frequency
MAS	Marker-assisted selection
Mb	Mega basepair
<i>MET1</i>	<i>METHYLTRANSFERASE 1</i>
mg	milligrams
min	minutes
miRNA	micro RNA
MM	Module membership
mMS	modified Murashige and Skoog
mRNA	messenger RNA
MS	Mildew susceptibility
NB	Night-break

NGS	Next-generation sequencing
PCR	Polymerase chain reaction
PCV	Phenotypic coefficients of variation
PD	Panicle density
PE	Pair end
PEBP	Phosphatidyl ethanolamine-binding proteins
PH	Plant height
<i>PHYA</i>	<i>PHYTOCHROME A</i>
<i>PHYB</i>	<i>PHYTOCHROME B</i>
PL	Panicle length
PPT	Phospinotricin
QTL	Quantitative trait locus
RAD	Restriction site-associated genomic DNA
RAPD	Random amplified polymorphic DNA
<i>RFT1</i>	<i>RICE FLOWERING LOCUS T 1</i>
RIL	Recombinant inbred line
RIN	RNA integrity number
RNA	Ribonucleic acid
RNA-seq	RNA-sequencing
RNPs	Ribonucleoproteins
ROS	Reactive oxygen species
RRS	Reduced-representation sequencing
Rt-qPCR	Real-time quantitative PCR
<i>RUG3</i>	<i>RCC1 DOMAIN-CONTAINING PROTEIN 3</i>
s	seconds
SAAT	Sonication-Assisted- <i>Agrobacterium</i> -mediated gene Transfer system
SAM	Shoot apical meristem
SC	Saponin content
SD	Short day
<i>SEP3</i>	<i>SEPALLATA</i>
Skim-seq	Skim sequencing
SN	Seed number per plant
SNPs	Single-nucleotide polymorphisms
<i>SOC1</i>	<i>SUPPRESSOR OF CONSTANS OVEREXPRESSION 1</i>
<i>SPL</i>	<i>SQUAMOSA PROMOTER BINDING PROTEIN-LIKE</i>
SSRs	Simple sequence repeats
<i>SUF4</i>	<i>SUPPRESSOR OF FRI 4</i>
SVD	Singular value decomposition
SW	Seed weight per plant
<i>TK11</i>	<i>TSL-KINASE INTERACTING PROTEIN 1</i>
TKW	Thousand kernel weight
<i>TPS9</i>	<i>TREHALOSE-PHOSPHATASE/SYNTHASE 9</i>
<i>TSAR2</i>	<i>TRITERPENE SAPONIN BIOSYNTHESIS ACTIVATING REGULATOR 2</i>
<i>TSF</i>	<i>TWIN SISTER of FT</i>
<i>TSL</i>	<i>TOUSLED</i>

VCF	Variant calling format
VE	Environmental variance
VG	Genotypic variance
VIGS	Virus-induced gene silencing
VP	Phenotypic variance
W	Week
WES	Whole-exome sequencing
WGCNA	Weighted Gene Co-expression Network Analysis
WGRS	Whole-genome resequencing
<i>WRKY13</i>	<i>WRKY TRANSCRIPTION FACTOR 13</i>
<i>WUS2</i>	<i>WUSCHEL 2</i>
ZT	Zeitgeber time
<i>ZTL</i>	<i>ZEITLUPE</i>
μmol	micromoles

1. General introduction

1.1. Quinoa botany, diversity, and importance

Quinoa (*Chenopodium quinoa* Willd.) is a pseudocereal native to the Andean region of South America. Along with potato and maize, quinoa is one of the three major pre-colonial domesticated crops. This crop is still one of the major staple foods in the Andean territory and, among other uses, its seeds are used to make flour, soup, cereal, and alcohol (Peterson & Murphy, 2015). Besides, quinoa is grown for animal consumption and medicinal purposes. Traditional quinoa processing and products evolved and are being adopted into novel food processing and modern food products, such as puffed quinoa, noodles, and ready-to-eat products (Angeli et al., 2020).

Botanically, quinoa is a dicotyledonous annual species from the family *Amaranthaceae*, which includes other economically important species such as sugar beet (*Beta vulgaris*). Quinoa's stem can be straight or with branches with colors changing from white, yellow, or light brown to red. It can grow up to 3 m and its leaves can be rhomboidal or triangular with entire, dentate, or serrate margins (Jacobson et al., 2021). Quinoa inflorescence is a panicle and has both hermaphrodite flowers and female flowers. This crop is a predominantly self-pollinated species with varying rates of cross-pollination (10–17%) and some cultivars are male sterile (partially or in all the flowers) (Ward, 1998).

C. quinoa shows a broad genetic variation, which is reflected in its extensive distribution and adaptation to biotic and abiotic stresses. In South America, quinoa grows from 20 °N in Colombia to 40 °S in Chile and Argentina, and from sea level to over 4,000 m altitude (Garcia, Condori, & Del Castillo, 2015). Quinoa exhibits resistance to insects and diseases, as well as tolerance to frost, drought, and salinity (Vita et al., 2021). Hence, because of its great adaptability to extreme climatic and soil conditions, quinoa can be grown on marginal lands not currently suitable for other major crops (Alandia, Rodriguez, Jacobsen, Bazile, & Condori, 2020).

The physicochemical, nutritional, and functional properties of quinoa seeds make them ideal for human consumption. This crop has been called a pseudocereal because, unlike cereals, its seeds have high protein content. Lysine and eight of the other essential amino acids are present in quinoa seeds in balanced amounts (Melini & Melini, 2021). Therefore, this crop is considered a "functional food" that contributes to human nutrition and lowers the risk of several diseases (Ali, 2019), and presents a great potential to improve world food security (Alandia et al., 2020).

Quinoa is an allotetraploid species ($2n = 4x = 36$), with a genome size of 1.45-1.5 Gb, that resulted from a spontaneous hybridization event between two diploid species approximately 3.3 to 6.3 million years ago (Jarvis et al., 2017). In terms of genetic variability, this crop is classified as an oligocentric species that has a wide center of origin in the Andean region and has undergone multiple diversification routes (Bhargava & Srivastava, 2013). Genotypes have been selected by ancient people according to their particular ecological environments and used over time. This process resulted in five main ecotypes at a continental scale, each with its diversity subcenter. In addition, there remain several wild species and relatives in South America (Bazile, Fuentes, & Mujica, 2013).

Germplasm collections provide a unique genetic resource for breeding programs. The quinoa genebank network covers 30 countries and holds 16,422 accessions of the quinoa species and its wild relatives (*C. quinoa*, *Chenopodium album*, *Chenopodium berlandieri*, *Chenopodium hircinum*, *Chenopodium petiolare*, and *Chenopodium murale*). Genebanks in the Andean region conserve over 88% of the quinoa and its wild relatives' accessions (Rojas et al., 2015). The largest seed bank is the Bolivian National Collection at the Fundación para la

Promoción e Investigación de Productos Andinos (PROINPA) comprises 4,312 quinoa accessions preserved under *ex situ* conditions (FAO, 2010).

1.2. Worldwide quinoa cultivation and breeding

Quinoa production, trade, and consumption across the world have grown rapidly in recent decades because of high consumer demand for healthy, nutritious, and gluten-free food products. Consequently, 161,415 tons of quinoa were produced worldwide in 2019 as compared to 32,435 tons in 1961; the harvested area increased from 52,555 ha in 1961 to 184,585 ha in 2019 (Ahmadzai, 2020; FAO, 2021). Quinoa cultivation has transcended continental boundaries, being now present in Europe, Africa, and Asia. The number of quinoa-producing countries increased from eight countries in 1980 to 95 countries in 2015. Moreover, because of population growth and growing demand for nutritious, healthy, and gluten-free foods, quinoa production, as well as quinoa production facilities, will keep on increasing in the near future (Bazile, Jacobsen, & Verniau, 2016; FAO, 2021).

Despite the large history of this crop and the growing worldwide interest, quinoa breeding is still in its infancy. In quinoa breeding, the main aim is to develop high-yielding varieties with high protein and low saponin content (Alandia et al., 2020; Bhargava, Shukla, & Ohri, 2006). From an agronomic perspective, this crop is being improved for short, non-branching plants with compact panicles and increased tolerance for abiotic and biotic stresses. In particular, downy mildew (*Peronospora variabilis*) can cause 90 percent yield reductions in susceptible cultivars (Danielsen & Munk, 2004). Importantly, one of the primary goals of quinoa breeding is to change the seasonal timing of reproduction to create unique varieties that are more suited to local conditions. Because this species is primarily a short-day plant, it must be adapted to long-day conditions when grown in temperate climates and high latitudes such as Europe, North America, and China (Patirange et al., 2020).

1.3. Flowering time regulation and photoperiod response in the model plant *Arabidopsis thaliana*

Yield potential is mainly determined by the flowering processes, which constitute the first step toward seed formation. Hence, improved understanding and subsequent modification of the flowering processes, including flowering time, is essential to aid in the development of breeding programs (Golicz, Steinfort, Arya, Singh, & Bhalla, 2020). Unfortunately, features of photoperiod response and flowering time genes, as well as their associated pathways, remain elusive in quinoa. However, these genes and pathways have been comprehensively studied in the model plant *Arabidopsis thaliana*. Hence, I made use of the knowledge established for *Arabidopsis* to present a brief overview of the flowering time regulation mechanisms (Figure 1).

A complex network of genes engaged in many regulatory mechanisms, including photoperiod, vernalization, circadian clock, gibberellin, autonomous, and plant aging, influence the flowering process (Mouradov, Cremer, & Coupland, 2002). These pathways converge to regulate floral integrator genes, which combine the outputs and start floral transition under favorable conditions.

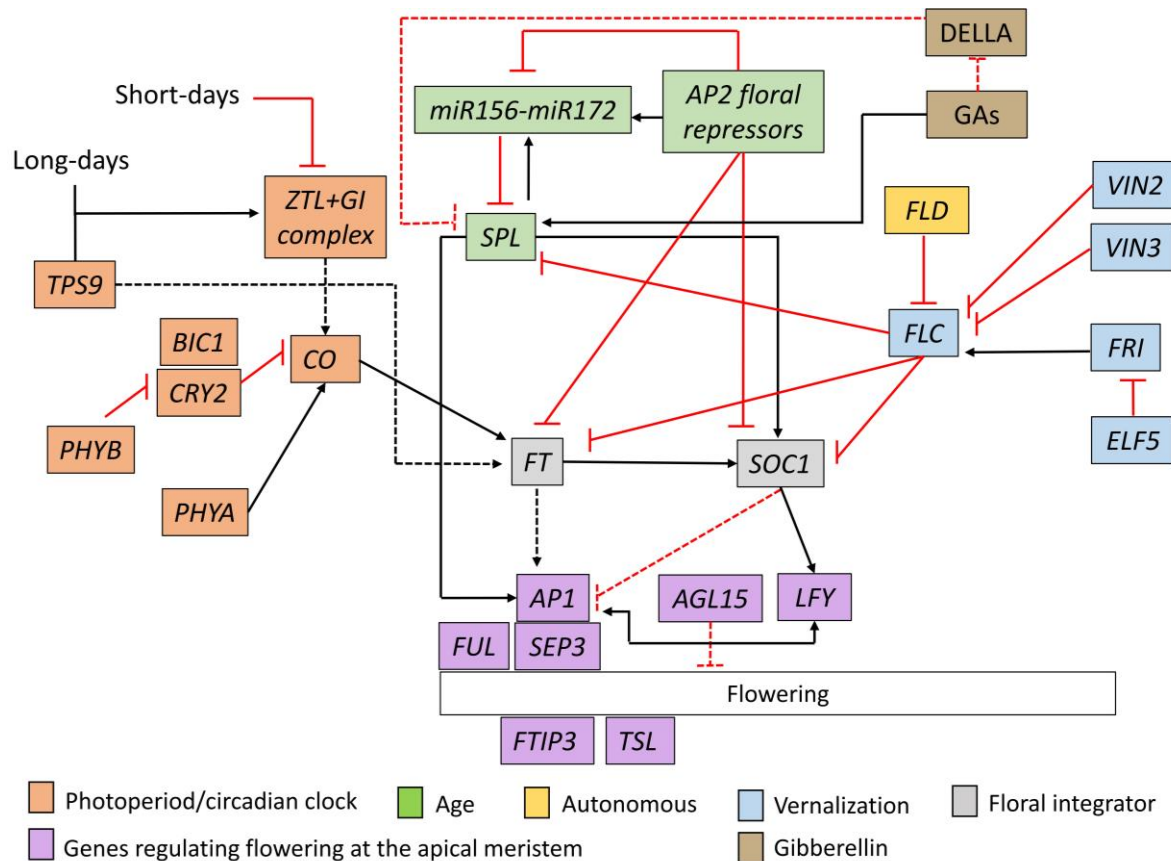


Figure 1. Simplified scheme of flowering time regulation in *A. thaliana*. Genes involved in photoperiod (orange boxes), age (green boxes), autonomous (yellow boxes), gibberellin (brown), and vernalization (light blue) pathways are shown. Grey boxes represent the main floral integrators. Genes acting at the meristem are stated in purple. Solid and dotted lines show direct or indirect regulation, respectively. Black arrows and red T-ends show positive or negative regulation, respectively. Modified from Leijten et al. (2018).

In *A. thaliana*, the circadian clock genes sense the photoperiod and coordinate the expression of *CONSTANS* (*CO*), a zinc finger transcription factor. A *CO*-dependent transcriptional activation of *FLOWERING LOCUS T* (*FT*) takes place in the leaf tissue (Valverde et al., 2004). *CO* transcription depends on the activity of a protein complex formed by *GIGANTEA* (*GI*) and *ZEITLUPE* (*ZTL*). *ZTL* is attached to a chromophore and perceives light signals (Andrés & Coupland, 2012). Phytochromes, on the other hand, act upstream of these genes and play fundamental roles in photo perception of the environmental light (red/far-red light-absorbing phytochromes) (Wickland, Battu, Hudson, Diers, & Hudson, 2017). Following, *BLUE-LIGHT INHIBITOR OF CRYPTOCHROMES 1* (*BIC1*) acts as an inhibitor of plant cryptochromes by binding to *CRYPTOCHROME 2* (*CRY2*) (Wang et al., 2016). Furthermore, a very recent representative of flowering promoters responding to long-day conditions is *TREHALOSE-PHOSPHATASE/SYNTHASE 9* (*TPS9*). This gene probably accelerates the floral transition by activating *FT* expression (Tian et al., 2021).

The long-distance florigenic signal is conveyed by moving FT from leaves via vascular tissues to the shoot apical meristem. FT, together with *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*), activates the meristem identity genes *LEAFY* (*LFY*), *APETALA1* (*AP1*), *SEPALLATA3* (*SEP3*), and *FRUITFULL* (*FUL*), indicating that transition of the vegetative meristem into a reproductive meristem has started (Blümel, Dally, & Jung, 2015). In the meristem tissue, another group of genes plays a role. For instance, *AGAMOUS-LIKE 15* (*AGL15*), acts as a flowering repressor (Adamczyk, Lehti-Shiu, & Fernandez, 2007), while *FT INTERACTING PROTEIN 3* (*FTIP3*) is required for the maintenance of inflorescence SAMs (Liu et al., 2018). Besides, *TOUSLED* (*TSL*) function is required in the floral meristem for the correct initiation of floral organ primordia (Roe, Rivin, Sessions, Feldmann, & Zambryski, 1993).

In the vernalization pathway, *FLOWERING LOCUS C (FLC)* is one of the major flowering repressor genes in *Arabidopsis* (Michaels & Amasino, 1999). *FRIGIDA (FRI)* is a major upstream element of *FLC*, and together, they contribute to a large proportion of flowering time variation in winter-annual *Arabidopsis* (Shindo et al., 2005). Another gene playing a role in this pathway is *EARLY FLOWERING 5 (ELF5)*, which acts as an *FRI* repressor causing early flowering (Noh, Bizzell, Noh, Schomburg, & Amasino, 2004). To continue, the autonomous pathway controls flowering time in parallel to the vernalization pathway. Although multiple genes have been identified as being involved in this pathway, their role in flowering converges to regulations of *FLC* (Wu, Fang, Zhu, & Dean, 2020). For instance, *FLOWERING LOCUS D (FLD)* acts in this pathway by repressing *FLC*; thus, promoting flowering. Besides the autonomous and vernalization pathways, the age pathway mechanism also affects floral transition. In this pathway, miR156 and miR172 are important factors. With age, miR156 abundance decreases, while miR172 abundance increases. There is an inverse relationship between *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL)* expression patterns and miR156, which shows that *SPL* is regulated by miR156/172 (Zhou et al., 2013). As the plants get older, the synchronized regulation of miRNA156, miRNA172, and *SPLs* promote flowering (Luo, Guo, & Li, 2013).

A group of hormones (gibberellic acids) also promote *Arabidopsis* flowering. There are over 130 described gibberellic acids (GA), of which GA1, GA3, GA4, GA5, and GA6 are associated with flowering (Yamaguchi, 2008). For instance, it is known that artificial GA3 treatment restores the *FT* expression in mutants, showing delayed flowering because of the suppression of *FT* (Galvão, Horrer, Küttner, & Schmid, 2012). Additionally, DELLA proteins are an important part of the gibberellin pathway since they are involved in GA-mediated *FT* regulation. These proteins can bind to *CO* directly to suppress its expression and are degraded by a complex containing GA (Murase, Hirano, Sun, & Hakoshima, 2008). DELLA proteins are involved in regulating the floral meristem identity genes indirectly through *SPL* (Bao, Hua, Shen, & Yu, 2020).

1.4. Flowering time regulation and photoperiod response in quinoa and related species

Quinoa is primarily a short-day plant, which makes it difficult to grow in temperate and high-latitude environments (Patirange et al., 2020). Therefore, a better knowledge of photoperiod response and flowering time can help quinoa adapt to various conditions by adjusting the seasonal timing of reproduction. However, up to date, little is known about the regulation of these processes in quinoa.

Regarding quinoa's response to photoperiod, it is known that cultivars originating from the tropics present major sensitivity to photoperiod and longer vegetative phase (short-day behavior), while cultivars from the Altiplano of Peru and Bolivia and sea level of southern Chile are the least sensitive to photoperiod and have the shortest vegetative phase (day-neutral behavior) (Risi, Galwey, Wickens, Haq, & Day, 1989). A recent study by Patirange et al. (2020) showed that 276 different quinoa accessions depict substantial differences in photoperiod sensitivity, where short-day, long-day, and day-neutral behavior could be found. However, most quinoa accessions (67.39%) depicted short-day behavior and only 7.6% were defined as long-day accessions.

The genetics behind *C. quinoa* photoperiod response and flowering time have been an object of only a few studies. Relevant conclusions from those studies include that quinoa *FLOWERING LOCUS T (FT)* and *TWIN SISTER* homologs are involved in flowering regulation under short-day conditions (SD). These conclusions were made based on *in silico* analysis of targets of long non-coding RNA (lncRNA) after night break (NB) under SD and expression analyses under SD (Patirange et al., 2021; Wu et al., 2021). An NB under SD causes down-regulation of *CONSTANS-like (COL)*, *LATE ELONGATED HYPOCOTYL*

(*LHY*), *EARLY FLOWERING 3 (ELF3)*, and *ELONGATED HYPOCOTYL 5 (HY5)* homologs and up-regulation of *PHYTOCROME A (PHYA)* and *CRYPTOCHROME1 (CRY1)* homologs (Wu et al., 2021). Besides, *COL* homologs were associated with flowering time variation under long days (LD) (Patiranage et al., 2021).

Flowering time regulation has been studied in other members of the *Chenopodium* genus. *Chenopodium rubrum*, a short-day species, has three homologs of *FT*, *CrFTL1*, *CrFTL2*, and *CrFTL3*. From these, only *CrFTL1* was required for flowering to begin, operating as a single floral activator (Cháb, Kolář, Olson, & Štorchová, 2008; Drabešová et al., 2016; Drabešová, Cháb, Kolář, Haškovcová, & Štorchová, 2014). Furthermore, *Chenopodium ficifolium* and *Chenopodium suecicum*, two near relatives of the diploid progenitors of quinoa, have recently been examined for *FT* homologs. Under SD, in *C. suecicum* and *C. ficifolium*, expression analysis revealed a link between *FTL1* and floral induction (Štorchová et al., 2019). Finally, in a segregating F₂ population of *C. ficifolium*, marker-trait associations with flowering time, plant height, and the number of branches were identified for an ortholog of *FTL1* (Subedi, Neff, & Davis, 2021).

Flowering timing in *Beta vulgaris*, a quinoa-related *Amaranthaceae* species, has been analyzed extensively. To flower, *B. vulgaris*, unlike quinoa, requires vernalization followed by exposure to LD. Mutasa-Göttgens et al. (2010) found that, as a direct response to LD, plants containing the dominant *B allele* behave as annuals. Contrarily, homozygous recessive plants at the *B locus* act like biennials and require vernalization to bolt and flower. Two *Arabidopsis* homologous of *FLOWERING LOCUS T (FT)* drive the flowering process in *B. vulgaris*. *BvFT2* promotes flowering while *BvFT1* inhibits it. *BvFT1* downregulation is critical for the vernalization response (Pin et al., 2010). In addition, the *BOLTING TIME CONTROL 1 (BTC1)* and the *DOUBLE B-BOX TYPE ZINC FINGER 19 (BvBBX19)* genes also take part in the flowering regulation pathway of sugar beet. *BvBTC1*'s function is to promote early bolting by repressing *BvFT1* (Pin et al., 2012). Complementarily, *BvBBX19* may act together with *BvBTC1* similarly to *CONSTANS (CO)* in *A. thaliana* (Dally, Eckel, Batschauer, Höft, & Jung, 2018).

1.5. QTL mapping as a tool for unraveling the genetic basis of agronomically important traits

QTL is a statistical method, which attempts to explain genetic variation in complex traits by linking phenotypic and genotypic data (Myles & Wayne, 2008). Therefore, it is a useful tool to unravel the genetic mechanisms of important agronomical traits, such as flowering time in crops. The basis of QTL mapping lies in finding significant differences between the groups concerning the trait of interest. For this purpose, molecular markers are used to partition the mapping population into different genotypic groups. When the phenotypic means of the genotypic groups differ significantly, it indicates that the marker is linked to the QTL (Collard, Jahufer, Brouwer, & Pang, 2005).

The most used method for detecting QTL is interval mapping. Within this method, simple interval mapping considers intervals between adjacent pairs of linked markers along chromosomes to examine for QTL, while composite interval mapping combines interval mapping with linear regression on markers located outside the interval under consideration, as covariates. By using composite interval mapping, the residual variation is reduced and the power and precision of QTL detection are improved (Lander & Botstein, 1989; Zeng, 1994).

QTL mapping considers both marker genotypes and phenotypic data from a segregating population. QTL analysis in plants is generally based on populations that segregate for different traits, such as F₂ or backcross populations. These populations are relatively uncomplicated to build for self-pollinating species (difficulties when crossing parental lines may arise) and are considered the simplest type of mapping population (Collard et al., 2005).

However, the use of F_2 populations is limited because they cannot be assessed as replicates given that each plant is genetically unique (Singh & Singh, 2015). In this scenario, the use of phenotypic data from succeeding generations, such as F_3 , can considerably improve the accuracy of QTL detection. With an increased number of individuals studied in the F_3 generation, the environmental variance is reduced. Hence, by combining F_3 phenotypic data with F_2 genotypic data, QTL can be detected with greater precision (Zhu, Huang, & Zhang, 2007).

In quinoa, few mapping studies are available up to date. The betalain color locus was mapped by Jarvis et al. (2008). For this purpose, a genetic map constructed with 216 simple sequence repeats (SSR) comprising 38 linkage groups (LGs) and covering 913 cM was used. Later, Cervantes and van Loo (2017) studied 94 individuals from an F_2 population from a cross between ‘Carina Red’ (bitter) and ‘Atlas’ (non-bitter). QTL for color, flowering time, and yield-related traits were reported. Based on a linkage map constructed with 1,076 SNPs, Cervantes and van Loo (2017) located two major QTL, one for days to flower bud appearance on chromosome Cq6B, and another for seed characters on chromosome Cq2B. They detected 135 QTL, most of them for color traits. QTL were detected for each of the spectral bands of each color of seeds and leaves by measuring the color wavelengths of different samples. Moreover, another linkage map constructed with 14,178 SNPs (KASPar genotyping) from two RIL populations is available. This map comprises 29 linkage groups spanning 1,404 cM (Maughan et al., 2012). The most recent quinoa linkage map by Jarvis et al. (2008) combined the map from Maughan et al. (2012) with two new linkage maps. The resulting map contains 6,403 markers on 18 linkage groups spanning 2,034 centimorgans (cM).

1.6. Next-generation sequencing enables the generation of high-quality genetic maps and identification of QTL with high precision

The relative abundance and codominant nature of SSRs have caused them to be extensively used across a wide range of crops for genetic mapping and diversity studies over the last few decades. Other markers used for the same purpose were random amplified polymorphic DNA (RAPD), inter-simple sequence repeats (ISSRs), and amplified fragment length polymorphisms (AFLPs). These markers, however, provide a relatively low coverage of the genome; require large amounts of gel electrophoresis, are time-consuming, and have a high cost (Huang et al., 2009). With next-generation sequencing (NGS) and single-nucleotide polymorphisms (SNPs) markers, these problems have been addressed to a large extent (Kumar et al., 2021).

Among NGS approaches, reduced-representation sequencing (RRS) signifies low-cost and effective genotyping. The RRS approaches use restriction digestion, ligation of adaptors, and/or pooling of DNA samples as previous steps to sequencing. Later, an adaptor containing barcode sequences can be used to isolate sequenced reads from different individuals using bioinformatics tools. Examples of RRS are genotyping by sequencing (GBS) and restriction site-associated genomic DNA (RAD). RAD sequencing is a method to sequence DNA next to restriction sites (Baird et al., 2008). Like RAD sequencing, GBS relies on restriction enzymes to generate a reduced representation of the genome for sequencing. However, the GBS library preparation protocol involves more than one restriction enzyme and requires less DNA (Kyriakidou, Tai, Anglin, Ellis, & Strömvik, 2018). Different from RRS techniques, skim sequencing (Skim-seq) avoids the complexity reduction step. In Skim-seq, the entire genome is sequenced with low depth. Besides, Skim-seq involves genotyping of large populations to address the limitations related to missing data (Kale et al., 2015). These large populations are sequenced in depths from ~0.01x to ~8.0x and population-based algorithms, which often need information from the parental lines, are used for the imputation of missing data. In terms of costs, Skim-seq is less expensive than standard GBS; but the bioinformatic analysis required is more thorough (Table 1). This approach has a wide range of downstream applications and

has been used in several crops (Table 2). For example, Tong et al. (2021) sequenced the whole genome of 381 pedigree populations at 1.8x depth. By this approach, they identified 43 QTL for six agronomical traits in tobacco.

Table 1. Comparison between standard genotyping-by-sequencing and Skim-sequencing. Based on Kumar et al. (2021) and Kyriakidou et al. (2018).

	Genotyping by sequencing	Skim-sequencing
Cost per sample	High	Low
SNP discovery rate	High	Low to moderate
Restriction enzyme	Required	Not required
Analysis of complexity	Moderate	High
Reference genome availability	Not essential	Essential
Genome coverage	More coverage of genome with high resolution	Low genome coverage with high resolution on a whole-population basis
Imputation	Optional	Critical
Applications	De novo SNPs discovery and genetic mapping	

1.7. Whole transcriptome sequencing as a tool to unravel flowering time regulation in crops

A transcriptome is a collection of all messenger RNA (mRNA) molecules expressed by an organism. Transcriptome assemblies provide a powerful tool for high-throughput research that enables efficient screening and selection of differentially expressed genes (DEGs) in a specific experimental setting (Byrne, Cole, Volden, & Vollmers, 2019; Koboldt, Steinberg, Larson, Wilson, & Mardis, 2013). In non-model plants, such as quinoa, combining transcriptome sequencing and differential expression analysis can provide a sensitive and efficient analysis of gene expression changes in complex processes such as flowering time regulation, where many genes interplay (Golicz et al., 2020).

A current limitation of transcriptome sequencing is that the read length of the sequencers is too short to capture entire transcripts from end to end. Hence, incomplete fragments of transcripts must be computationally assembled into full-length isoforms. For this purpose, there are two strategies: first, performing *de novo* assembly using powerful algorithms (e.g. Trinity, rnaSPAdes) (Bankevich et al., 2012; Grabherr et al., 2011); second, genome-guided transcriptome assemblies (e.g. Cufflinks, StringTie) (Pertea et al., 2015; Trapnell et al., 2010). The advantage of *de novo* assembly relies on resolving alternative transcript features (e.g. splicing variants). However, *de novo* assembly algorithms have key drawbacks. For instance, when a transcript has two alternative splice sites too far apart (~1000 bp), no algorithm will connect those two events. Although computational methods that take this into account have been developed, they still cannot detect high complex isoform mixtures. On the other hand, genome-guided assembly is more accurate in terms of annotated genes assembly. Conversely, not all gene isoforms are detected by this method (Kanitz et al., 2015).

Table 2. Investigations showing diverse Skim sequencing applications in a wide range of crops using varied plant materials.

Application	Crop	Plant material	Number of variants or bases	Coverage ¹	Reference
Genome assembly	Finger millet	PR202	~1,366 million bases	0.05-0.91X	Hatakeyama et al. (2018)
Evolutionary studies	Wheat	300 accessions	~234 million SNPs	3.80X	Zhou et al. (2020)
Variant discovery in mutants	Cotton	14 CRISPR edited plants	4,188 to 6,404 SNPs and 312 to 745 InDels	~3.00X	Li et al. (2019)
	Banana	Gamma irradiated Novaria cultivar	11 newly-discovered InDels	2.20-3.47X	Datta, Jankowicz-Cieslak, Nielen, Ingelbrecht, and Till (2018)
Genome-Wide Association Study (GWAS)	Rice	203 rice varieties	~2,3 million SNPs	1.50X	Wang et al. (2016)
	Tomato	246 accessions	~2 million SNPs, InDel, inversions, and translocations	> 6.00X	Voichkek and Weigel (2020)
	Maiz	150 inbred lines	~35 million SNPs, InDel, inversions and translocations	> 6.00X	Voichkek and Weigel (2020)
QTL mapping	Rice	RILs <i>O. nivara</i> x <i>Indica</i>	181,957 SNPs	~2.83X	Ma et al. (2016)
	Peanut	RILs Middleton x Sutherland	67,000 SNPs	0.5X	Korani et al. (2021)
	Tabacum	Backcrossing BC4F3 population	24,142 SNPs	~1.00X	Tong et al. (2021)
	Maize	RILs X178 x HuangC	20,000 SNPs	0.03X	Wang et al. (2018)
	Chickpea	RILs ICC 4958 x ICC 1882	53,223 SNPs	~0.72X	Kale et al. (2015)
	Soybean	Six F ₂ populations	4,408 to 6,704 SNPs	~1.00X	Fang et al. (2019)
	Pepper	120 F ₂ individuals	171,199 SNPs	8.00X	Wei et al. (2020)

RIL: recombinant inbred line.

¹Average number of reads per individual, accession, variety or line, multiplied by read length and divided by genome size.

In recent years, transcriptome sequencing and differential expression analysis have aided to unravel the genetic mechanism behind the flowering time regulation of several crops. For instance, a recent study in cotton generated transcriptomes for early and late flowering varieties at different developmental stages. After differential expression and co-expression analyses, 46 candidate genes regulating the cotton floral bud differentiation were identified. From these, functional analysis of one candidate gene led to the construction of a new flowering time regulatory pathway in cotton (Cheng et al., 2021). A further example is given for rapeseed. In this crop, leaf transcriptome analysis under vernalized and non-vernalized conditions provided novel evidence of neofunctionalization of flowering time regulators (Shah et al., 2018). On the other hand, the transcriptome profile of a rice transgenic line helped to unravel circadian clock genes delaying rice flowering time in an expression level-dependent manner. Altogether, these studies show the utility of transcriptome analysis to unravel flowering mechanisms in crops (Liu et al., 2018).

By publishing the reference genome of quinoa (Jarvis et al., 2017), transcriptional studies became possible. As a result, several researchers have already used this technology for a variety of studies. A recent experiment examined salt stress responses in the roots of two quinoa cultivars (QQ056 and 37TES) by transcriptome analysis (Shi & Gu, 2020). Besides, Wu et al. (2019) studied the transcriptome of the LL-1 cultivar to understand how inflorescences develop. Moreover, researchers recently produced transcriptome data as part of a study to understand the inhibition of quinoa seed germination by exogenous abscisic acid (Wu et al., 2020).

1.8. Increasing genetic variation for quinoa breeding

Genetic analysis tools such as QTL mapping and transcriptome sequencing offer a first insight into the genetic mechanisms of a targeted biological process. However, scaling up to investigate the function of major genes regulating these processes is essential. For this aim, crop transformation is the ultimate tool to provide evidence of the function of such genes. Apart from the utility of crop transformation as a tool for functional analysis, it can also be used for the direct introduction of useful genes into important crops; thus, supplying new and genetically diverse plant materials to breeding programs (Keshavareddy et al., 2018).

The introduction of novel genes into the nuclear genomes of diverse plant species can be accomplished using several methods that have been proven stable. Among these methods, delivery systems are the most employed. At present, there are two classes of delivery systems: non-biological systems and biological systems (Darbani, Farajnia, Toorchi, Zakerbostanabad, & Noeparvar, 2008). Non-biological systems encompass, for instance, particle bombardment, sonication, and electroporation. Meanwhile, the biological systems include transformation by a viral vector and *Agrobacterium*-mediated transformation. *Agrobacterium*-mediated transformation is preferred for several crops since it reports higher transformation efficiencies than non-biological systems and signifies a stable transformation (Keshavareddy et al., 2018). *Agrobacterium*-mediated transformation can be carried out by cocultivation with explant tissue; thus, requiring a well-established tissue culture-based regeneration system. Another possibility is *in planta* transformation, which delivers the foreign genetic material through *Agrobacterium* directly to the target plant tissue, abolishing the need for tissue culture. Examples of *in planta* transformation include vacuum infiltration, the transformation of germinating seeds, and floral dipping (Clough & Bent, 1998; Feldmann & David Marks, 1987).

In quinoa, there are several reports on successful *in vitro* regeneration of shoots from different quinoa explants, where direct as well as indirect organogenesis have been studied. Investigated tissues for *in vitro* culture include cotyledon and hypocotyl, as well as callus produced from roots (Hesami & Daneshvar, 2016; Hesami, Naderi, & Yoosefzadeh-Najafabadi, 2018; Telahigue & Toumi, 2017). Dark conditions were claimed to be appropriate

for callus regeneration from hypocotyl and different concentrations of Murashige and Skoog (MS) medium in different concentrations of plant growth regulators have been tested (Hesami & Daneshvar, 2016). Besides, Hesami et al. (2018) reported a shooting rate of 93.33% (percentage of regenerated shoots from the total number of explants) through direct organogenesis using cotyledonary node explants MS medium supplemented with 2.0 mg/l 6-benzyl amino purine (BAP).

Despite the several reports on quinoa *in vitro* regeneration, there is only one report about quinoa transformation. Komari (1990) tested four different *A. tumefaciens* strains to infect leaf disks, callus, and cell suspension cultures. The DNA isolated from selective transformants was analyzed by Southern hybridization. As an outcome, Komari (1990) reported on transgenic cells in the suspension culture only. However, all attempts to regenerate plants from these cells were unsuccessful.

Notwithstanding, a recent Virus-Mediated transient expression technique (VIGS) protocol offers an alternative for functional studies given the lack of transformation protocols for quinoa (Ogata et al., 2021). Ogata et al. (2021) claimed that the apple latent spherical virus can induce gene silencing of quinoa. They carried out the silencing of genes involved in betalain production. They claimed that the apple latent spherical virus could be transmitted to the progeny of quinoa plants.

1.9. Hypotheses, objectives, and scientific questions

I present my study in three chapters. In the first chapter, I aimed to map QTL for agronomically important traits. For this aim, I used F₂ and F₃ segregating populations grown in the greenhouse and field, respectively. In the second chapter, I aimed to identify genes differentially responding to photoperiod and involved in flowering time by transcriptome analysis. For this purpose, I used the short-day accession D-12082 and the day-neutral accession PI-614886. Finally, in the third chapter, I aimed to investigate the feasibility of *Agrobacterium*-mediated transformation in quinoa.

Overall, I aimed to answer the following questions:

1. Which QTL control flowering time, maturity, plant height, panicle length, thousand kernel weight, and seed yield in quinoa?
2. When is the flowering initiation stage in quinoa accessions under short-day and long-day conditions?
3. Is the flowering time differently regulated under short-day and long-day conditions?
4. Do the candidate genes for photoperiod sensitivity express differently between photoperiod-sensitive and insensitive genotypes?
5. Can quinoa be transformed by the *Agrobacterium*-mediated transformation system?

2. High-density mapping of QTL controlling agronomically important traits in quinoa (*Chenopodium quinoa* Willd.)

Published in *Frontiers in Plant Science*, 2022
<https://doi.org/10.3389/fpls.2022.916067>

Nathaly Maldonado-Taipé¹, Federico Barbier¹, Karl Schmid², Christian Jung¹, Nazgol Emrani^{1*}

¹Plant Breeding Institute, Christian-Albrechts-University of Kiel, Olshausenstr. 40, D-24098 Kiel, Germany

²Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, Fruwirthstr. 21, 70599 Stuttgart, Germany

2.1. Abstract

Quinoa is a pseudocereal originating from the Andean regions. Despite quinoa's long cultivation history, genetic analysis of this crop is still in its infancy. We aimed to localize QTL contributing to the phenotypic variation of agronomically important traits. We crossed the Chilean accession PI-614889 and the Peruvian accession CHEN-109, which depicted significant differences in days to flowering, days to maturity, plant height, panicle length, thousand kernel weight (TKW), saponin content, and mildew susceptibility. We observed sizeable phenotypic variation across F₂ plants and F₃ families grown in the greenhouse and the field, respectively. We used Skim-seq to genotype the F₂ population and constructed a high-density genetic map with 133,923 SNPs. Fifteen QTL were found for ten traits. Two significant QTL, common in F₂ and F₃ generations, depicted pleiotropy for days to flowering, plant height, and TKW. The pleiotropic QTL harbored several putative candidate genes involved in photoperiod response and flowering time regulation. This study presents the first high-density genetic map of quinoa that incorporates QTL for several important agronomical traits. The pleiotropic loci can facilitate marker-assisted selection in quinoa breeding programs.

2.2. Introduction

Quinoa (*Chenopodium quinoa* Willd.) is a pseudocereal native to the Andean region of South America. It is an allotetraploid species ($2n = 4x = 36$), with a genome size of 1.45-1.50 Gb (Jarvis et al. 2017). Quinoa is characterized by its broad genetic variation and adaptation to biotic and abiotic stresses. It exhibits resistance to insects and diseases and tolerance to frost, drought, and salinity. Furthermore, quinoa seeds have outstanding physicochemical, nutritional, and functional properties for human consumption. Due to these unique qualities, quinoa is considered an option to improve world food security (Alandia et al. 2021).

Quinoa cultivation has transcended continental boundaries and it is present in Europe, Africa, and Asia (Alandia et al. 2020). However, substantial breeding efforts are still needed to explore all quinoa qualities and to expand its cultivation worldwide. Quinoa breeding aims for short, non-branching plants with a compact panicle, as well as increased tolerance to abiotic and biotic stresses. Nevertheless, the main breeding objective in quinoa remains to be the development of high-yielding varieties and, in temperate regions and high latitudes of Europe, North America, and China, the adaptation to long-day conditions (Murphy et al. 2018; Patiranage et al. 2021). Thus, for breeding quinoa, a better understanding of the molecular regulation of flowering time and day-length responsiveness is essential since yield potential and local adaptation are largely determined by these processes.

Despite being a domesticated crop, quinoa has not yet reached its full potential but molecular and genetic technologies may help change this situation (Alandia et al. 2021). In this scenario,

QTL mapping is useful to understand the genetic basis of quantitative traits. The use of sequencing technologies and computational analysis has made QTL detection easier. In skim sequencing (Skim-seq), genomes are sequenced at low coverage and sequence variants are called after mapping to a reference genome. Later, imputation is performed based on genetic linkage. Due to the large size of linkage blocks, Skim-seq is a suitable method for genotyping F_2 and F_3 segregating populations (Golicz et al. 2015; Kumar et al. 2021).

To date, only a few *C. quinoa* linkage maps are available. The first quinoa linkage map was constructed using 216 SSR (simple sequence repeats) markers using a recombinant inbred line (RIL) population. The map consisted of 38 linkage groups (LGs) covering 913 cM (Jarvis et al. 2008). Another linkage map contained 14,178 SNPs (KASPar genotyping) mapped in two RIL populations. This map consisted of 29 LGs spanning 1,404 cM (Maughan et al. 2012). A recent linkage map by Jarvis et al. (2017) combined the map from Maughan et al. (2012) with two new linkage maps. The resulting map contains 6,403 markers on 18 LGs spanning 2,034 centimorgans (cM). A few studies have attempted to identify loci for agronomically important traits in quinoa so far. Cervantes and van Loo (2017) identified QTL for color, flowering time, and yield-related traits using an F_2 population of 94 individuals from a cross between ‘Carina Red’ (bitter, dark seed) and ‘Atlas’ (non-bitter EU variety). They used a linkage map constructed with 1,076 SNPs and localized two major QTL, one for days to floral bud appearance on chromosome Cq6B, and another one for seed characters on chromosome Cq2B. In addition, a recent genome-wide association study with 2.9 million markers uncovered significant marker-trait associations for days to flowering, days to maturity, plant height, and panicle length on chromosome Cq2A (Patirange et al. 2020).

In this study, we aimed to create a high-density linkage map and localize QTL for agronomically important traits. A high-density linkage map was constructed with an F_2 population from a cross between a Chilean and a Peruvian accession. Agronomic traits were assessed in the F_2 population and the F_3 generation derived thereof. We mapped a number of highly significant QTL and we identified candidate genes within the QTL confidence intervals. Molecular markers tightly linked to the QTL can be helpful for marker-assisted selection in quinoa breeding programs.

2.3. Methods

2.3.1. Plant material and growth conditions

The Chilean quinoa accession PI-614889 (seed parent; seed code 171115) was crossed with the Peruvian inbred line CHEN-109 (pollen parent, seed code 170876) by applying hot water emasculation (Emrani et al. 2020). The F_1 plant was selfed to give rise to the F_2 population (seed code: 190031). The F_3 population (seed codes: 191203-191562) consisting of 334 families, was produced by selfing F_2 plants (Supplementary Table 1). A total number of 336 F_2 individuals and 10 plants of each parent were grown in 13 cm pots from March to October 2019 in a greenhouse under long-day conditions in Kiel, Germany. Seeds were harvested from August to October 2019. Three hundred thirty-four F_3 families and their parental lines were mechanically sown in a plant-to-row scheme in the field in 2020 (10.0°E 54.3°N, Achterwehr, Germany). One hundred fifty seeds were sown in two-meter single rows (one cm sowing depth) with 50 cm spacing between rows under a complete randomized block design with two blocks. Mechanical weeding was carried out four weeks after sowing using a row crop cultivator and hand weeding was performed five and seven weeks after sowing. Thinning was performed six weeks after sowing, aiming at 20 plants per row distanced at 10 cm.

2.3.2. Phenotypic evaluation

The following traits were assessed in both populations: days to flowering (DTF), days to maturity (DTM), plant height (PH), panicle length (PL), panicle density (PD), saponin content

(SC), and thousand kernel weight (TKW) (Supplementary Table 2). Mildew susceptibility (MS) was recorded only in the F₃ population in the field, while seed weight per plant (SW) and seed number per plant (SN) were recorded only in the F₂ population. In the F₂ trial, 336 individual plants and 10 plants of each parental line were phenotyped. In the field, ten plants per block and family and the parental lines were phenotyped (a total of 6,720 plants). Plants with significant biotic stress damage in the field (e.g., insect damage) were excluded from phenotyping (Supplementary Table 1). We carried out phenotyping at different BBCH stages, which were defined by Sosa-Zuniga et al. (2017). Additionally, to verify genetic segregation in the F₂ generation, we phenotyped red axil pigmentation in all 336 individuals.

2.3.3. DNA isolation and PCR

We verified genetic segregation by molecular marker analysis. Leaf genomic DNA was isolated from 48 F₂ plants and 194 F₃ plants by the standard CTAB method (Porebski et al., 1997). We used the InDel marker JASS5 (Supplementary Table 3) (Zhang et al. 2017). Polymerase chain reaction (PCR) and agarose gel electrophoresis were carried out following the details presented in Supplementary Table 4.

For whole-genome sequencing, we sampled young leaves from 336 F₂ plants at BBCH 22 and freeze-dried them. We extracted DNA from these samples by a modified protocol of the Genomic Micro AX Blood Gravity kit (A&A Biotechnology, Gdansk, Poland). The modifications to the protocol include a starting sample quantity of 20 mg of ground sample treated with 800 µl of lysis buffer for 30 min. Additionally, RNase treatment was performed: 10 µl RNase A solution (6 mg RNase A in 600 µl of water). We verified the quality of the isolated DNA by agarose gel electrophoresis (0.8%, 60 V, 60 min).

2.3.4. Whole-genome sequencing and bioinformatics

Whole-genome sequencing libraries were constructed using the protocol of Baym et al. (2015) and normalized for equimolarity using a BioTec Synergy HTC multimode plate reader. The library was sequenced by Illumina NovaSeq PE150. We aimed to ~1x coverage per sample of whole-genome sequences of the F₂ individuals (Skim-seq). The genome sequences of both parents, CHEN-109 and PI-614889, were already available with a coverage of 7.45x and 8.00x, respectively (Patirange et al. 2020). We trimmed raw reads with Trim_galore v 0.6.4 (parameters -q 30 -fastqc -paired) (Krueger 2015), sorted and indexed them with SAMTOOLS 1.10 (Li et al. 2009) and deduplicated them with MarkDuplicates (parameter REMOVE_DUPLICATES = TRUE tool of PICARD v2.21.9) (Broad-Institute 2019). Quality control was done with FastQC and MultiQC by removing reads containing N > 10% (N represents the percentage of the nucleotides that cannot be determined) and a quality base filter of Qscore ≤ 5 (over 50% of the total base). We mapped the reads to the Quinoa Reference Genome QQ74_V2 (CoGe Genome ID: id60716). We called variants using HaplotypeCaller in -ERC GVCF mode (Poplin et al. 2017). Markers were named as “chromosome number_physical position” (e.g., chr12_2345937). We kept only homozygous loci within each parent and considered only SNPs with a minimum base quality of 30 (minQ=30) and minor allele frequency (maf): 0.1. Then, we imputed the missing data by FSFHap (maf: 0.1 MaxMissing: 0.8; Window: 50) (Swarts et al. 2014). After imputation, we applied the following filters: min-alleles: 2; max-alleles: 2 max-missing: 0.3; maf: 0.1. Finally, we transformed the data to a parent-based format (.abh) by using the GenosToABH plugin from TASSEL (Glaubitz et al. 2014), using the codes A: pollen parent, B: seed parent, H: heterozygous. We performed quality control of the imputed data in .abh format by segregation distortion and percentage of missing data (ABHgenotypeR R package) (Reuscher and Furuta 2016). The bioinformatics pipeline is illustrated in Supplementary Figure 1A.

2.3.5. Linkage map construction

First, we performed final filtering of the F₂ population (Supplementary Figure 1B). We excluded F₂ plants with more than 30% missing data. Only markers present in more than 302 F₂ plants and fitting a 1:2:1 ($\alpha=0.05$) segregation ratio were used for linkage studies. We also excluded “identical” individuals with >95% sequence similarity. Then, we constructed the genetic map by MSTMap (Wu et al. 2008) with the following parameters: Kosambi function, cut-off p-value = 1e-09, no_map_dist: 15, no_map_size: 2, missing_threshold: 0.1. Markers with an estimated genetic distance $\leq 1.00\text{E-}04$ cM were clustered into bins. Finally, we performed several analyses for quality control of our genetic map. To begin with, we checked segregation distortion and estimated the number of crossovers and double-crossover following the guidelines given by Broman (2010). In addition, we analyzed parental allele frequencies and collinearity of our linkage map with the physical map. Moreover, we used a heatmap of our linkage map to look for switched alleles. We used LinkageMapViewR, ASMapR, and R/qtl for quality control of the genetic map (Broman et al. 2003; Ouellette et al. 2018; Taylor and Butler 2017).

2.3.6. QTL mapping and pleiotropy analysis

We carried out QTL mapping by composite interval mapping using the software package R/qtl version 1.46-2 (Haley-Knott with forward selection to three markers and a window size of 10 cM) (Broman and Sen 2009). The threshold for the logarithm of odds (LOD) for a significant QTL declaration was calculated by 1000 permutations of the genome-wide maximum LOD. The 95th and 99th percentile of this distribution were used as the genome-wide LOD thresholds (5% and 1% LOD thresholds). The confidence intervals were calculated using the 95% Bayes credible interval method. QTL effects were calculated with the nearest markers as the phenotypic differences between marker genotypes. The percentage of phenotypic variation explained by each QTL (R^2) was estimated by the “drop-one-QTL-at-a-time” analysis. A simple additive model for multiple QTL was generated for each trait using the multiple imputation method and the Haley-Knott regression. When a putative pleiotropy was observed, it was confirmed by the qtlpvl R package, and a multiple trait QTL analysis was performed (Tian et al. 2016). After confirmation of pleiotropy, pleiotropic sites were analyzed as single multitrait QTL (scanone.mvn) to obtain Bayes intervals and R^2 values.

2.3.7. Epistasis analysis

A genome-wide epistasis analysis was performed to describe how alleles influence each other. For this purpose, we used the cape R package (Tyler et al. 2013). To facilitate the analysis in terms of computational time, the software decomposed the phenotypes matrix into eigentrains (ET) by singular value decomposition (SVD). Then, we selected the two ET capturing the highest total variance among the traits to perform a pairwise scan of the variants (SNPs). From this scan, the software found interactions between alleles (epistasis), and the epistatic models were combined across ET to find allelic effects on the phenotypes included in the ETs. Positive and negative allelic effects refer to the reparametrized coefficient (either <0 or >0) from the pairwise regression as described by Tyler et al. (2013). Ultimately, the results of this analysis describe how alleles influence each other, in terms of enhancement (positive coefficient) and suppression (negative coefficient), as well as how gene variants influence phenotypes. The results of this analysis were plotted as heatmaps. ET contribution to the phenotypes was estimated for all bins of our genetic map, and heatmaps were constructed with 1,000 randomly selected markers. Effect calculations were performed in reference to the seed parent (PI-614889) allele.

2.3.8. Candidate gene identification and haplotype analyses

We retrieved annotated genes from the reference genome within the regions of the confidence intervals of each QTL to explore possible candidate genes (.gff from QQ74_V2; CoGe Genome ID: id60716). We selected preliminary candidate genes using the UniProt Knowledgebase database (UniProtKB). A gene was considered a candidate when a related function to the identified QTL was already described in other plant species. Then, we searched for variants (SNPs and InDels) within the parental sequences. From this search, we kept homozygous genes and gave preference to those variants with a putative effect on the function of the encoded protein. Following, we evaluated the haplotype of the selected variants in the F₂ population, as follows: we clustered the phenotypes according to the corresponding genotype at the variant site; later, we performed t-tests ($\alpha=0.05$) to compare DTF, PH, and TKW among the created clusters. To further evaluate the phenotypic effect of the variants, we used whole-genome sequencing and phenotypic data of 310 quinoa accessions grown in a two-year experiment in Kiel, Germany. This dataset comprises 2.9 million high confidence SNP and 414,891 InDel loci (Patirange et al. 2020). We followed the same procedure as for the haplotype evaluation in the F₂ population. We assigned letters for each allele to describe the genotypes (e.g., A_1A_1 homozygous, A_1A_2 heterozygous). A complete description of the nomenclature is given in Supplementary Table 5.

2.3.9. Heritability estimates and statistical analysis

The phenotypic (V_P), genotypic (V_G), environmental variances (V_E), and the broad sense heritabilities (h^2) were estimated using F₃ data (Falconer 1996). The heritability values were classified as low (below 30%), medium (30-60%), and high (above 60%) as suggested by Johnson et al. (1955). Genotypic coefficients of variation (GCV), phenotypic coefficients of variation (PCV), environmental coefficients of variation (ECV), and genetic advance with a selection intensity of 5% (GA) were calculated as described by Singh and Chaudhary (1977). In addition, phenotypic correlation coefficients (Pearson's r) of quantitative traits within and between the F₂ and F₃ populations were estimated using the phenotypic value of each F₂ plant and the average value of each F₃ family.

The mean of each F₃ family or parental line was used to estimate the corresponding phenotypic value. The percentage of transgression was defined as the number of plants with a phenotypic value exceeding the mean value of the superior parent or below the mean value of the inferior parent divided by the total number of phenotyped plants for the corresponding trait. We performed t-tests ($\alpha=0.05$) to detect significant differences between phenotypic values of the parental lines.

2.4. Results

2.4.1. Segregation and phenotypic analysis of F₂ and F₃ populations

We verified the expected 1:2:1 genetic segregation in the F₂ population by two approaches: phenotyping of the red axil pigmentation (complete dominance of red color over green color) (Simmonds, 1971) and molecular marker analysis. We phenotyped red axil pigmentation in all 336 F₂ individuals while genotyping was carried out for 48 individuals. Likewise, the expected segregation in the F₃ generation (3:2:3) was verified by molecular markers. One hundred ninety-four plants were genotyped from 20 randomly selected F₃ families (Supplementary Table 6 and Supplementary Figure 2).

Both populations, F₂ and F₃, exhibited a vast phenotypic variation under field and greenhouse conditions (Supplementary Figure 3). Moreover, substantial transgressive segregation was found for all traits (Table 3). The highest transgression percentage was found for TKW in the F₂ generation. On the other hand, heritabilities ranged between 38.02 and 91.06% with TKW

exhibiting the highest heritability value (91.06%). Besides, DTF, DTM, PD, and MS showed high heritability (79.77 to 82.99%) while PH and PL exhibited moderate heritability (38.44% and 38.02%, respectively) (Table 4). Importantly, only 34.9% of the F₃ plants reached maturity before harvesting in the field; thus, fewer plants were phenotyped for DTM in the F₃ population (Supplementary Table 1).

Table 3. Results from the quinoa populations grown in the greenhouse (F₂) and the field (F₃) and their parental lines. DTF: days to flowering, DTM: days to maturity, PH: plant height, PL: panicle length, PD: panicle density, TKW: thousand kernel weight, SW: seed weight per plant, SN: seed number per plant, SC: Saponin content, MS: Mildew susceptibility. NA: data not available due to late-maturity genotype.

Character	Population / parent	Mean \pm SD		Minimum	Maximum	Transgression (%)
		2019	2020			
DTF	CHEN-109	61 \pm 2 ^a	72 \pm 3 ^a	-	-	-
	PI-614889	76 \pm 1 ^b	96 \pm 4 ^b	-	-	-
	F ₂	69 \pm 5	-	55	87	29.17
	F ₃	-	82 \pm 9	67	77	14.63
DTM	CHEN-109	177 \pm 7 ^b	NA	-	-	-
	PI-614889	110 \pm 2 ^a	129 \pm 4	-	-	-
	F ₂	137 \pm 14	-	110	194	11.91
	F ₃	-	152 \pm 9	123	165	NA
PH (cm)	CHEN-109	151.40 \pm 7.06 ^b	164.74 \pm 27.56 ^b	-	-	-
	PI-614889	83.90 \pm 6.00 ^a	137.11 \pm 13.98 ^a	-	-	-
	F ₂	136.3 \pm 20.63	-	83.00	193.00	22.02
	F ₃	-	163.14 \pm 27.85	50.00	265.00	66.49
PL (cm)	CHEN-109	37.20 \pm 4.32 ^b	32.37 \pm 6.74 ^a	-	-	-
	PI-614889	16.30 \pm 2.49 ^a	29.47 \pm 9.41 ^a	-	-	-
	F ₂	24.07 \pm 6.14	-	14.00	50.00	8.33
	F ₃	-	32.85 \pm 10.4	10.00	80.00	76.31
PD	CHEN-109	3.00 \pm 0.00 ^a	1.05 \pm 0.23 ^a	-	-	-
	PI-614889	5.00 \pm 0.00 ^b	2.68 \pm 0.58 ^b	-	-	-
	F ₂	3.61 \pm 1.49	-	1	7	14.88
	F ₃	-	2.49 \pm 1.05	1	5	77.17
TKW (g)	CHEN-109	2.00 \pm 0.14 ^b	1.52 \pm 0.06 ^b	-	-	-
	PI-614889	2.44 \pm 0.31 ^a	2.64 \pm 0.03 ^a	-	-	-
	F ₂	2.75 \pm 0.49	-	1.18	3.91	77.97
	F ₃	-	2.62 \pm 0.29	1.58	3.45	48.17
SW (g)	CHEN-109	1.07 \pm 0.26 ^b	-	-	-	-
	PI-614889	5.90 \pm 0.88 ^a	-	-	-	-
	F ₂	4.10 \pm 1.48	-	0.06	7.22	13.29
SN	CHEN-109	633 \pm 161 ^b	-	-	-	-
	PI-614889	2,432 \pm 328 ^a	-	-	-	-
	F ₂	1469 \pm 471	-	51	3,251	6.25
SC	CHEN-109	6.84 \pm 0.20 ^b	6.5 \pm 2.12 ^b	-	-	-
	PI-614889	3.20 \pm 1.09 ^a	2.00 \pm 0.00 ^a	-	-	-
	F ₂	5.68 \pm 4.22	-	2.00	24.00	47.02
	F ₃	-	12.61 \pm 8.28	2.00	30.00	57.97
MS	CHEN-109	-	2.95 \pm 0.23 ^b	-	-	-
	PI-614889	-	2 \pm 0.64 ^a	-	-	-
	F ₃	-	2.34 \pm 0.69	1	3	59.42

Different letters ^{a,b} indicate significant differences between parental lines in each year (F₂:2019; F₃:2020) (t-test, $\alpha=0.05$).

Table 4. Statistical parameters calculated for eight phenotypic characters measured in the F₃ population. V_P: phenotypic variance, V_G: genotypic variance, V_E: environmental variance, PCV: phenotypic coefficient of variation, GCV: genotypic coefficient of variation, ECV: environmental coefficient of variation, h²: broad-sense heritability, GA: genetic advance as percentage of the mean. DTF: days to flowering, DTM: days to maturity, PH: plant height, PL: panicle length, PD: panicle density, TKW: thousand kernel weight, SC: saponin content, MS: Mildew susceptibility.

Parameter	DTF	DTM	PH	PL	PD	MS	TKW	SC
V _P	72.54	81.01	775.68	108.15	1.10	0.48	0.02	8.28
V _G	60.20	64.62	298.20	41.11	0.90	0.40	0.02	7.22
V _E	12.34	16.39	477.49	67.03	0.20	0.08	0.002	1.06
PCV (%)	10.34	5.94	17.07	31.66	42.06	41.67	5.70	22.81
GCV (%)	9.42	5.30	10.59	19.52	38.12	37.93	5.44	21.30
ECV (%)	4.26	2.67	13.39	24.93	17.78	17.26	1.70	8.16
h ² (%)	82.99	79.77	38.44	38.02	82.14	82.85	91.06	87.19
GA (%)	17.68	9.76	13.52	24.79	71.16	71.12	10.70	40.97

Then, we calculated correlations between all evaluated traits within years. The highest correlation was found between DTF and PH (Pearson's *r* in F₂: 0.69; Pearson's *r* in F₃: 0.63) (Figure 2). Both traits, PH and DTF, were significantly correlated with DTM. Furthermore, DTF showed a high negative correlation with TKW (F₂ and F₃) and with SN and SW (only F₂). In general, taller plants flowered later, reached maturity later, and depicted a reduction in the yield traits values, while shorter plants flowered earlier, reached maturity earlier, and showed higher values for the yield-related traits. Additionally, significant correlations for DTF, DTM, PH, PD, and PL between the F₂ plants grown under greenhouse conditions and their F₃ progenies were calculated, with the highest values for DTF (0.73) and PH (0.66). Surprisingly, SC showed a low correlation between years (Pearson's *r*: 0.17).

2.4.2. Sequencing the F₂ population revealed millions of SNPs

We sequenced the genomes of 336 F₂ plants, the parents of the F₃ families grown in the field. Skim-Seq by Illumina NovaSeq PE150 resulted in a total data output of 4.98 million raw PE reads on average per individual (~1.07x coverage). All reads passed the quality base filter requirement (Qscore ≤ 5), and 0.0053% of the raw data were removed due to a high number of nucleotides that could not be determined (N>10%)

Seventeen million SNPs were obtained after mapping and variant calling (Supplementary Figure 4). First, these SNPs were filtered by maf: 0.1 and minQ30, producing a data set of 4 million high-quality SNPs with a high percentage of missing data (~86%) (Supplementary Figure 5). After imputation, the proportion of missing data was reduced from ~86.0% to ~11.0%. Following the next filtering steps, we obtained a set of 249,744 high-quality biallelic SNPs with 4.2% missing data (Supplementary Figure 6), 21.2% homozygous markers for the pollen parent allele, 26.5% homozygous markers for the seed parent allele, and 52.3% heterozygous markers (Figure 3). We used this set of markers for genetic map construction.

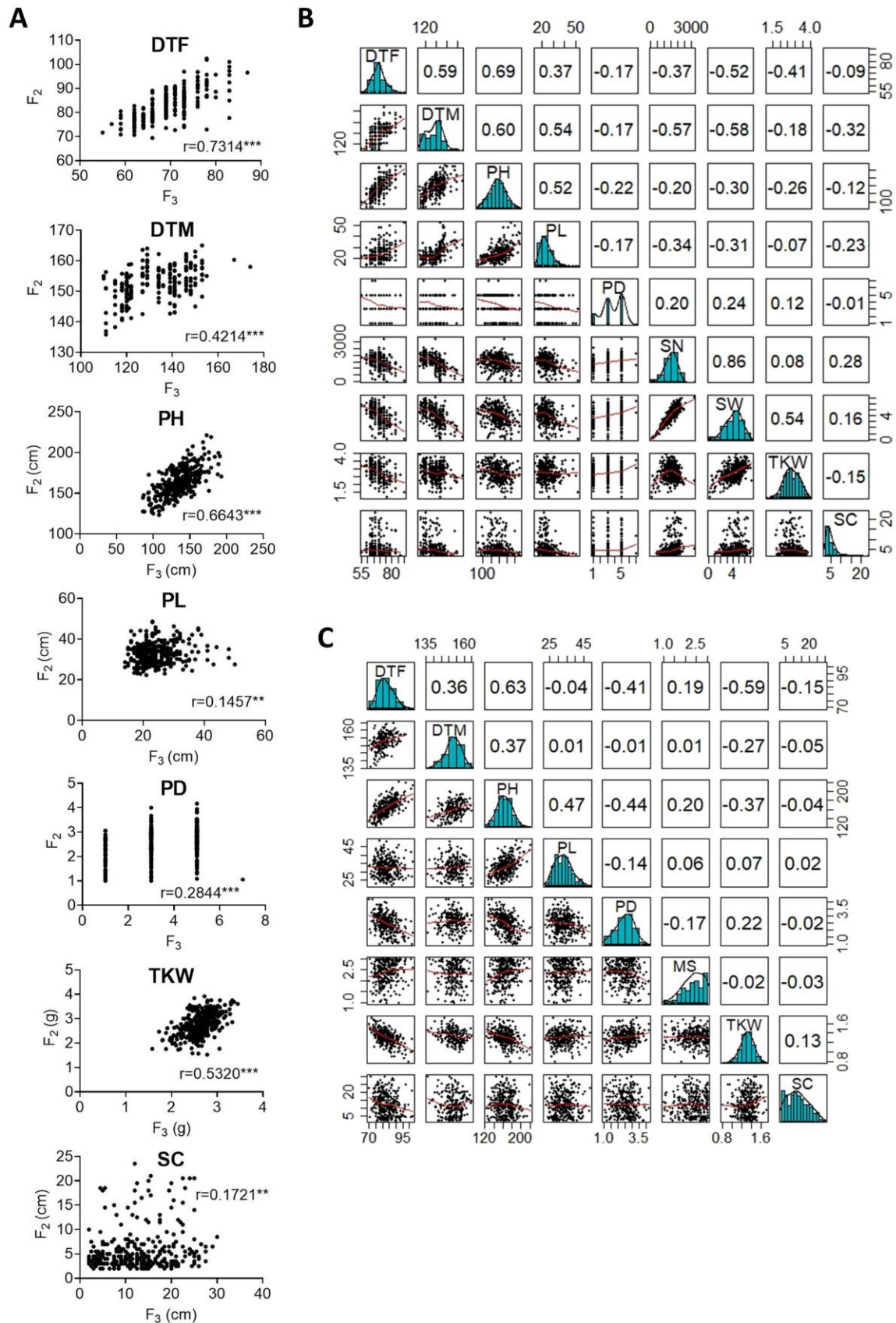


Figure 2. Correlation between phenotypic traits measured in the F_2 and F_3 populations. (A) Pearson's r correlations between F_2 and F_3 populations. (B) Correlations between nine traits measured in the F_2 population (C) Correlations between eight traits measured in the F_3 population. In (B) and (C), bivariate scatter plots are shown below the diagonal, histograms on the diagonal, and the Pearson correlation above the diagonal. DTF: days to flowering, DTM: days to maturity, PH: plant height, PL: panicle length, PD: panicle density, TKW: thousand kernel weight, SW: seed weight per plant, SN: seed number per plant, MS: Mildew susceptibility, SC: saponin content.

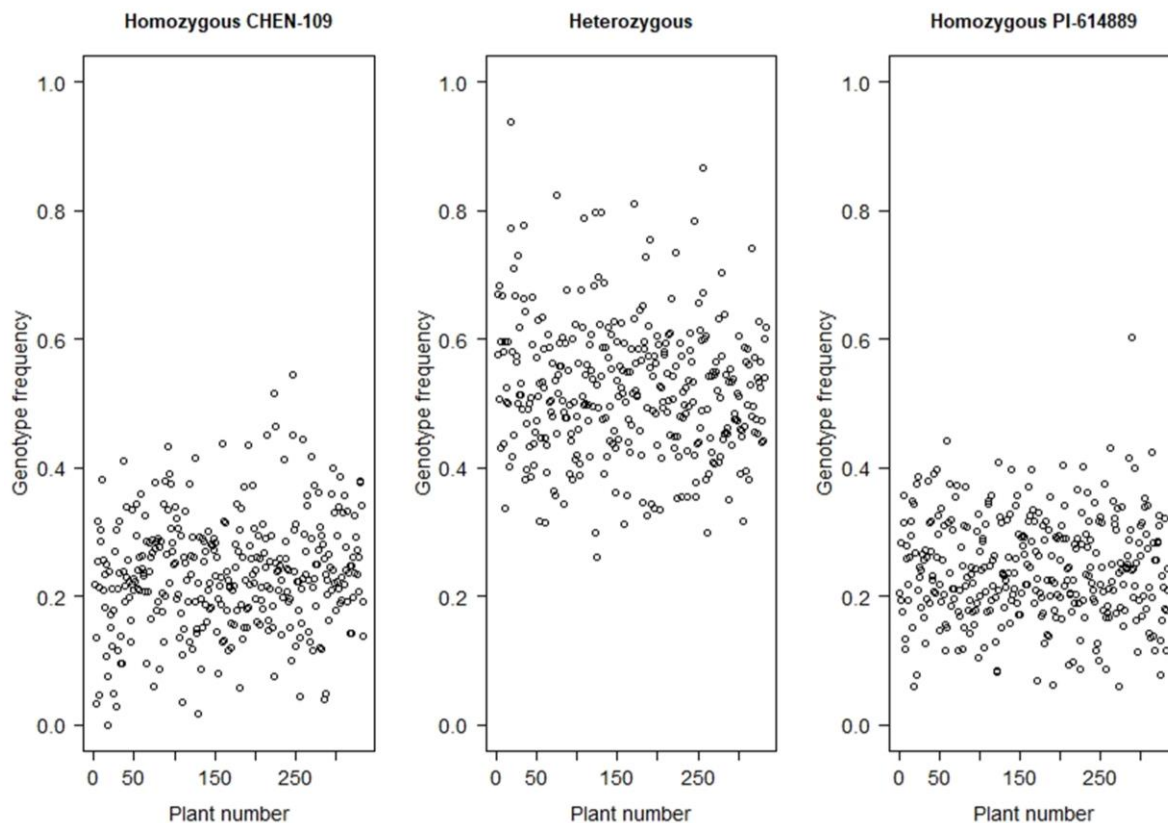


Figure 3. F₂ individuals' genotype frequencies: homozygous for the CHEN-109 parental allele, heterozygous and homozygous for the PI-614889 parental allele.

2.4.3. Construction of a high-density linkage map

Ahead of genetic map construction, the F₂ population sequences were cleaned based on our quality criteria. Two plants were removed due to >30% missing data (Supplementary Figure 6A), and 15,933 markers were removed because they were missing from >10% of the population (Supplementary Figure 6B). Another 99,898 markers were removed because they did not segregate in the expected 1:2:1 manner. No plants had to be removed due to high (>95%) sequence similarity to another F₂ plant (Supplementary Figure 6C). As an outcome of our final filtering, 334 F₂ plants were used to construct a genetic map with 133,913 markers, resulting in an average density of one marker per ~8.97 Kb. The resulting genetic map consists of 21 linkage groups (LG), with the chromosomes 5B, 6B and 8B split into two LGs each (Table 5, Supplementary Table 7). Moreover, the linkage map has an average density of ~88 markers per cM, where one cM corresponds to ca. 0.83 Mb (Supplementary Figure 8). For further steps, we created 5,219 bins where the markers with a genetic distance $\leq 1.00\text{E-}04$ cM were clustered into.

2.4.4. QTL mapping, pleiotropic loci identification, and epistasis calculation

We mapped QTL for ten agronomically important traits using phenotypic data of 334 F₂ plants and 328 F₃ families, which had passed our quality check (Supplementary Figure 6). Fifteen QTL were identified, ranging from one to three QTL per trait (Table 6). We found pleiotropy at seven QTL, which were named with the prefix “*pleio*” (Figure 6 and Supplementary Figure 11). Two QTL (*pleio4.1* and *pleio14.1*) were in common between F₂ and F₃, whereas six and seven QTL were found only in F₂ or F₃ populations, respectively. Together, *pleio4.1* and *pleio14.1* explained 22.01% of the phenotypic variation for TKW, PH, and DTF, being this the strongest effect observed among all QTL. *pleio20.1* and *pleio4.1* showed the highest additive and dominance effect, respectively.

Table 5. Summary statistics of the quinoa linkage map based on F₂ plants derived from a cross between CHEN-109 and PI-614889. The physical size of each chromosome was taken from the reference genome QQ_74 (CoGe Genome ID: id60716).

Chromosome	Size (Mb)	Linkage group	Size (cM)	No. of markers	No. of bins
Cq1A	57.13	1	82.3	11,679	323
Cq1B	71.68	2	79.9	2,377	180
Cq2A	53.75	3	87.3	11,397	310
Cq2B	73.59	4	110.4	5,637	340
Cq3A	57.38	5	97.5	8,827	349
Cq3B	72.83	6	77.5	8,770	253
Cq4A	59.38	7	114.3	9,473	380
Cq4B	71.74	8	94.3	9,634	360
Cq5A	64.66	9	90.8	10,824	323
Cq5B	78.03	10	22.4	1,162	103
		11	32.3	2,405	120
Cq6A	66.92	12	116.6	5,559	317
Cq6B	87.28	13	24.1	1,923	119
		14	25.3	873	109
Cq7A	57.68	15	51.9	3,879	205
Cq7B	75.96	16	91.2	9,004	273
Cq8A	59.27	17	104.1	12,687	502
Cq8B	75.04	18	7.00	570	42
		19	46.5	1,636	117
Cq9A	54.51	20	83.5	8,551	228
Cq9B	63.80	21	80.9	7,046	265
Total	1,200.71	-	1,520.1	133,913	5,218

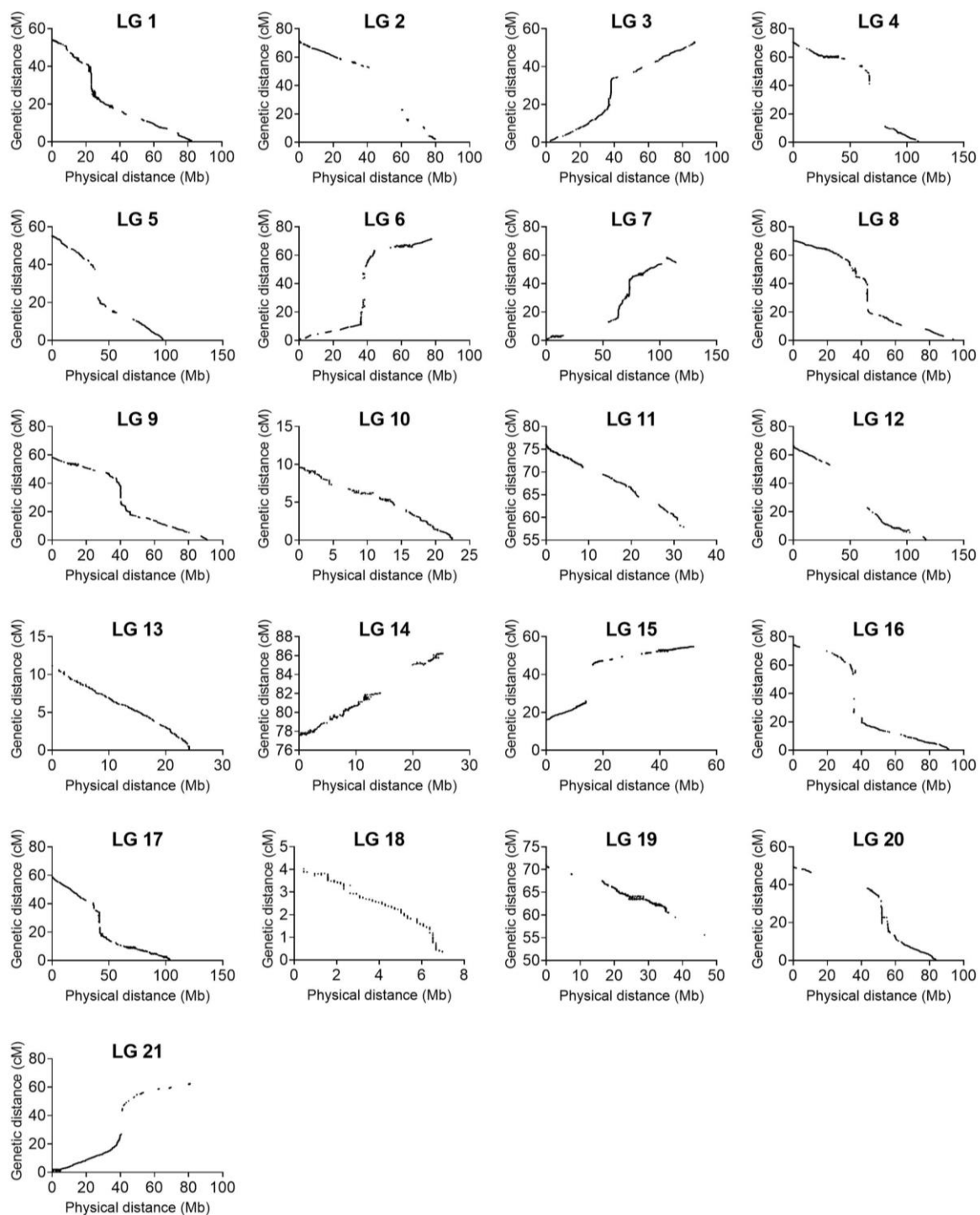


Figure 4. Collinearity between the linkage map constructed in this study and the physical map from the reference genome QQ74_V2 (CoGe Genome ID: id60716). The graphs were constructed with 133,913 non-binned markers.

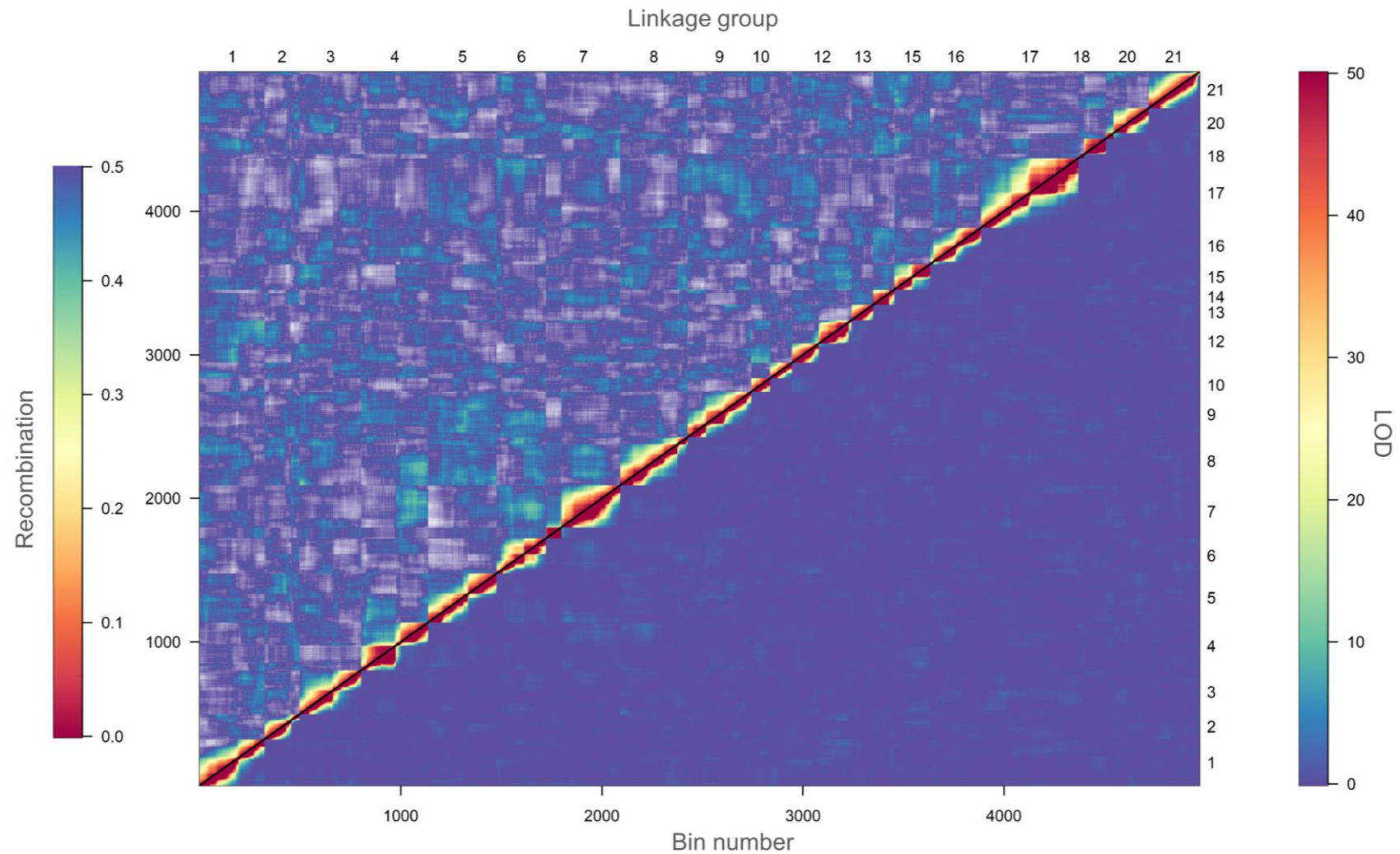


Figure 5. Heatmap of pairwise recombination fractions and LOD scores based on 5,218 bins. Estimated recombination fractions between binned markers are shown above the diagonal and LOD scores below the diagonal. Red colors indicate closely linked binned markers (high LOD score and low recombination fraction) whereas blue colors indicate non-linked binned markers (low LOD score and high recombination fraction). A LOD score of 50 corresponds to a recombination fraction of zero.

Table 6. Summary statistics of QTL mapping with the F₂ population and 328 F₃ families. The trait acronyms are explained in the methods section. LG: linkage group; Chr: chromosome, R²: estimated percentage of the phenotypic variance explained by the QTL.

Trait	Population	QTL name	Confidence Interval (cM)	Flanking markers	LG	Chr	R ² (%)		p-value	Effect ^a		No. of genes ^b
							QTL	Trait		Additive	Dominant	
DTF, PH, TKW	F ₂ and F ₃	<i>pleio4.1</i>	69.10-64.44	chr4_53552445, chr4_50421644	4	Cq2B	10.55	22.01	3.68E-03	-7.17	8.74	104
		<i>pleio14.1</i>	8.95-16.00	chr12_8042206, chr12_82045210	14	Cq6B	6.52		3.68E-03	6.73	-3.17	178
PH, PL, PD	F ₃	<i>pleio20.1</i>	10.39-22.00	chr17_4653672, chr17_38090534	20	Cq9A	10.86	10.86	6.47E-09	-9.13	-3.25	660
DTF, PH, PD, TKW		<i>pleio4.2</i>	69.00-64.59	chr4_53811433, chr4_50411615	4	Cq2B	12.44	12.44	3.43E-10	-6.79	7.68	110
DTM		<i>dtm3.1</i>	35.64-43.00	chr3_17349205, chr3_37087692	3	Cq2A	8.73	8.73	3.42E-05	-0.58	0.32	470
MS		<i>ms4.1</i>	1.68-7.00	chr4_69599485, chr4_66031413	4	Cq2B	0.38	0.43	0.264	-0.17	0.11	410
		<i>ms5.1</i>	70.00-76.69	chr5_14713288, chr5_9505959	5	Cq3A	0.06		0.955	-0.22	0.10	407
PD		<i>pd16.1</i>	42.32-44.72	chr14_1851842, chr14_17111206	16	Cq7B	10.80	20.44	1.99E-15	-0.28	0.03	47
DTF, DTM, PH, PL, SW, SN, TKW	F ₂	<i>pleio4.3</i>	59.62-62.00	chr4_53467737, chr4_50761666	4	Cq2B	16.21	16.21	1.78E-13	-3.21	0.41	88
DTF, PH, PD, SW		<i>pleio7.1</i>	5.32-20.00	chr7_2959770, chr7_3472874	7	Cq4A	10.97	10.97	5.02E-09	-2.31	-1.31	56
SN		<i>sn6.1</i>	35.89-37.54	chr6_12791379, chr6_23647528	6	Cq3B	2.77	5.46	0.010	1.02	0.15	313
TKW		<i>tkw17.1</i>	18.98-23.78	chr15_4808587, chr15_45144697	17	Cq8A	4.02	8.14	0.001	-0.96	-1.01	156
Saponin		<i>sap10.1</i>	0.00-2.63	chr10_9590242, chr10_8685894	10	Cq5B	6.60	21.98	3.68E-03	-1.51	-0.69	80
		<i>sap13.1</i>	14.61-20.93	chr12_5194791, chr12_2345937	13	Cq6B	9.31		3.68E-03	-1.39	-1.68	262
	<i>sap17.1</i>	2.25-10.21	chr15_5709440, chr15_52868334	17	Cq8A	6.63	3.68E-03		-1.43	0.60	365	

^a Relative to the PI-614889 allele

^b Genes within the confidence interval (retrieved from quinoa genome QQ_74 V2, CoGe Genome ID: id60716).

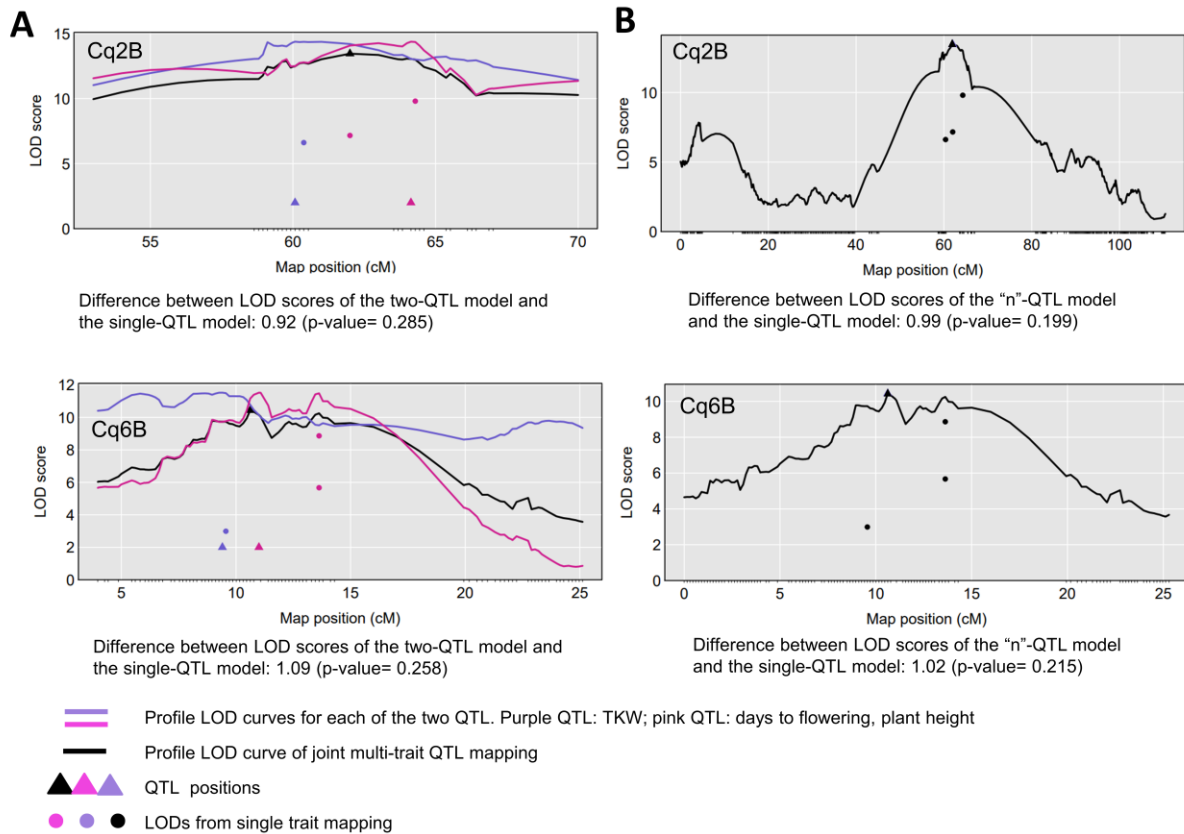


Figure 6. Comparative QTL analysis to detect pleiotropic QTL. Two tests were performed: (A) one vs. two QTL and (B) one vs. "n" QTL. Tests were performed considering the traits involved in the QTL found in common for F_2 and F_3 populations (top graphs: *pleio4.1*; bottom graphs: *pleio14.1*). The black curve is the LOD score curve for the single-QTL model, with the estimated QTL location indicated by a black triangle. The blue and pink curves are profile LOD score curves for the two-QTL model. Dots indicate the LOD score for the traits considering a single-QTL model. The LOD scores for each model were also calculated and the difference between these LOD scores is reported below the plots.

We performed a genome-wide epistasis analysis to investigate how alleles influence each other in terms of enhancement and suppression and also examined how different alleles influence phenotypes (DTF, PH and TKW). As the first step of this epistasis analysis, the phenotype matrix was decomposed by singular value decomposition (SVD) into eigentraits (ET). Two ET captured 69.00% and 21.00% of the total variance among DTF, PH, and TKW and were selected to perform a pairwise scan of the SNPs (Figure 7A). Then, we constructed a heatmap where 46,52% of the alleles at the genome level had a minor simultaneous effect (>-1 or <1) on DTF, PH and TKW. Moreover, alleles located within the pleiotropic region *pleio4.1* showed 252 significant interactions with other alleles in all LGs, except for LGs 14, 15 and 18. Interestingly, we found that 96.82% of the alleles located within *pleio4.1* (source 4 in Figure 7B) had suppressive interaction with other alleles at the genome level (reparametrized coefficient <0). Moreover, while PI-614889 alleles at *pleio4.1* (source 4 in Fig 6B) had a negative effect on DTF and PH and a positive effect on TKW, PI-614889 alleles at *pleio14.1* (source 14 in Figure 7B) had a positive effect on DTF and PH and a negative effect on TKW. Positive and negative SNP effects refer to the reparametrized coefficient (either <0 or >0) from the pairwise regression as described by Tyler et al. (2013).

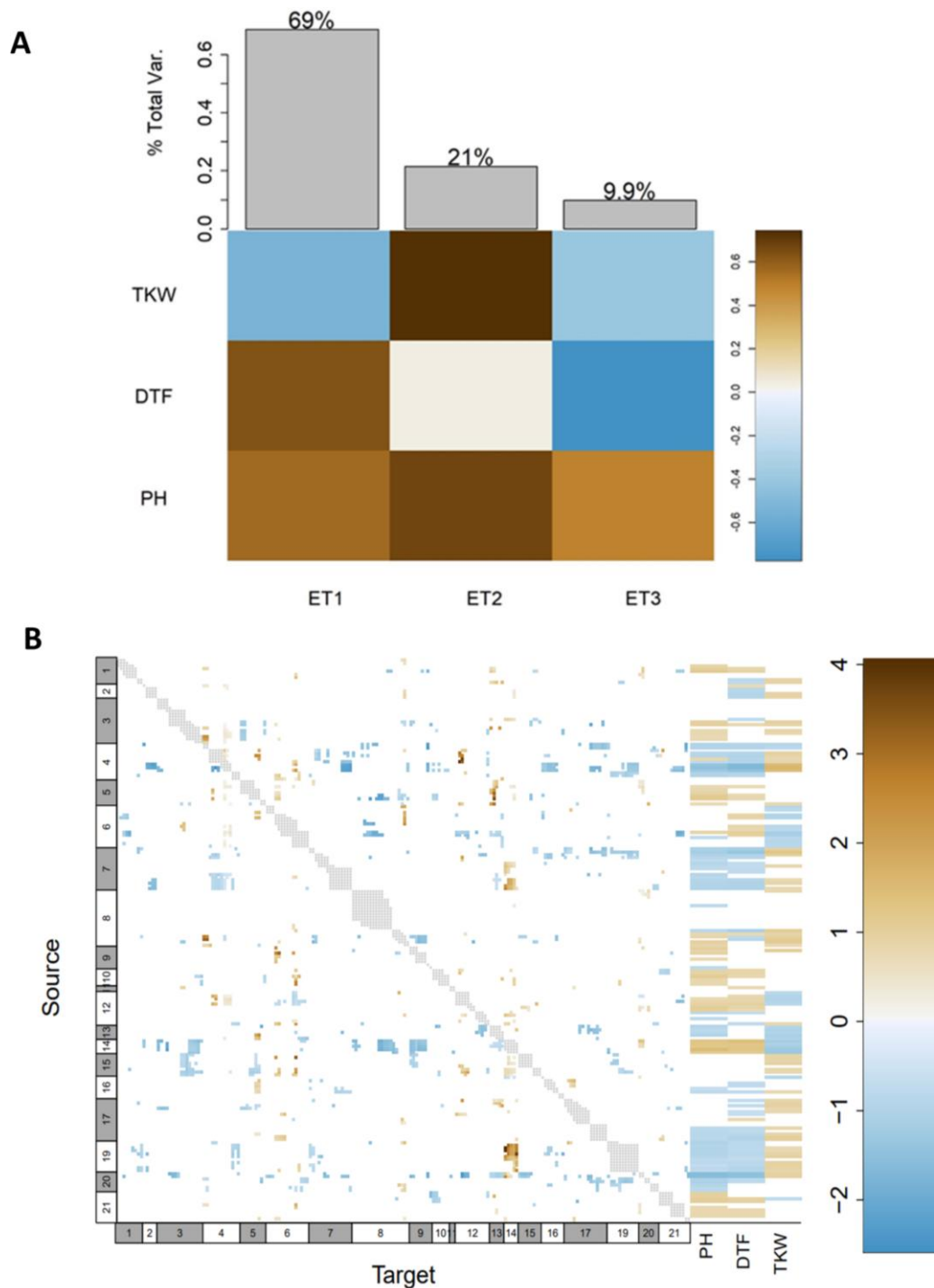


Figure 7. Genome-wide epistasis analysis and effects of alleles at genome level on days to flowering (DTF), plant height (PH) and thousand kernel weight (TKW). (A) Decomposition of phenotypes into eigentraits (ET). The colors of the heatmap correspond to the global variance fraction of each ET. (B) Heatmap showing interactions between all pairs of alleles at the genome level, resulting from a pairscan analysis. Every allele was assigned as “source” and “target” for the pairscan analysis. To the right of the heatmap, the interaction of every allele (assigned as source) on the phenotypes involved in the pleiotropy is shown. The heatmap scale represents the reparametrized coefficient calculated by the software and might be interpreted as the direction of the allelic effect. Grey dots show marker pairs that were not included in the pairwise scan due to complete linkage. Numbers at the x and y axes in white and grey boxes correspond to linkage groups.

2.4.5. Identification of putative candidate genes controlling agronomically important traits

We searched for candidate genes within the confidence intervals of all QTL. We reasoned that trait-related SNPs could be found within or close to the genes contributing to quantitative variation (quantitative trait genes, QTG). Altogether, 1,874 genes were found within non-overlapping confidence intervals of fifteen QTL (Supplementary Table 7). Nevertheless, we focused on the QTL *pleio4.1* and *pleio14.1* because of their pleiotropic effects on multiple traits and because they were common in F₂ and F₃ populations. Accordingly, the QTL *pleio4.1* and *pleio14.1* contributed to the phenotypic variation of three traits: DTF, PH and TKW, and 282 genes were identified within their confidence intervals.

Among the 282 genes described above, we found 41 genes with a previously described function related to flowering-time, photoperiod, and yield regulation in other plant species. Later, we compared the sequences of these genes between both parents of the population (Supplementary Table 8). From all the SNPs and InDels that differed between the parents for the selected genes, we chose those that were homozygous for each parent and had a putative effect on the function of the encoded protein. From 83 selected variants, we could identify only seven SNPs in the sequencing data of the F₂ population (Supplementary Figure 12). To assess the possible effects of the variants, we grouped the F₃ plants according to the corresponding F₂ genotype at the variant locus and performed t-tests ($\alpha=0.05$) to compare DTF, PH, and TKW between the groups. As result, none of the analyzed variants explained the phenotypic variance observed for DTF, PH and TKW (Supplementary Figure 12).

Afterward, we used available whole-genome sequencing and phenotypic data of a quinoa diversity set (310 quinoa accessions grown in a two-year experiment in Kiel, Germany) (Patirange et al. 2020) to perform the same analysis. Namely, we grouped the diversity set based on the genotypes of the F₂ parents (either CHEN-109 or PI-614889) at the variant locus and performed t-tests ($\alpha=0.05$) to compare DTF, PH, and TKW among the created groups. As result, we observed several significant phenotypic differences for PH and/or TKW and/or DTF when we grouped the quinoa diversity set based on the genotypes of our seed and pollen parents (Supplementary Figure 13). Interestingly, a missense SNP variant at *TSL-KINASE INTERACTING PROTEIN 1 (TKI1)* had simultaneous significant effects on DTF, PH and TKW (Figure 8). Accessions in the quinoa diversity set that carried the PI-614889 genotype (*N₁N₁*) at chr12_81633685 flowered earlier, were shorter and showed higher TKW values than those carrying the CHEN-109 (*N₂N₂*) genotype. Furthermore, a frameshift variant and a disruptive in-frame deletion at *DNA (CYTOSINE-5)-METHYLTRANSFERASE 1 (MET1b)* had simultaneous significant effects on DTF, PH and TKW. Accessions in the quinoa diversity set carrying the deletion and the frameshift variant (CHEN-109 genotype) flowered later, were taller and had lower TKW values than those accessions without the deletion and the frameshift variant (PI-614889 genotype). A similar scenario was observed for a disruptive in-frame insertion at *RICESLEEPER3*, where the accessions carrying the insertion (CHEN-109 genotype) had higher values of DTF, PH and lower TKW values (Figure 8).

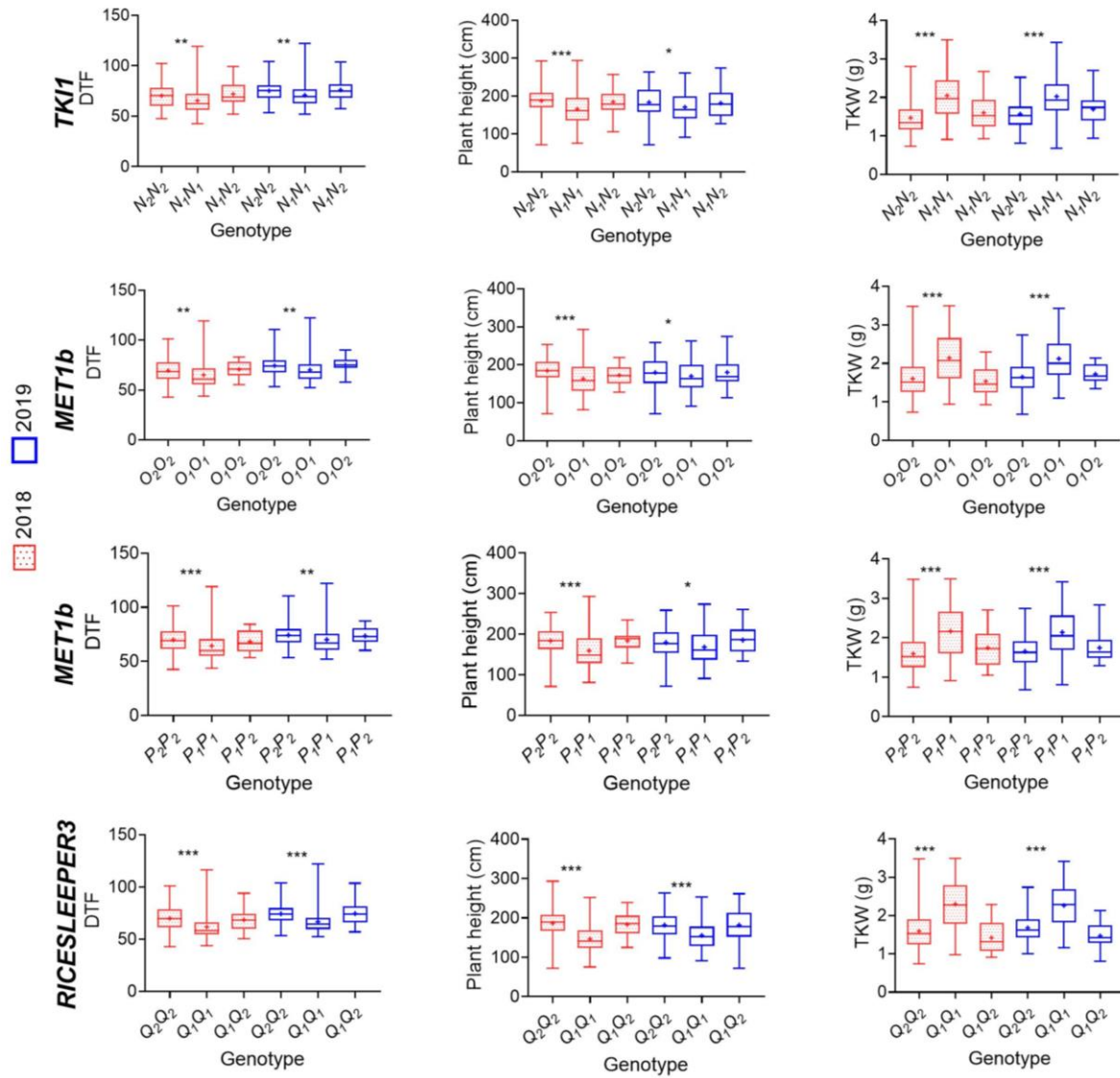


Figure 8. Evaluation of variant haplotypes using available whole-genome sequencing and phenotypic data of 310 quinoa accessions (Patiranage et al., 2020). Phenotypic effects of haplotype variations within four candidate genes are shown: *TSL-KINASE INTERACTING PROTEIN 1* (*TK11*) (SNP: chr12_81633685), *DNA (CYTOSINE-5)-METHYLTRANSFERASE 1* (*MET1b*) (frameshift variant chr4_56534732), *DNA (CYTOSINE-5)-METHYLTRANSFERASE 1* (*MET1b*) (disruptive inframe deletion chr4_56534915), and *RICESLEEPER3* (disruptive_inframe_insertion chr4_55091902). The variants genotypes correspond to, for instance, N_1N_1 (our homozygous parent PI-614889), N_1N_2 (heterozygous), N_2N_2 (our homozygous parent CHEN-109) and are described in Table S4. Significant differences between genotypes are shown by asterisks (t-test, $\alpha < 0.05 = **$, $\alpha < 0.01 = *$, $\alpha < 0.001 = ***$). Phenotypic data of different years are shown in different colors. DTF: days to flowering, TKW: thousand kernel weight.

2.5. Discussion

We exploited the recent advances in sequencing technologies and computational analysis methods to localize QTL for agronomically important traits in quinoa. A high-density genetic map was constructed with a segregating F_2 population, and 15 QTL were mapped with phenotypic data from two segregating generations. Candidate genes underlying the quantitative variation were identified within the QTL.

We calculated broad-sense heritabilities and genetic advance (GA) with a selection intensity of 5%, which resulted in moderate to high across all traits (excluding days to maturity). Interestingly, the high heritability coupled with high GA observed for days to flowering, panicle density and saponin content indicates that selection may be more effective for these traits. Moreover, previous studies reported heritability values for days to flowering of 70.1%

(two-year experiment with five quinoa genotypes) (Al-Naggar et al. 2017) and 91.0% (quinoa diversity set phenotyped for two years) (Patirange et al. 2020). The same studies calculated heritabilities of 89.7% for TKW and 68.0% for panicle density. Thus, the stated values in our study are in accordance with previous reports. Nevertheless, Al-Naggar et al. (2017) and Patirange et al. (2020) reported values of 85.0% and 60.7% for plant height. Differently, in our study, the traits plant height and panicle length showed moderate heritability, both with values of around 38%. A possible explanation is that the pollen parent showed a wide range and large variability in plant height, resulting in lower estimates of heritability.

Our study considered Skim-seq as a genome complexity-reduction method for constructing a genetic map for quinoa. In our hands, genotyping by Skim-seq was effective for QTL mapping and could be applied for a minor crop like quinoa, for which available resources and commercial interest are currently limited (Böndel and Schmid 2021). We showed, by several quality controls, that the challenge of calling high-quality heterozygous SNPs at low sequencing coverage ($\sim 1.07\times$) could be overcome by modifications to conventional bioinformatics pipelines and imputation. Moreover, our results showed that whole-genome sequencing with coverage as low as $\sim 1.0\times$ would be sufficient for QTL mapping. QTL mapping using whole-genome low-coverage sequencing has been successfully applied in chickpea and tobacco. In these studies, RILs and backcross populations were sequenced with depths from ca. 0.75 to $1.0\times$, and the number of markers for mapping ranged from $\sim 4,000$ to $\sim 53,000$ (Kale et al. 2015; Tong et al. 2021). Although Skim-seq was sufficient for constructing a high-density genetic map, there are limitations to this method. First, there are no tools available for the accurate imputation of InDels in F_2 populations of polyploid crops. Thus, further uses of our genotypic data are restricted only to SNPs. The second important drawback of this approach is that the centromeric regions cannot be accurately sequenced due to their repetitive nature. However, these problems might be addressed soon by the development of more sophisticated bioinformatics tools for imputation (Jordan et al. 2022).

Using Skim-seq and phenotyping, we mapped 15 QTL for ten different traits with a wide range of explained phenotypic variation (from 0.43% for MS to 22.01% for PH, DTF and TKW) and a wide range of the contribution of individual QTL to the total phenotypic variation (from 0.06% to 16.21%). QTL explaining the total phenotypic variance are unreachable because variation in quantitative traits is affected by many small-effect QTL (often undetectable) with additive and dominance effects, and QTL-QTL interactions. Importantly, in our study, the use of families in the F_3 generation allowed plants of the same family to be considered as replicates, allowing measurement of environmental variances and thus, significantly increasing the power and precision of QTL detection. Accordingly, the accuracy of our QTL analysis is supported by the QTL *sap10.1*, found for saponin content. This QTL, which was previously reported in other studies (Jarvis et al. 2017; Patirange et al. 2020), harbors a *TRITERPENE SAPONIN BIOSYNTHESIS ACTIVATING REGULATOR 2* (*TSAR2*) basic helix–loop–helix (bHLH) transcription factor, which likely controls the production of anti-nutritional triterpenoid saponins in quinoa seeds. However, no QTL was found for saponin content in the F_3 generation. This might be due to the method, which was used to phenotype saponin content in this generation. We bulked seeds from ten plants in the field corresponding to one F_3 family. From these bulks, we took samples of 20 seeds for saponin analysis. Thus, bulked samples of 20 seeds might not have been true representatives of thousands of seeds from an F_3 family. Moreover, the relatively low correlation coefficient between the saponin content measured in F_2 and F_3 generations might be also a result of the low efficiency of this method for measuring the saponin content in the F_3 population. We suggest measuring the saponin content separately in multiple single plants of every F_3 family to obtain more accurate values for this trait in F_3 .

We consider the common QTL among F_2 and F_3 populations, which depicted pleiotropy, as the most relevant QTL in our study. These QTL, explaining the phenotypic variation of days

to flowering, plant height and TKW, were located at Cq2B and Cq6B. In contrast, a putative pleiotropic locus controlling days to flowering, days to maturity, plant height and panicle length was found at Cq2A by Patirange et al. (2020). The different outcomes might respond to the different type of population that was used in our study as compared to Patirange et al. (2020). Accordingly, we could expect different alleles segregating in our biallelic F₂ population than the ones studied by Patirange et al. (2020), who used a quinoa diversity set comprising more than 300 accessions. Furthermore, we found a strong correlation between days to flowering, plant height, and TKW (Fig 1), which were the traits implicated in pleiotropy. This reinforces their QTL co-localization at Cq2B and Cq6B. Similar correlations, where quinoa plants flowering earlier are shorter and depict higher yield-related traits have been observed by Manjarres-Hernández et al. (2021) and Patirange et al. (2020).

We reported additive and dominance effects in a wide range of values, where the highest additive effect was observed for *pleio20.1* (-9.13), a QTL explaining the phenotypic variation of plant height, panicle length and panicle density, and the highest dominant effect was reported for *pleio4.1* (8.74). However, these results solely show the effects of single loci. Since quantitative variation in phenotypes results from highly complex gene networks and epistasis, we aimed to investigate the allele effects (relative to the PI-614889 allele) at genome-wide level on days to flowering, plant height, and TKW. We observed that nearly half of the studied alleles (46.52%) had a minor simultaneous effect on these three phenotypes; thus, confirming the nature of these traits as quantitative. From this analysis, we could also corroborate that *pleio4.1* and *pleio14.1* are major QTL, which themselves explain 10.55% and 6.52% of the phenotypic variance, respectively. Moreover, while PI-614889 alleles at *pleio4.1* had a negative effect on DTF and PH and a positive effect on TKW, PI-614889 alleles at *pleio14.1* had a positive effect on DTF and PH and a negative effect on TKW. This is correlated with the additive effects calculated for the markers with the highest LOD score within each QTL (Table 4). The observed contrasting effects of PI-614889 alleles at *pleio4.1* and *pleio14.1* would probably complicate breeding processes, whose aim is to reduce days to flowering and plant height and increase TKW, simultaneously (Murphy et al. 2018; Patirange et al. 2021). In a broader sense, our results from the genome-wide epistasis analysis revealed the complexity of the regulation of days to flowering, plant height, and TKW in quinoa, which was expected given the intricate processes, such as DNA methylation, histone modification, and non-coding RNA-associated gene silencing, which underlie these traits (Pandey et al. 2021).

A search for candidate genes within the confidence intervals of the QTL was performed. We reasoned that trait-related SNPs could be found within the genes contributing to quantitative variation. Within the 15 different QTL described in our study, we found homologs of known flowering time (*HEADING DATE 3A*, *WRKY TRANSCRIPTION FACTOR 13*, *FLOWERING LOCUS D*), plant architecture (*APETALA2-1*) and yield-related (*SMALL BASIC INTRINSIC 1-2*, *SWEET*) genes from other species (He et al. 2003; Li et al. 2016; Ma et al. 2017; Patil et al. 2015; Wang et al. 2015; Yuan et al. 2014). Besides, we identified *FLOWERING LOCUS T* (*CqFT2A*) within *pleio7.1*, a QTL found in the F₂ population, exclusively. Although *FT* genes are described in studies related to flowering time regulation in quinoa and *C. rubrum*, no flowering-time function has been specified particularly for the *CqFT2A* paralog (Patirange et al. 2021; Štorchová 2020). To continue towards the identification of candidate genes, we mainly focused on the two pleiotropic QTL detected in F₂ and F₃ populations and identified putative candidate genes based on their known functions in flowering time and yield regulation in other species. Then, we investigated the effect of different sequence variants in these genes in a quinoa diversity set. As an outcome and even though non-related accessions could have different haplotypes although they possess the same SNP or InDel in the candidate gene, we found several sequence variants that significantly explained the phenotypic variation of PH and/or TKW and/or DTF in the quinoa diversity set. Hence, the genes containing these variants (either SNP or InDel) might be interesting candidates for further studies. Among

these genes, we found that sequence variation at *TSL-KINASE INTERACTING PROTEIN 1* (*TKII*) had significant simultaneous effects on days to flowering, plant height and TKW, which were the phenotypes whose variation was partly explained by the QTL *pleio14.1* in our study. In Arabidopsis, *TKII* interacts with *TOUSLED* (*TSL*) and *TSL* loss of function mutations has pleiotropic effects on both leaf and flower development. Loss of *TSL* function also affects flowering time since it is required in the floral meristem for the correct initiation of floral organ primordia (Ehsan et al. 2004; Roe et al. 1993). Therefore, it is tempting to speculate that *TKII* is involved in the regulation of flowering time, flower development and consequently, seed set in quinoa. Moreover, we found that sequence variations at *MET1b* and *RICESLEEPER3* have similar simultaneous effects on days to flowering, plant height and TKW, as observed for the variant at *TKII*. In Arabidopsis, *MET1* homozygous mutants displayed late-flowering phenotypes caused by ectopic expression of *FLOWERING WAGENINGEN* (*FWA*), a regulator of flowering time. Hypomethylation, which correlates with the mentioned late-flowering phenotypes, is often accompanied by other developmental alterations (Kakutani et al. 1996; Saze et al. 2003; Soppe et al. 2000). Furthermore, *DAYSLEEPER*, the Arabidopsis homolog of *RICESLEEPER3*, encodes for a transposase-like protein essential for plant growth and development. Moreover, loss of function mutants of *DAYSLEEPER* showed retarded growth and delayed flowering (Bundock and Hooykaas 2005; Knip et al. 2012). Importantly, the evidence of the role of these genes in the regulation of different biological processes is given for the model plant Arabidopsis while the observed pleiotropic regulation of days to flowering, plant height, and TKW in our study might be controlled by quinoa-specific genes. Hence, if *TKII*, *MET1*, and *RICESLEEPER3* have the same function in quinoa can only be verified by further investigations. A first step towards elucidating the molecular mechanism governed by these genes might be expression analysis, for instance. Furthermore, haplotype analyses may focus on up-and downstream regulatory regions of the most relevant candidate genes in our study. Besides, despite the lack of reliable transformation protocols for quinoa, screening mutants and assessing their phenotypic effects seems to be another feasible approach for determining the function of these genes. Moreover, a recent study offers perspectives for functional studies in quinoa using virus-induced gene silencing (VIGS) (Ogata et al. 2021).

On the other hand, molecular markers linked to the pleiotropic QTL identified in the current study can be directly used in quinoa breeding programs for the simultaneous selection of different traits. Moreover, the provided information about QTL effects could guide breeders toward the selection of early, short, and high-yielding quinoa genotypes. Future work may address fine mapping of the interesting pleiotropic regions and characterization of candidate genes. Overall, the results presented in this study will help provide a framework for future research on the molecular mechanisms of flowering and other agronomically important traits and facilitate marker-assisted selection (MAS) in quinoa breeding programs.

2.6. Acknowledgements

We thank Monika Bruisch (Kiel University), Brigitte Neidhardt-Olf (Kiel University) and Elisabeth Kokai (University of Hohenheim) for technical assistance. Mireia Vidal-Villarejo (University of Hohenheim) provided support in the bioinformatic analysis of sequencing data. Prof. Dr. Mark Tester and Dr. Elodie Rey (King Abdullah University of Science and Technology) for providing access to the reference genome QQ74_V2 and the parental sequences, and for their continuous advice.

2.7. References

Al-Naggar A, Abd El-Salam R, Badran A, El-Moghazi MM (2017) Heritability and interrelationships for agronomic, physiological and yield traits of Quinoa (*Chenopodium quinoa* Willd.) under elevated water stress. Archives of Current Research International:1-15

- Alandia G, Odone A, Rodriguez JP, Bazile D, Condori B (2021) Quinoa—Evolution and future perspectives. *The Quinoa Genome*. Springer, pp 179-195
- Alandia G, Rodriguez J, Jacobsen S-E, Bazile D, Condori B (2020) Global expansion of quinoa and challenges for the Andean region. *Global Food Security* 26:100429
- Baym M, Kryazhimskiy S, Lieberman TD, Chung H, Desai MM, Kishony R (2015) Inexpensive multiplexed library preparation for megabase-sized genomes. *PloS one* 10:e0128036
- Böndel KB, Schmid KJ (2021) Quinoa Diversity and Its Implications for Breeding. *The Quinoa Genome*. Springer, pp 107-118
- Broad-Institute (2019) “Picard Tools.” Broad Institute. GitHub repository
- Broman K (2010) Genetic map construction with R/qtl. University of Wisconsin-Madison, Department of Biostatistics & Medical
- Broman KW, Sen S (2009) Two-dimensional, two-QTL scans. *A Guide to QTL Mapping with R/qtl*. Springer, pp 213-239
- Broman KW, Wu H, Sen S, Churchill G (2003) R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19:889-890
- Bundock P, Hooykaas P (2005) An Arabidopsis hAT-like transposase is essential for plant development. *Nature* 436:282-284
- Cervantes DP, van Loo E (2017) QTL mapping for agromorphological traits in Quinoa (*Chenopodium quinoa* Willd.)
- Ehsan H, Reichheld J-P, Durfee T, Roe JL (2004) *TOUSLED* kinase activity oscillates during the cell cycle and interacts with chromatin regulators. *Plant Physiology* 134:1488-1499
- Emrani N, Hasler M, Patiranage DS, Nathaly MT, Rey E, Jung C (2020) An efficient method to produce segregating populations in quinoa (*Chenopodium quinoa*). *Plant Breeding Reviews* 139:1190-1200
- Falconer DS (1996) Introduction to quantitative genetics. Pearson Education India
- Glaubitz JC, Casstevens TM, Lu F, Harriman J, Elshire RJ, Sun Q, Buckler ES (2014) TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PloS one* 9:e90346
- Golicz AA, Bayer PE, Edwards D (2015) Skim-based genotyping by sequencing. *Plant Genotyping*. Springer, pp 257-270
- He Y, Michaels SD, Amasino RM (2003) Regulation of flowering time by histone acetylation in Arabidopsis. *Science* 302:1751-1754
- Jarvis D, Kopp O, Jellen E, Mallory M, Pattee J, Bonifacio A, Coleman C, Stevens M, Fairbanks D, Maughan P (2008) Simple sequence repeat marker development and genetic mapping in quinoa (*Chenopodium quinoa* Willd.). *Journal of Genetics* 87:39-51
- Jarvis DE, Ho YS, Lightfoot DJ, Schmöckel SM, Li B, Borm TJ, Ohyanagi H, Mineta K, Michell CT, Saber N (2017) The genome of *Chenopodium quinoa*. *Nature* 542:307
- Johnson HW, Robinson H, Comstock RE (1955) Genotypic and phenotypic correlations in soybeans and their implications in selection. *Agronomy journal* 47:477-483

- Jordan KW, Bradbury PJ, Miller ZR, Nyine M, He F, Fraser M, Anderson J, Mason E, Katz A, Pearce S (2022) Development of the Wheat Practical Haplotype Graph database as a resource for genotyping data storage and genotype imputation. *G3* 12:jkab390
- Kakutani T, Jeddeloh JA, Flowers SK, Munakata K, Richards E (1996) Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. *Proceedings of the National Academy of Sciences* 93:12406-12411
- Kale SM, Jaganathan D, Ruperao P, Chen C, Punna R, Kudapa H, Thudi M, Roorkiwal M, Katta MA, Doddamani D (2015) Prioritization of candidate genes in “QTL-hotspot” region for drought tolerance in chickpea (*Cicer arietinum* L.). *Scientific Reports* 5:1-14
- Knip M, de Pater S, Hooykaas PJ (2012) The *SLEEPER* genes: a transposase-derived angiosperm-specific gene family. *BMC Plant Biology* 12:1-15
- Kumar P, Choudhary M, Jat B, Kumar B, Singh V, Kumar V, Singla D, Rakshit S (2021) Skim sequencing: an advanced NGS technology for crop improvement. *Journal of Genetics* 100:1-10
- Li W, Wang H, Yu D (2016) Arabidopsis WRKY transcription factors *WRKY12* and *WRKY13* oppositely regulate flowering under short-day conditions. *Molecular Plant* 9:1492-1503
- Ma L, Zhang D, Miao Q, Yang J, Xuan Y, Hu Y (2017) Essential role of sugar transporter *OsSWEET11* during the early stage of rice grain filling. *Plant Cell Physiology* 58:863-873
- Manjarres-Hernández EH, Arias-Moreno DM, Morillo-Coronado AC, Ojeda-Pérez ZZ, Cárdenas-Chaparro A (2021) Phenotypic Characterization of Quinoa (*Chenopodium quinoa* Willd.) for the Selection of Promising Materials for Breeding Programs. *Plants* 10:1339
- Maughan P, Smith S, Rojas-Beltran J, Elzinga D, Raney J, Jellen E, Bonifacio A, Udall J, Fairbanks D (2012) Single nucleotide polymorphism identification, characterization, and linkage mapping in quinoa. *The Plant Genome* 5
- Murphy KM, Matanguihan JB, Fuentes FF, Gómez-Pando LR, Jellen EN, Maughan PJ, Jarvis D (2018) Quinoa breeding and genomics. *Plant Breeding Reviews* 42:257-320
- Ogata T, Toyoshima M, Yamamizo-Oda C, Kobayashi Y, Fujii K, Tanaka K, Tanaka T, Mizukoshi H, Yasui Y, Nagatoshi Y (2021) Virus-Mediated Transient Expression Techniques Enable Functional Genomics Studies and Modulations of Betalain Biosynthesis and Plant Height in Quinoa. *Frontiers in Plant Science* 12
- Ouellette LA, Reid RW, Blanchard SG, Brouwer CR (2018) LinkageMapView—rendering high-resolution linkage and QTL maps. *Bioinformatics* 34:306-307
- Pandey SP, Benstein RM, Wang Y, Schmid M (2021) Epigenetic regulation of temperature responses: past successes and future challenges. *Journal of Experimental Botany* 72:7482-7497
- Patil G, Valliyodan B, Deshmukh R, Prince S, Nicander B, Zhao M, Sonah H, Song L, Lin L, Chaudhary J (2015) Soybean (*Glycine max*) *SWEET* gene family: insights through comparative genomics, transcriptome profiling and whole genome re-sequence analysis. *BMC genomics* 16:1-16
- Patiranage DS, Asare E, Maldonado-Taípe N, Rey E, Emrani N, Tester M, Jung C (2021) Haplotype variations of major flowering time genes in quinoa unveil their role in the adaptation to different environmental conditions. *Plant, Cell & Environment*

- Patirange DS, Rey E, Emrani N, Wellman G, Schmid K, Schmöckel SM, Tester M, Jung C (2020) Genome-wide association study in the pseudocereal quinoa reveals selection pattern typical for crops with a short breeding history. *bioRxiv*
- Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, Kling DE, Gauthier LD, Levy-Moonshine A, Roazen D (2017) Scaling accurate genetic variant discovery to tens of thousands of samples. *BioRxiv*:201178
- Porebski, S., Bailey, L. G., and Baum, B. R. (1997). Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Molecular Biology* 15, 8–15.
- Reuscher S, Furuta T (2016) ABHgenotypeR: Easy Visualization of ABH Genotypes. R Package 1
- Roe JL, Rivin CJ, Sessions RA, Feldmann KA, Zambryski PC (1993) The *Tousled* gene in *A. thaliana* encodes a protein kinase homolog that is required for leaf and flower development. *Cell* 75:939-950
- Saze H, Scheid OM, Paszkowski J (2003) Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. *Nature Genetics* 34:65-69
- Simmonds N (1971) The breeding system of *Chenopodium quinoa* I. Male sterility. *Heredity* 27:73-82
- Singh RK, Chaudhary BD (1977) Biometrical methods in quantitative genetic analysis. *Biometrical methods in quantitative genetic analysis*
- Soppe WJ, Jacobsen SE, Alonso-Blanco C, Jackson JP, Kakutani T, Koornneef M, Peeters A (2000) The late flowering phenotype of *fwa* mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Molecular Cell* 6:791-802
- Sosa-Zuniga V, Brito V, Fuentes F, Steinfort U (2017) Phenological growth stages of quinoa (*Chenopodium quinoa*) based on the BBCH scale. *Annals of Applied Biology* 171:117-124
- Štorchová H (2020) The evolution of the *FLOWERING LOCUS T-Like* (*FTL*) genes in the goosefoot subfamily Chenopodioideae. *Evolutionary biology—A transdisciplinary approach*:325-335
- Swarts K, Li H, Romero-Navarro J, An D, Romay-Alvarez M, Hearne S (2014) FSFHap (Full-Sib Family Haplotype Imputation) and FILLIN (Fast, Inbred Line Library Imputation) optimize genotypic imputation for low-coverage, next-generation sequence data in crop plants. *The Plant Genome* 7:1-12
- Taylor J, Butler D (2017) R package ASMap: efficient genetic linkage map construction and diagnosis. *Journal of Statistical Software* 79:1-29
- Tian J, Keller MP, Broman AT, Kendzierski C, Yandell BS, Attie AD, Broman KW (2016) The dissection of expression quantitative trait locus hotspots. *Genetics* 202:1563-1574
- Tong Z, Jiang S, He W, Chen X, Yin L, Fang D, Hu Y, Jiao F, Zhang C, Zeng J (2021) Construction of high-density genetic map and QTL mapping in *Nicotiana tabacum* backcrossing BC4F3 population using whole-genome sequencing. *Czech Journal of Genetics and Plant Breeding* 57:102-112
- Tyler AL, Lu W, Hendrick JJ, Philip VM, Carter GW (2013) CAPE: an R package for combined analysis of pleiotropy and epistasis. *PLoS Computational Biology* 9:e1003270

Wang L, Sun S, Jin J, Fu D, Yang X, Weng X, Xu C, Li X, Xiao J, Zhang Q (2015) Coordinated regulation of vegetative and reproductive branching in rice. *Proceedings of the National Academy of Sciences* 112:15504-15509

Wu Y, Bhat PR, Close TJ, Lonardi S (2008) Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph. *PLoS Genetics* 4:e1000212

Yuan M, Zhao J, Huang R, Li X, Xiao J, Wang S (2014) Rice *MtN3/saliva/SWEET* gene family: evolution, expression profiling, and sugar transport. *Journal of Integrative Plant Biology* 56:559-570

Zhang T, Gu M, Liu Y, Lv Y, Zhou L, Lu H, Liang S, Bao H, Zhao HJ (2017) Development of novel InDel markers and genetic diversity in *Chenopodium quinoa* through whole-genome re-sequencing. *BMC Genomics* 18:1-15

2.8. Supplementary Data

Supplementary Table 1. Plant material used in QTL mapping. DTF: days to flowering, DTM: days to maturity, PH: plant height, PL: panicle length, PD: panicle density, TKW: thousand kernel weight, SW: seed weight per plant, SN: seed number per plant, SC: Saponin content, MS: Mildew susceptibility.

Supplementary Table 2. Methods for phenotypic evaluation.

Supplementary Table 3. Primers used in this study.

Supplementary Table 4. Polymerase chain reaction (PCR) and agarose gel electrophoresis description.

Supplementary Table 5. Allele and genotype nomenclature used in this study.

Supplementary Table 6. Genetic and phenotypic segregation for two traits in the F₂ and F₃ populations. Red axil pigmentation was determined five weeks after sowing. The InDel marker JAASS5 was described by Zhang et al. (2017). *R*₁ and *R*₂ represent the 189 bp and 164 bp alleles, respectively.

Supplementary Table 7. Complete list of genes located within the non-overlapping QTL confidence intervals.

Supplementary Table 8. Genes within *pleio4.1* and *pleio14.1* confidence intervals with a previously described function related to flowering time, photoperiod response or yield regulation and their gene variants. Gene variants between CHEN-109 (pollen parent) and PI-614889 (seed parent) and the reference genome QQ74_V2 are shown.

Supplementary Figure 1. Flow charts depicting the (A) bioinformatics pipeline and (B) the genetic map construction pipeline. The main steps of the pipelines are shown in solid-line boxes, method/parameters are shown in dashed boxes and the number of markers is shown in red dashed boxes. PE= paired-end.

Supplementary Figure 2. Agarose gel electrophoresis of PCR products from the InDel marker JASS5. (A) Homozygous F₃ plants. From lane 1 to 10, plants are homozygous for allele *R*₁ (parent CHEN-109; 189 bp); and from lane 11 to 20, plants are homozygous for allele *R*₂ (parent PI-614889; 164 bp). (B) Plants from segregating F₃ families. L = middle range DNA ladder, P1 = parent CHEN-109, P2 = parent PI-614889, C = negative control. Agarose gels were run for 60 min at 100 V.

Supplementary Figure 3. Phenotypic variation in the F₂ population. (A) Variation of panicle density. The numbers in yellow represent the scoring scale used for phenotyping. (B)

Variation of plant height illustrated by seven F₂ individuals (A to G) and the parental lines. (C) Variation in days to maturity is illustrated by the different colors of the panicle. PI-614889 had reached the maturity stage and was ready to harvest. CHEN-109 had not reached the seed filling stage. Individual C is at the seed filling stage. All pictures were taken 16 weeks after sowing.

Supplementary Figure 4. SNP densities across 18 quinoa chromosomes. The number of SNP within a 1 Mb window size is shown by different colors. Densities were calculated by the CMplot R package using the raw data (~17 million SNP).

Supplementary Figure 5. Percentage of genome-wide missing data. Black dots show missing markers. Black horizontal lines represent several continuous missing markers. Chromosomes are separated by black vertical lines. The numbers above correspond to the quinoa chromosomes (even numbers: Cq1A to Cq9A; uneven numbers: Cq1B to Cq9B).

Supplementary Figure 6. Data filtering before map construction. (A) Number of markers per individual. (B) Number of genotyped individuals for each marker. (C) Histogram of the proportion of markers for which pairs of individuals have matching genotypes. Dashed red lines show the filtering thresholds. Data below the dashed red line in A and B, and to the right of the dashed line in C was removed.

Supplementary Figure 7. Quinoa linkage map based on 334 plants from an F₂ population derived from a cross between CHEN-109 and PI-614889. The map consists of 133,913 markers arranged in 5,219 bins and it was drawn with the LinkageMapView R package. The numbers above indicate linkage groups (LG) and chromosome numbers. Horizontal blue lines show the location of the first marker of each bin followed by the number of markers in each bin in parenthesis. Marker names are coded as “S” + Chromosome number + “_” + physical position of the marker. Genetic distances in cM are written to the left of each LG.

Supplementary Figure 8. Map density across 21 linkage groups from the F₂ population derived from a cross between CHEN-109 and PI-614889. Densities were recorded by the LinkageMapView R package. Binned markers are shown by horizontal black lines.

Supplementary Figure 9. Quality control of the marker data used for linkage map construction. (A) Marker’s segregation distortion and missing proportion (number of individuals that are missing a marker in a specific locus / 334 individuals) calculated by ASMap R package; different colors correspond to each linkage group. (B) Estimated number of crossovers and double crossovers calculated by ASMap R package; blue dots represent the estimated values for each F₂ individual.

Supplementary Figure 10. Frequencies of the parental alleles calculated from the F₂ population using the program ABHgenotype R package in each of the linkage groups. Different alleles are shown by different colors. Linkage group numbers are shown to the right. The x-axis shows the physical positions of the 133,913 markers according to the reference genome QQ74_V2. Hetero: heterozygous genotype.

Supplementary Figure 11. Comparative QTL analysis to detect pleiotropy for (A) *pleio20.1*, (B) *pleio4.2*, (C) *pleio4.3* and (D) *pleio7.1*. Two tests were performed: one vs. two QTL (to the left) and one vs. “n” QTL (to the right). The black curve is the LOD score curve for the single-QTL model, with the estimated QTL location indicated by a black triangle. The blue and pink curves are profile LOD score curves for the two-QTL model. Dots indicate the LOD score for the traits considering a single-QTL model. DTF: days to flowering, DTM: days to maturity, PH: plant height, PL: panicle length, PD: panicle density, SN: seed number per plant, SW: seed weight per plant, TKW: thousand kernel weight.

Supplementary Figure 12. Evaluation of variant haplotypes using the sequences of 328 F₂ individuals and the corresponding phenotypic data of 328 F₃ families. Phenotypic effects of haplotype variations within several candidate genes are shown: *DEXH-BOX ATP-DEPENDENT RNA HELICASE DEXH10 (HEN2)*, *SUPPRESSOR OF FRI 4 (SUF4)*, *RCC1 DOMAIN-CONTAINING PROTEIN 3 (RUG3)*, *WRKY TRANSCRIPTION FACTOR 13 (WRKY13)* and *TSL-KINASE INTERACTING PROTEIN 1 (TK11)*. The variants genotypes correspond to, for instance, *A₁A₁* (our homozygous parent PI-614889), *A₁A₂* (heterozygous), *A₂A₂* (our homozygous parent CHEN-109) and are described in Table S4. Significant differences between genotypes are shown by asterisks (t-test, $\alpha < 0.05 = ***$, $\alpha < 0.01 = **$, $\alpha < 0.001 = ***$). DTF: days to flowering, TKW: thousand kernel weight, PH: plant height.

Supplementary Figure 13. Evaluation of variant haplotypes using available whole-genome sequencing and phenotypic data of 310 quinoa accessions. Phenotypic effects of haplotype variations within several candidate genes are shown: *ATHB-15/CORONA*, *DEXH-BOX ATP-DEPENDENT RNA HELICASE DEXH10 (HEN2)*, *NRT1/ PTR 2.6 (NPF2.6)*, *RCC1 DOMAIN-CONTAINING PROTEIN 3 (RUG3)*, *WRKY TRANSCRIPTION FACTOR 13 (WRKY13)*, *ETHYLENE-RESPONSIVE TRANSCRIPTION FACTOR 113 (ERF113)* and *FLOWERING LOCUS D (FLD)*. The variants genotypes correspond to, for instance, *H₁H₁* (our homozygous parent PI-614889), *H₁H₂* (heterozygous), *H₂H₂* (our homozygous parent CHEN-109) and are described in Table S4. Significant differences between genotypes are shown by asterisks (t-test, $\alpha < 0.05 = ***$, $\alpha < 0.01 = **$, $\alpha < 0.001 = ***$). Phenotypic data of different years are shown in different colors. DTF: days to flowering, TKW: thousand kernel weight.

3. Leaf and shoot apical meristem transcriptomes of quinoa (*Chenopodium quinoa*) in response to photoperiod and plant development

3.1. Abstract

In this study, I aimed to identify candidate genes for flowering time regulation and photoperiod response in quinoa. For this purpose, I used two accessions with contrasting life cycles, which were grown under two photoperiod regimes (short- and long-days). I examined the histology of the shoot apical meristem at different stages of development, and analyzed the leaf and shoot apical meristem transcriptomes. I investigated four groups of differentially expressed genes, (i) genes differentially responding to photoperiod, (ii) genes regulating flowering time under long-day conditions, (iii) genes regulating flowering time under short-day conditions, and (iv) genes controlling floral transition within the shoot apical meristem. My results confirm the fundamental claim that two distinct pathways regulate flowering time under short- and long-day conditions and that two quinoa *FT* orthologues are key components of these pathways.

3.2. Introduction

The pseudocereal species quinoa (*Chenopodium quinoa* Willd.) is an allotetraploid species ($2n = 4x = 36$) with a genome size of 1.45-1.50 Gb (Jarvis et al., 2017). It originated from the Andean region of South America and, in the past decade, it has gained increasing attention as a human diet due to its favorable nutritional properties. Quinoa has a high protein content and an exceptional balance between oil, protein and fat content (Granado-Rodríguez et al., 2021; Iqbal et al., 2020). Quinoa is tolerant against frost, drought, and salinity (Shi & Gu, 2020). Because of these advantages, quinoa is an interesting alternative to improve food security, and research and production of this crop have been started in over 120 countries in the world (Alandia et al., 2020).

However, the short-day nature of this crop remains an obstacle to cultivation in temperate regions and high latitudes of Europe, North America, and China (Murphy et al., 2018; Patirange et al., 2020). A better understanding of the genetic mechanisms underlying floral transition, including photoperiodic regulation of flowering time, will help addressing this problem. By using this knowledge, breeding programs aiming to adapt quinoa to new environments can benefit from the profound effect of flowering time control on crop adaptation and yield potential (Gaudinier & Blackman, 2020).

Fuller (1949) gave the first report on quinoa's response to photoperiod. Since then, only a few projects have worked on the genetic mechanisms for the regulation of flowering time in quinoa. Golic et al. (2020) used a computational approach to identify hundreds of putative orthologues of *Arabidopsis thaliana* flowering time genes. They found overall conservation of flowering pathways, except for the vernalization pathway, where there was a lack of orthologues of key genes in quinoa. In another study, long non-coding RNA (lncRNA) was analyzed under short-day (SD) conditions, including a night-break (NB) of 60 min (Wu et al., 2021). Twenty-four lncRNA involved in flowering time regulation were identified, some of which targeted *FLOWERING LOCUS T (FT)* and *TWIN SISTER of FT (TSF)* homologs (*in silico* analysis). Besides, they found that quinoa homologs of *CONSTANS-like (COL)*, *LATE ELONGATED HYPOCOTYL (LHY)*, *EARLY FLOWERING 3 (ELF3)*, and *ELONGATED HYPOCOTYL 5 (HY5)* were down-regulated after NB, while homologs of *PHYTOCROME A (PHYA)* and *CRYPTOCHROME1 (CRY1)* were up-regulated after NB. Furthermore, the role of *FT* and *COL* homologs in the adaptation of quinoa to different day-length conditions was highlighted by Patirange et al. (2021). In this study, the haplotypes of 12 genes were analyzed using a set of 276 accessions grown under long-day (LD) and SD conditions. As a

result, *CqFT1A*, *CqCOL2B*, *CqCOL4A-1*, and *CqCOL5B* were associated with flowering time variation under LD conditions, but not under SD. The regulation of flowering time has also been studied in closely related species. In *Chenopodium rubrum*, *FLOWERING LOCUS T-LIKE 1* (*CrFTL1*) is SD-inducible and inhibited by an NB of 15 min, whereas *CrFTL2* is constitutively expressed (Cháb et al., 2008). Moreover, two *CrCOL* genes were down-regulated during the light period regardless of the length of the preceding dark period. The same behavior was observed for activation of the floral promoter *CrFTL1* (Drabešová et al., 2014). *Chenopodium ficifolium* is the most likely donor of the quinoa B genome (Walsh et al., 2015). In a segregating F₂ generation of this species, marker-trait associations with flowering time, plant height, and the number of branches were identified for an ortholog of *FTL1* (Subedi, Neff, & Davis, 2021).

The quinoa reference genome (QQ74_V1, URL: <https://www.cbrc.kaust.edu.sa/chenopodiumdb>) encompasses 44,776 gene models (Jarvis et al., 2017). Several projects have already taken advantage of this resource. A root transcriptome study unveiled 117 genes, which likely control salt stress responses (Shi & Gu, 2020). Another transcriptome study unraveled expression profiles of ~1,400 genes expressed during germination (Wu et al., 2020). Wu et al. (2019) analyzed inflorescence transcriptomes, which resulted in the identification of ~1,900 candidate genes possibly regulating inflorescence architecture. A recent study attempted to identify the effect of NB in quinoa flowering under SD conditions. Wu et al. (2021) analyzed leaf transcriptomes of two quinoa cultivars before and after NB. However, this study did not identify genes regulating floral transition because the investigated cultivars did not differ in their time to flowering.

I aimed to determine the time point when the shoot apical meristem (SAM) is turned into a floral meristem. Furthermore, I wanted to unravel genes controlling flowering time and responding to photoperiod. I expected that those genes display differential expression profiles between different photoperiods and between accessions with contrasting life cycle regimes. Accordingly, I analyzed the development of the SAM in two accessions differentially responding to day length. Then, I studied leaf and SAM transcriptomes from plants grown under two photoperiod conditions (SD and LD). I identified thousands of differentially expressed genes (DEGs). I focused on photoperiod-responsive genes and genes with contrasting expression profiles between the accessions. My study provides new insights into flowering time and photoperiod regulation in quinoa.

3.3. Materials and Methods

3.3.1. Plant material and growth conditions

I investigated two quinoa accessions in this study, D-12082 from Peru (seed code: 182301) and PI-614886 from Chile (reference genome QQ74, seed code: 182283), which were differentially flowering under SD and LD conditions as reported by Patiranage et al. (2021). Plants were grown in two experiments (LD and SD). One hundred five plants of each accession were grown in 3 × 3 cm 35-multiwell palettes (Hermann Meyer KG, Germany) in a growth chamber under long-day conditions (LD) (22 °C and 16 h light; 900 μmol·m⁻²·s⁻¹, Son-T Agro 400 W, Koninklijke Philips Electronics N.V., Eindhoven, The Netherlands). Likewise, 105 plants of each accession were grown under short-day conditions (SD) (22 °C and 8 h light; 900 μmol·m⁻²·s⁻¹).

3.3.2. Phenotyping and histological analysis

In each experiment (LD and SD), I phenotyped ten plants per accession for days to bolting (days until the floral bud is visible) and days to flowering (days until the first flower opens) (Supplementary Table 9 and Supplementary Figure 14). For histological analysis, I sampled seven apices weekly until plants reached the bolting stage. Apices were fixed in 4.0% FAA (4.0% formaldehyde, 50.0% ethanol and 5.0% acetic acid) overnight. Then, the samples were

dehydrated by an ethanol series and embedded in Paraplast by a standard protocol. Apices were sectioned at 8 μm , using a Leica rotational microtome, and stained with 0.05% toluidine blue.

3.3.3. Sample preparation and RNA isolation

From each experiment, I took leaf samples from seven plants (A and B samples) starting 14 days after sowing and continuing weekly until plants reached the flowering stage. Shoot apical meristems (SAMs) were sampled weekly, starting 14 days after sowing until the bolting stage. SAMs were dissected under a binocular microscope using a scalpel and forceps for sampling. Ten to 14 SAMs were pooled to make one biological replicate. I harvested leaves at ZT-9 (Zeitgeber-time) and SAMs at ZT-9 to ZT-12. Leaf and SAM samples were immediately placed in liquid nitrogen.

Then, I isolated RNA from 72 samples (48 leaf and 24 SAM samples.) Leaf samples corresponded to three biological replicates of the two accessions under SD and LD conditions at two (W2), three (W3), four (W4), and five (W5) weeks after sowing. The SAM samples corresponded to three biological replicates of the two accessions under SD and LD conditions at two (W2) and three (W3) weeks after sowing. RNA isolation and DNase treatment were carried out according to the instruction provided by the peqGold Total RNA Kit (PeqLab, Erlangen, Germany).

3.3.4. Transcriptome analysis

Seventy-two DNase-treated RNA samples (RIN number >6.9) were sequenced using NovaSeq PE150. Library preparation and RNA sequencing were carried out at Novogene Company Limited (Cambridge, United Kingdom). Directional mRNA library preparation was carried out by the Poly(A) enrichment method. To obtain read-count data for each transcript from the raw reads, the protocol described by Pertea et al. (2016) was followed. In brief, reads from each sample were mapped to the reference genome and gene annotation of *C. quinoa* cv. QQ74_V2 (CoGe Genome ID: id60716) with HISAT2 (v.2.1.9) (Kim, Langmead, & Salzberg, 2015). Then, novel and known transcripts were assembled for each sample and the resulting annotations were merged into one comprehensive annotation using StringTie (v.2.2.0) (Pertea et al., 2015). The transcriptome was then retrieved from the merged annotation using the gffread function of GFF utilities (v.0.11.1) and read counts were obtained by mapping reads back to the transcripts using bowtie2 (v.2.3.5) leveraged by rsem (v.1.3.1) with ‘very_sensitive’ stringency level. Next, I extracted ‘expected counts’ from each sample to build the gene matrix containing both known and newly-assembled transcripts. Newly-assembled transcripts are those sequences assembled as transcripts by StringTie, which are not found as transcripts (annotated genes) in the reference genome QQ74_V2. They include putative gene isoforms as well as artifacts from the algorithm. The differential expression analysis was then carried out at the gene level using R (version 4.1.1) package edgeR (version 3.36.0) (Chen et al., 2020). First, I filtered out lowly expressed genes defined as those with less than five overlapping mapped reads (min.count=5) and normalized the gene expression to take library composition into account (TMM = Treamead mean method). Later, as quality control, I performed unsupervised samples clustering of transcriptome data of all samples and also of only the leaf samples separately. I constructed MDS plots to identify which dimension had the most significant effect, as well as to verify that biological replicates clustered together. I used the Trimmed Mean log-ratios (\log_2 fold change of M values) (Robinson & Oshlack, 2010). Next, I used the dgeFitGLM function to fit a generalized linear model to the count data and conduct gene-wise statistical tests for my contrasts. I defined a differentially expressed gene (DEG) as the one with a false discovery rate of < 0.05 and a \log_2 fold change (FC) ≥ 0.25 or ≤ -0.25 .

In a first analysis, I investigated the DEGs of transcripts annotated in the reference genome QQ74. For this, I examined my leaf transcriptome data in four main steps. First, I looked for genes differentially responding to photoperiod to regulate flowering time (i). With this aim, I used the contrasts between SD and LD conditions in D-12082 and investigated the spatial expression pattern of the DEGs for the two accessions in leaves (DEGs between W2-W5, W3-W5, and W4-W5). Second, I looked for genes regulating flowering time under LD conditions (ii). For this purpose, I used the contrasts between D-12082 and PI-614886 under LD in leaves. Third, I searched for genes regulating the time to flower under SD conditions (iii). With this aim, I again investigated the spatial expression pattern of the DEGs for the two accessions in leaves (DEGs between W2-W5, W3-W5, and W4-W5). Finally, I analyzed the DEGs from the SAM to identify flowering time candidate genes in this tissue (iv). For this purpose, I used the contrasts between D-12082 and PI-614886 under LD at the SAM. I performed a Gene Ontology (GO) analysis of the resulting DEGs of each of the four analyses (i, ii, iii, and iv). For this, I used the NCBI Basic Local Alignment Search Tool (BLAST) service to perform blastx (Taxid: 58023; E-value: 1.0E-3, number of Blasts Hits: 3, word size: 6). Then, I used Blast2GO (version 6.0) to map and annotate GO terms (annotation cut-off: 55, GO-weight: 5) (Götz et al., 2008). Later, I performed Enzyme Code annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) mapping with the Blast2GO tool (Kanehisa & Goto, 2000).

In a second analysis, I used co-expression networks to associate newly-assembled transcripts to a putative biological process based on the annotation provided for the annotated genes within the same co-expression cluster. For this purpose, I first performed leaf transcriptome data investigation in the same way as described for annotated genes (i, ii, and iii). Later, I combined the results from the newly-assembled DEGs with the results from the annotated DEGs. Then, I used the R package BioNERO (Almeida-Silva & Venancio, 2022) to perform co-expression analysis in the set of data containing annotated and newly-assembled DEGs. The Weighted Gene Co-expression Network Analysis (WGCNA) algorithm was used (Langfelder & Horvath, 2008) and pairwise correlations were computed using Pearson correlation analysis. Later, I constructed co-expression networks for each of the modules resulting from the previous step. Then, for each module, I selected the top 10% of newly-assembled transcripts with the highest degree and with module membership (MM) > 0.8. Degree is defined as the sum of connection weights of a gene to all other genes in the module, while MM is the correlation of a gene to its module eigengene (Almeida-Silva & Venancio, 2022). Following, I blasted and GO annotated the selected newly-assembled transcripts. For annotation, I used the NCBI Basic Local Alignment Search Tool (BLAST) service to perform blastx (Taxid: 58023; E-value: 1.0E-3, number of Blast Hits: 3, word size: 6). Then, I used Blast2GO (version 6.0) to map and annotate GO terms (annotation cut-off: 55, GO-weight: 5).

3.3.5. Real-time qPCR

DNase-treated RNA was reverse transcribed with the First Strand cDNA Synthesis Kit (Thermo Fisher Scientific Inc., Waltham, United States), according to the manufacturer's instructions. Then, I performed RT-qPCR by Bio-Rad CFX96 Real-Time System, which has a built-in Bio-Rad C1000 Thermal Cycler (Bio-Rad Laboratories GmbH, Munich, Germany). Every PCR reaction had a total volume of 20 µl: 10 µl of Platinum SYBR Green qPCR SuperMix-UDG with ROX (Invitrogen by Life Technologies GmbH, Darmstadt, Germany), 1 µl of each forward and reverse primer (10 µg), 2 µl of diluted cDNA (1:20 dilution) and 6 µl of ddH₂O. The amplification conditions were: 95°C for 3 min as initial denaturation and 40 cycles of 10 s at 95°C, 20 s at primer pair annealing temperature and 30 s at 72°C. Three technical replicates per sample and negative control (water) were also loaded. Cq values were obtained by setting the baseline threshold to 100 and expression levels were calculated by the comparative $\Delta\Delta C_t$ method (Livak & Schmittgen, 2001). Primers used for gene expression analysis are listed in Supplementary Table 3. *CqPTB* and *CqIDH-A*, which had been identified

in a previous study as suitable housekeeping genes for RT-qPCR analysis in quinoa, were used as reference genes (Maldonado-Taipe, Patirange, Schmöckel, Jung, & Emrani, 2021). Following, to validate the results obtained by the bioinformatics pipeline, I calculated the Pearson's R correlation between the differential expression results from the transcriptome bioinformatics pipeline and the results from RT-qPCR. For this calculation, I used the \log_2 FC (fold change) values calculated for different conditions, accessions, and developmental stages.

3.4. Results

3.4.1. The transition of the shoot apical meristem into a floral meristem

I grew two accessions with contrasting life cycle regimes in a growth chamber under SD and LD conditions. I phenotyped flowering time as the first flower opens and as bolting. When phenotyping flowering as the first flower opens, the accession PI-614886 flowered much earlier under both SD (seven days earlier) and LD (18 days earlier) conditions, as expected (Table 7). However, phenotyping for bolting revealed that, under SD conditions, both accessions transit to flowering at the same time. Therefore, I analyzed longitudinal sections of apices under the microscope to determine the precise time when the SAM turned into a floral meristem.

Table 7. Results from growth experiments with two accessions. D-12082 is a short-day accession from Peru and PI-614886 is a day-neutral accession from Chile. Plants of each accession were grown in a growth chamber at 22°C under long-day (LD, 16 h light) and short-day conditions (SD, 8 h light); das: days after sowing; SAM: shoot apical meristem.

Accession	Photoperiod	Floral bud appearance (das)	First flower opens (das)	Floral transition based on SAM morphology (das)
D-12082	LD	46±2	74±3	28
	SD	36±2	56±3	21
PI-614886	LD	31±3	46±3	21
	SD	34±2	49±3	21

I dissected apices over six weeks, beginning with the second week after sowing. I determined the time of floral transition when a “dome shape” at the SAM became visible (apical dominance release) (Dun, Ferguson, & Beveridge, 2006). The first flower opened 25 (PI-614886, SD) to 46 days (D-12082, LD) after transitioning the SAM into a floral meristem (Table 7). The floral transition was strikingly different between both accessions under LD. The accession PI-614886 already underwent floral transition after three weeks under both SD and LD conditions. In contrast, it took one more week to observe a floral meristem in the accession D-12082 under LD than under SD conditions (Figure 9). The histological observations were in line with the plant growth because PI-614886 as a day-neutral genotype did not respond to varying day lengths, whereas D-12082 showed a delay in transition to flowering under LD. Interestingly, morphological changes indicating the change of the SAM into a floral meristem were observed when the plants were at a relatively early stage of development, between BBCH11 (first pair of leaves visible) and BBCH12 (second pair of leaves visible) (Sosa-Zuniga, Brito, Fuentes, & Steinfort, 2017).

3.4.2. Transcriptome sequencing

I sequenced 48 leaf and 24 SAM RNA libraries from plants grown under SD and LD conditions. ~68.66 million reads per library were generated with an average read length of 150 bp amounting to 20.59 Gbp per library (Supplementary Table 10). I found 132,965 gene models, of which 85,634 leaf gene models (64.40%) and SAM 74,640 gene models (56.14%) were kept after excluding the non-expressed genes (those with less than five overlapping mapped reads).

As a quality control, I performed unsupervised samples clustering with the filtered and normalized transcriptome data from the leaf and SAM tissues. As expected, I observed that the samples were grouped by tissue (19.0% of contribution to the differences in expression given in logFC) and by accession (13.0% of contribution to the differences in expression given in logFC) (Supplementary Figure 15). Then, to investigate the leaf transcriptome data more closely, I conducted unsupervised clustering using only leaf data. Following my expectations, I observed that the samples belonging to the same accessions, and the biological replicates corresponding to the same tissue/treatment clustered together (Supplementary Figure 15).

3.4.3. Identification of differentially expressed genes regulating flowering time

I first analyzed the transcripts annotated in the reference genome QQ74. I aimed to identify gene expression profiles that could be correlated to the morphological changes at the meristem. I focused on W2 and W3 developmental stages because my study of the SAM at the histological level showed that floral transition occurred at those stages. I analyzed my leaf transcriptome data in four main steps. First, I looked for genes differentially responding to photoperiod to regulate flowering time (i). Second, I looked for genes regulating flowering time under LD conditions (ii). Third, I searched for genes regulating the time to flower under SD conditions (iii). Finally, I analyzed the DEGs from the SAM to identify flowering time candidate genes in this tissue (iv).

To discover genes differentially responding to photoperiod (i), I investigated the DEGs between SD and LD conditions at W2 and W3 in D-12082. I focused on those stages and that accession because the morphological changes at the SAM showed differential transition to a generative SAM under LD and SD conditions. As an outcome, 3,769 genes were differentially expressed between SD and LD (Supplementary Figure 16A). However, I performed further filtering of the resulting DEGs based on the morphological changes at the SAM. First, I expected that a gene that differentially responds to photoperiod to regulate flowering time was differentially expressed between photoperiod regimes at W3 in D-12082, while at W2 the differential expression might or might not be present yet. Therefore, from the 3,769 genes, I kept 1,799 DEGs that were down- or up-regulated in W3 regardless of their status in W2 (Supplementary Figure 16B). Besides, given the morphological differences observed in my histological analysis, the differential response to photoperiod of my candidate genes must be different between W2 and W3. I used the criteria $-0.25 < (\log_2 \text{FC at W2} - \log_2 \text{FC at W3}) < 0.25$ and filtered out genes whose differential expression did not differ between W2 and W3. As result, I got 1,505 genes differentially responding to photoperiod, which likely regulate flowering time in D-12082 (Supplementary Figure 16C).

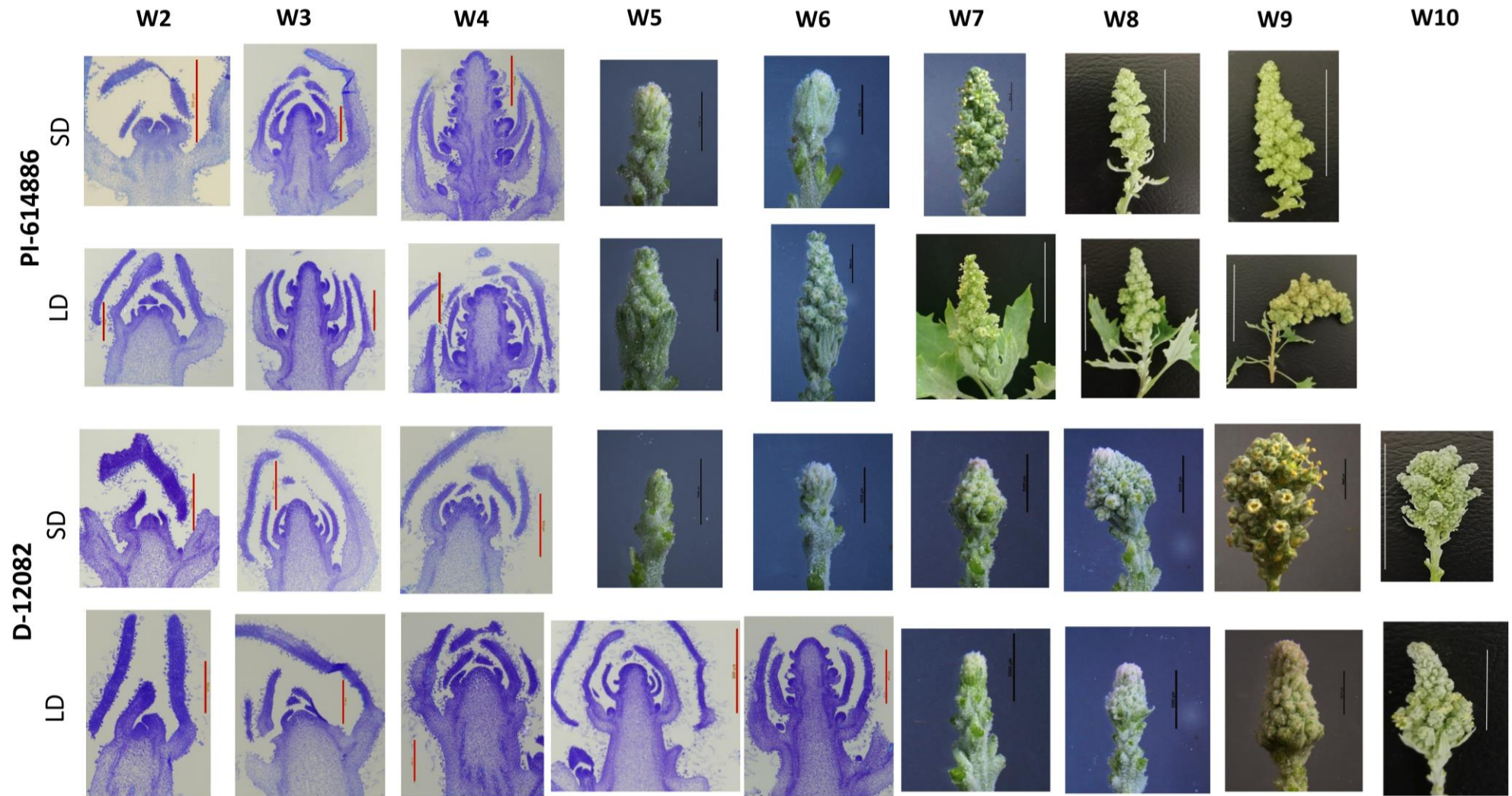


Figure 9. Shoot apical meristem (SAM) and inflorescence development from two quinoa genotypes. PI-614886 and D-12082 are day-neutral and short-day accessions, respectively. The SAM longitudinal sections were stained with toluidine blue. Plants were grown in a growth chamber at 22°C under long-day conditions (LD, 16 h light), and short-day conditions (SD, 8 h light). Photos were taken two to 10 weeks after sowing (W2-W10). A vegetative SAM was identified by an “edged” shape (e.g. W2 accession D-12082, LD), whereas a reproductive meristem was identified by a “dome shape” (e.g. W4 accession D-12082, LD). Scale bars: 500 μ m (red), 5000 μ m (black), 37.5 mm (white).

Based on the morphological changes at the SAM, I expected that a gene differentially responding to photoperiod to regulate flowering time shows different spatial expression patterns between SD and LD in D-12082 but not in PI-618886. Hence, I investigated the spatial expression pattern of the DEGs in both accessions by obtaining the DEGs between W2-W5, W3-W5, and W4-W5 (Supplementary Figure 16D, Supplementary Figure 16E). Then, I removed the resulting DEGs in the day-neutral accession (3,069 genes) from the resulting DEGs for the short-day accession (2,439 genes). As result, I found 1,527 genes with a spatial pattern of a gene that differentially responds to photoperiod in D-12082 and likely control the time to flower (Supplementary Figure 16F). I argued that genes identified based on their spatial expression pattern must also be DEGs between SD and LD in the short-day accession to be the genes that differentially respond to photoperiod to control flowering time. Therefore, as a final step, I selected the genes that overlapped between the 1,505 DEGs obtained in a previous step (DEGs between SD and LD conditions) and the 1,527 genes identified based on their spatial expression pattern. As an outcome, I obtained 222 DEGs that differentially respond to photoperiod to putatively control flowering time in quinoa (Figure 10A). Remarkably, as expected from my SAM histological analysis, most of the candidate DEGs did not show a differential expression after W4, when flowering had already taken place. Moreover, among the resulting DEGs, I found several genes with a known function in photoperiodic regulation of flowering time in other species like *TREHALOSE-PHOSPHATASE/SYNTHASE 9* (*CqTPS9*), *PHYTOCHROME B* (*CqPHYB*), *ZEITLUPE* (*CqZTL*), and *BLUE-LIGHT INHIBITOR OF CRYPTOCHROMES* (*CqBIC1*) (Figure 10A and Supplementary Table 11) (Blázquez & Weigel, 1999; Más, Kim, Somers, & Kay, 2003; Tian et al., 2021; Wang et al., 2016).

To continue, I searched for genes expressed in the leaves, which could regulate flowering time under LD conditions (ii). These genes might or not be downstream targets of the DEGs differentially responding to the photoperiod, which were found in my previous analysis. I assumed that putative flowering time genes should fulfill the following criteria. First, down- or up-regulated between accessions either at W2 or W3: second, the degree of this up- or down-regulation had to be different between W2 and W3 ($-0.25 > (\log_2 \text{FC W2} - \log_2 \text{FC W3}) > 0.25$). I assumed that the main flowering time integrators are the same for the day-neutral and the short-day accessions. Accordingly, I investigated the DEGs between PI-614886 and D-12082 under LD at W2 and W3. From 2,687 DEGs (Supplementary Figure 17A), 1,812 followed my criteria as putative flowering time genes under LD conditions. The obtained down- and up-regulated genes at LD.W2 (long-day conditions at week 2) could be putative flowering repressors and promoters, respectively (Figure 10B). Within the obtained DEGs, I found several genes previously reported as involved in flowering time regulation in quinoa and other species: *FLOWERING LOCUS T 1* (*CqFT1A*), *FLOWERING LOCUS T 2* (*CqFT2B*), *HEADING DATE 3A* (*CqHD3AB*), *CONSTANS-LIKE 16* (*CqCOL16*), *CONSTANS-LIKE 4* (*CqCOL4*), *FRIGIDA INTERACTING PROTEIN* (*CqFIP1*), *FRIGIDA* (*CqFRL4A*), *EARLY FLOWERING 5* (*CqELF5*) (Figure 10B) (Choi et al., 2011; Kim et al., 2008; Komiyama, Ikegami, Tamaki, Yokoi, & Shimamoto, 2008; Patiranage et al., 2021; Pin et al., 2010; Suarez-Lopez et al., 2001).

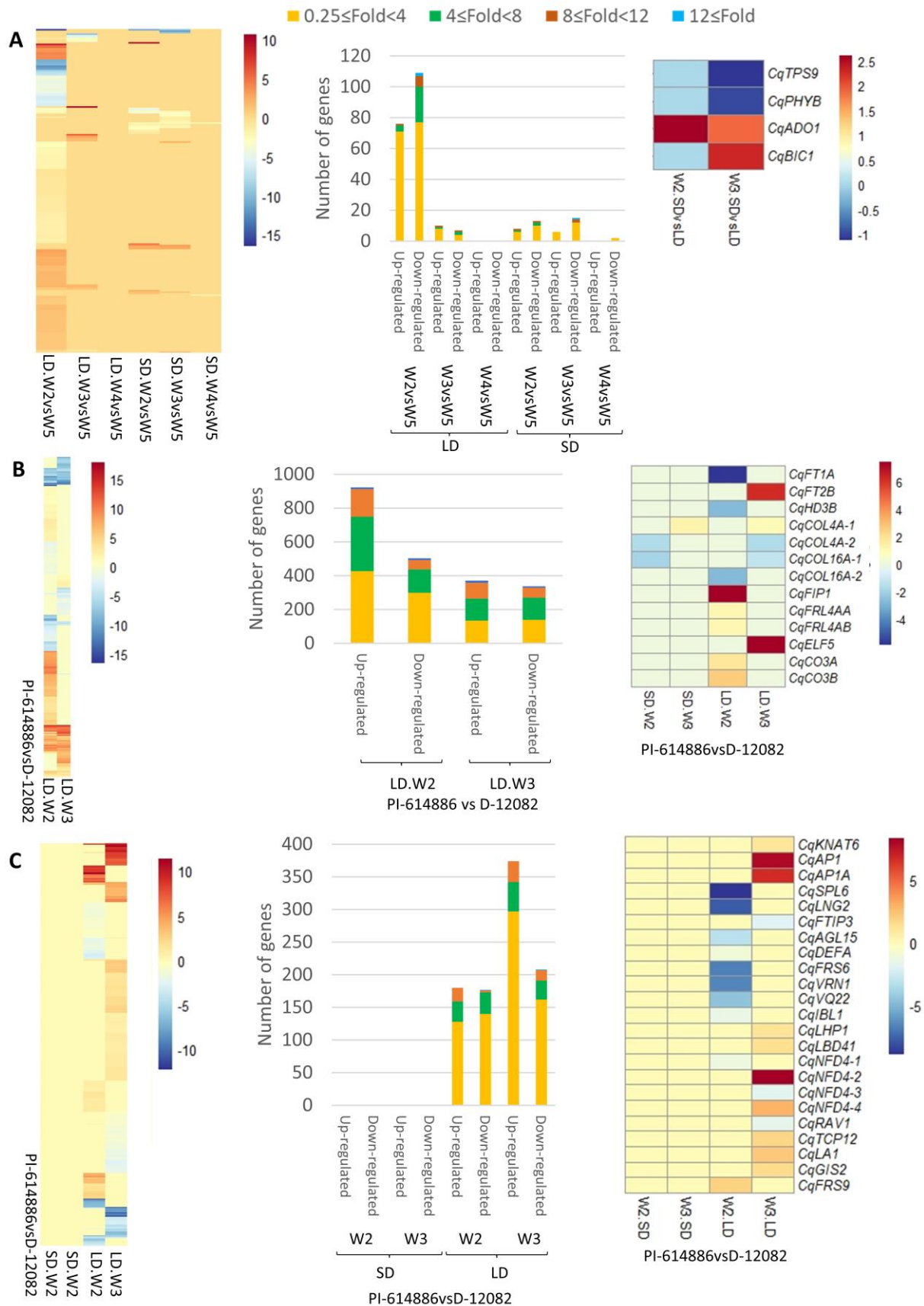


Figure 10. Comparative gene expression between accessions, growth conditions, and developmental stages. Plants from the accessions PI-614886 and D-12082 were grown under long-day conditions (LD, 16 h light), and short-day conditions (SD, 8 h light). RNA was isolated from leaf and shoot apical meristem (SAM) tissue. (A) Heatmap of the 222 differentially expressed genes (DEGs) between SD and LD in D-12082, which differentially respond to photoperiod to likely regulate flowering time. DEGs between W2-W5, W3-W5, and W4-W5 under SD and LD conditions are shown. (B) Heatmap of the 1,812 DEGs between PI-614886 and D-12082 at W2 and W3 under LD, which is putatively controlling flowering under LD. (C) Heatmap of the 911 DEGs at SAM between PI-614886 and D-12082 at W2 and W3, which are likely involved in the regulation of the time to flower

at the SAM. The number of up- and down-regulated genes in each comparison is shown by column graphs to the right of the heatmaps. Comparative expression profiles of selected genes with a known flowering time regulation function in other species are shown to the right of the column graphs. LD.W2.PI-614886vsD-12082 corresponds, for instance, to \log_2 (normalized reads in PI-614886/ normalized reads in D-12082). W=week. Heatmaps were constructed with the pheatmap R package.

In the next step, I searched for those genes, expressed in the leaves, likely regulating the time to flower under SD conditions (iii). Given that morphological changes at the SAM indicated that both accessions transit to flowering at the same time under SD, I focused on the spatial expression pattern of the DEGs to find the genes involved in flowering time regulation under this photoperiod regime. Accordingly, I examined once more the DEGs between W2-W5, W3-W5, and W4-W5. In these genes, I searched for the DEGs having similar spatial expression patterns under SD conditions between the day-neutral accession and the short-day accession, which showed similar morphological changes at the SAM at W2 and W3 (Supplementary Figure 17B). Among the 57 resulting genes, I found *SUPPRESSOR OF CONSTANS OVEREXPRESSION 1* (*CqSOC1*), which, in *Arabidopsis* and other species, integrates multiple flowering signals derived from the age-dependent and gibberellin pathways (J. Lee & Lee, 2010). Interestingly, I found only three genes in common between the DEGs identified as putative flowering time regulators under LD (1,812) and those identified under SD (57 genes).

To continue, I searched for the genes at SAM regulating flowering time (iv). I set two requirements for these genes to be involved in flowering regulation based on my SAM histological analysis. First, they must be differentially expressed between accessions under LD at W2 and W3 when the floral transition took place at the SAM. Second, given the morphological differences observed in the sections of the SAM, up- or down-regulation should differ between W2 and W3 ($-0.25 > (\log_2 \text{FC W2} - \log_2 \text{FC W3}) > 0.25$) under LD. I found 988 genes fulfilling my first requirement (Supplementary Figure 18) and 911 genes fulfilling both requirements (Figure 10C). As expected, these DEGs showed differential expression between the accessions under LD conditions but not under SD, where both accessions flowered simultaneously. Interestingly, among the 911 genes, which putatively regulate flowering time in quinoa at the SAM, I found examples of homologs of known regulators in other species (Supplementary Table 11). For instance, a quinoa homolog of *AGAMOUS-LIKE 15* (*CqAGL15*), a flowering repressor in *Arabidopsis* (Adamczyk, Lehti-Shiu, & Fernandez, 2007), was down-regulated in PI-614886 compared to D-12082 at W2 under LD conditions (Figure 10C). On the other hand, a quinoa homolog of *APETALA1* (*CqAPI1*), a flowering promoter in soybean and wheat (Chen et al., 2020; Murai, Miyamae, Kato, Takumi, & Ogiwara, 2003), was up-regulated in PI-614886 at W3 under LD conditions. Moreover, a homolog of *FT INTERACTING PROTEIN 3* (*CqFTIP3*), required for maintenance of inflorescence SAMs in *Arabidopsis* (Liu et al., 2018), was down-regulated in PI-614886 compared to D-12082 at W3.

3.4.4. Functional annotation and classification of the DEGs

I performed GO analysis for the four different groups of DEGs identified in this study (i, ii, iii, and iv). The GO analysis of the 222 DEGs which differentially respond to photoperiod to likely control flowering time (i) revealed the GO terms “Regulation of gene expression” and “Transcription, DNA-templated” (GO terms: 0010468 and 0006351, respectively) under the Biological Process category (Supplementary Table 12). Then, to further understand the gene functions involved in the metabolic pathways, I annotated the 222 DEGs responding to photoperiod to regulate flowering time against the KEGG database. Fifty-six genes (25.22%) were mapped into 38 signaling pathways. From these 38 pathways, glycine, serine, and threonine metabolism (7.14%) and glycolysis/gluconeogenesis (5.36%) were among the most enriched ones (Supplementary Table 13).

GO analysis of the 1,812 DEGs likely regulating flowering time under LD conditions (ii) (Supplementary Table 14) showed “Regulation of Gene Expression” GO term under the Biological Process category and “Binding” as one of the main GO terms under the Molecular Function category. Besides, glycine, serine, and threonine metabolism (3.70%) and pyruvate metabolism (4.63%) were found to be among the most enriched pathways based on KEGG annotation (Supplementary Table 15). Following this, I carried out GO analysis of the 57 DEGs (iii), which would regulate flowering time under SD conditions in quinoa (Supplementary Table 16). The functional annotation of these DEGs against the KEGG database showed that 13 genes (22.80%) were coding for enzymes acting in 12 different pathways (Supplementary Table 17). Finally, I performed GO analysis of the DEGs likely regulating flowering time at the SAM (iv) (Supplementary Table 18). In the SAM, 74 DEGs (8.26%) were mapped to the KEGG database (Supplementary Table 19). Importantly, GO terms such as “Transcription regulator activity”, “Positive regulation of nucleic acid-templated transcription” and “Regulation of gene expression” were found anew for DEGs putatively regulating flowering time under SD conditions and the DEGs at SAM tissue.

3.4.5. Co-expression analysis of the DEGs

In a second approach, I used the newly-assembled transcripts to perform co-expression analysis with the genes annotated in the reference genome QQ74. I used co-expression networks to associate newly-assembled transcripts to a putative biological process based on the annotation provided for the annotated genes within the same co-expression cluster. With this aim, I first used the newly-assembled transcripts to perform leaf transcriptome data dissection in the same way as described for the annotated genes. Thus, I obtained three groups of newly-assembled DEGs: (i) those differentially responding to photoperiod, (ii) those likely regulating flowering time under LD, and (iii) those putatively controlling the time to flower under SD. Later, I combined the results from the newly-assembled DEGs with the results from the annotated DEGs. As result, I grouped 89 newly-assembled DEGs differentially responding to photoperiod to control flowering time with the DEGs described in Figure 10A, which are annotated DEGs differentially responding to photoperiod (data set 1). I also grouped 1,777 newly-assembled DEGs that might regulate flowering time under LD with the annotated DEGs that might have the same role (Figure 10B) (data set 2). Besides, I grouped 127 newly-assembled DEGs that might regulate flowering time under SD conditions with the 57 DEGs described in Supplementary Figure 17B, which are annotated DEGs putatively regulating flowering time under SD (data set 3). To continue, I performed co-expression analysis in data sets 1, 2, and 3. As an outcome, I identified six co-expressed modules in data set 1 (Figure 11A). The red module contained the highest number of DEGs (88 DEGs) (Supplementary Figure 19). Besides, I identified 20 and three co-expressed modules in data set 2 and 3, respectively (Figure 11B and Figure 11C). The highest number of DEGs harbored by a module was 1,957 in data set 2 (grey module) and 72 (grey module) in data set 3 (Supplementary Figure 19). Interestingly, the red module (20 newly-assembled DEGs) from data set 1, harbored *CqZTL* and *CqBIC1*. Besides, the grey module of data set 2 contained 884 newly-assembled DEGs together with *CqFT1A*, *CqFT2B*, *CqCOL16*, and *CqELF5*. In data set 3, the grey module clustered 56 newly-assembled DEGs with *CqSOC1*.

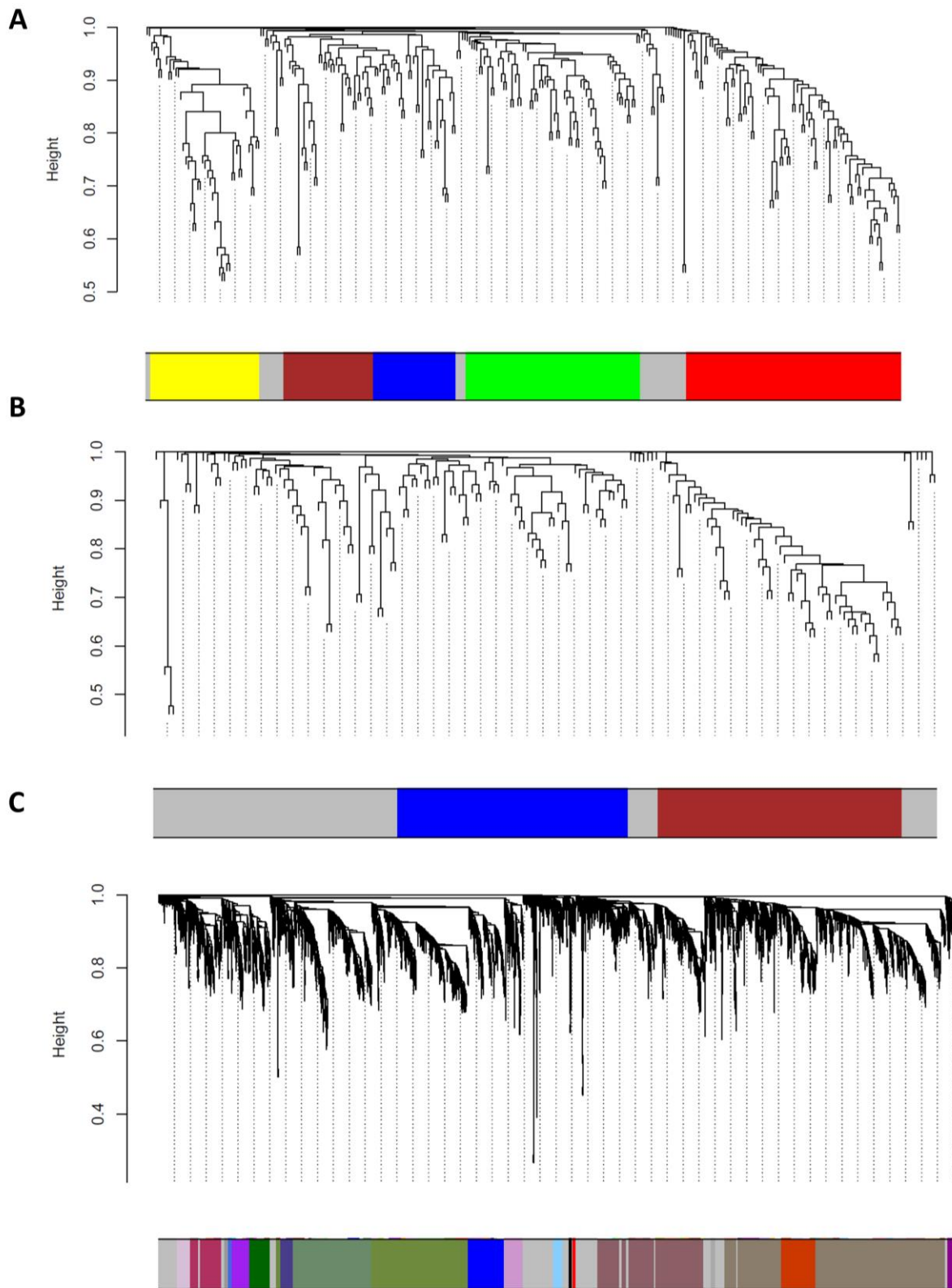


Figure 11. Co-expression analysis of annotated and newly-assembled DEGs in leaves. Newly-assembled transcripts are those sequences assembled by the algorithm, whose transcripts differed from the ones found in the reference genome (referred to as annotated). (A) Differentially expressed genes (DEGs) between short-day (SD) and long-day (LD) conditions in D-12082, which differentially respond to photoperiod to likely regulate flowering time. (B) DEGs between PI-614886 and D-12082 at W2 and W3 under LD, which are putatively controlling flowering under LD. (C) DEGs with similar spatial expression patterns between PI-614886 and D-12082, which likely control the time to flower under SD. Hierarchical clustering dendrograms of the DEGs are shown at the top. Modules, which are designated by color code at the bottom of each tree, are the branches of the clustering tree. The height axis displays the distance between clusters.

To continue, I constructed co-expression networks for each of the modules (Supplementary Figure 20, Supplementary Figure 21). For further examination, I selected the top 10% of newly-assembled transcripts in each module with the highest degree that have MM > 0.8. Next, I blasted and GO-annotated these genes (Supplementary Table 20, Supplementary Table 20). As an outcome, 90.73% of the genes depicted sequence similarity to one or more gene orthologs in other species by BLAST analysis, out of which 24.91% were coding for uncharacterized proteins (Supplementary Table 20). Importantly, none of the resulting genes showed similarity to *PEBP* or other flowering time gene families. Such similarity would indicate a putative role in flowering time regulation. However, like the annotated genes, the GO analysis revealed terms such as “Regulation of gene expression” and “Transcription, DNA-templated” (Supplementary Table 20).

3.4.6. Validation of RNA-seq pipeline by RT-qPCR

I validated the transcriptome results by performing expression analysis of *CqTPS9*, *CqZTL*, *CqFIP1*, *CqELF5*, *CqSOC1*, *CqFT2B*, and *CqHD3AB* by RT-qPCR (Figure 12). Then, I calculated the correlation between the differential expression values from RNA-seq and the values from RT-qPCR. As an outcome, the differential expression analyzed by RT-qPCR closely matched the results of RNA-seq with a correlation coefficient *R* of 0.82 (Supplementary Figure 22). The expression patterns obtained by RT-qPCR matched my expectations based on my transcriptome data dissection. For instance, *CqZTL* relative expression is higher in D-12082 than in PI-614886 under LD at W3. Thus, *CqZTL* expression is higher in the accession that differentially responds to photoperiod at the stage when flowering begins in LD, based on my histological analysis. These results are in line with a gene that responds differently to photoperiod to control flowering time. Besides, as expected, the relative expression of *CqZTL* remains relatively unchanged under both day length conditions in the day-neutral accession. A further example is given for *CqFT2B*. This gene expression was much higher in PI-614886 than in D-12082 under LD at W3, matching my observations by RNA-seq (Figure 10B). Interestingly, *CqFT2B* relative expression was higher towards W3 under SD in both accessions, showing a correlation of its expression with the time to flower (as detected in my histological analysis) under this photoperiod regime.

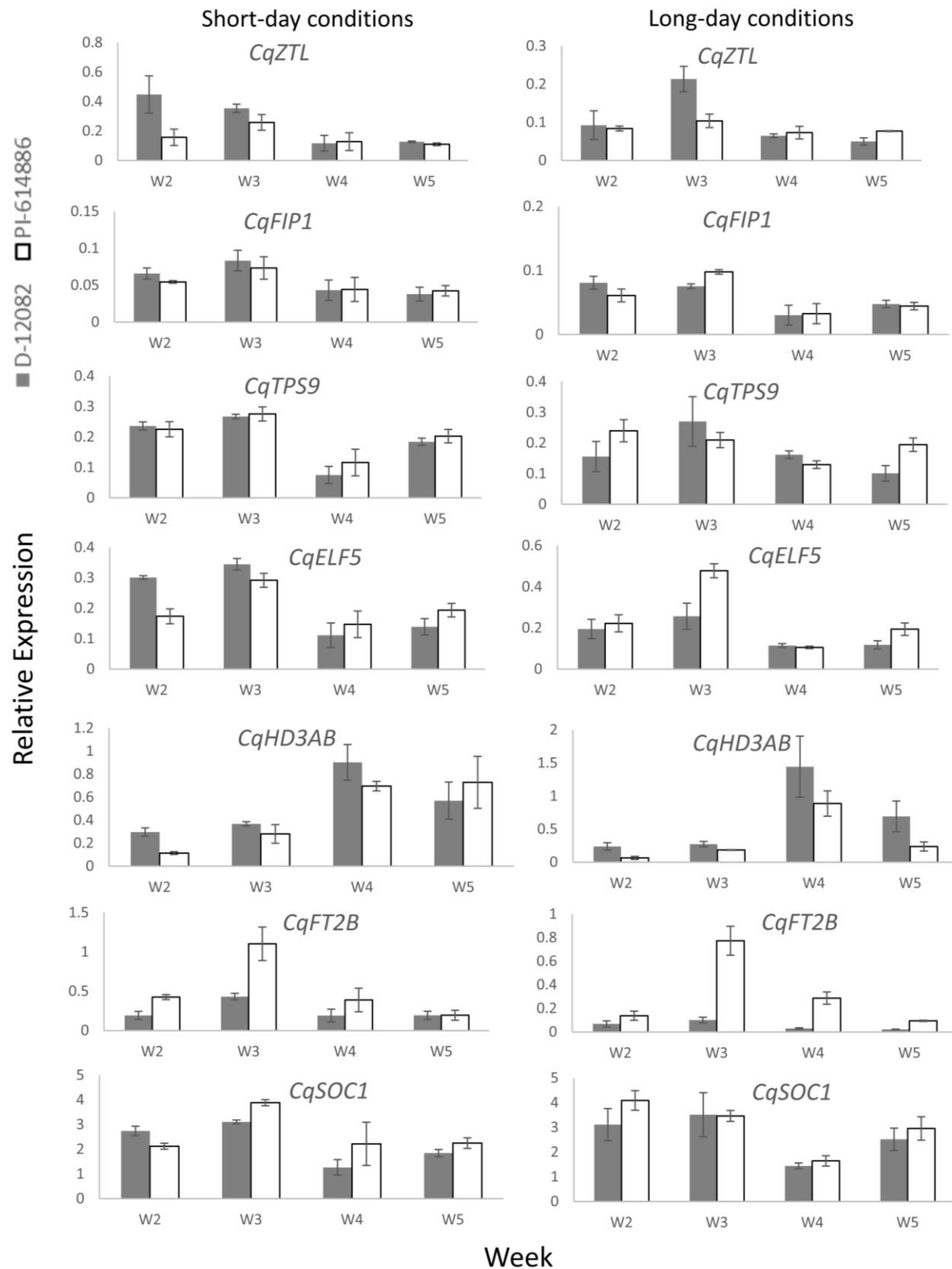


Figure 12. Leaf expression profiles of six genes at different stages of development obtained by RT-qPCR. Seven genes were used to validate the RNA-seq differential expression results. Plants from accessions PI-614886 and D-12082 were grown in a growth chamber at 22°C under short-day (SD, 8 h light) and long-day conditions (LD, 16 h light). Error bars: \pm SEM; data were normalized against the geometric mean of *CqPTB* and *CqIDH-A*. W=week.

3.5. Discussion

In this study, I determined the time point when the quinoa shoot apical meristem (SAM) turned into a floral meristem. Besides, I studied leaf and SAM transcriptomes in correlation with this morphological shift at the SAM. I investigated genes displaying differential expression profiles between different photoperiods and between accessions with contrasting life cycle regimes and obtained thousands of DEGs as an outcome.

I aimed to determine the time point when the shoot apical meristem (SAM) turned into a floral meristem. I investigated SAM development at different stages to identify the time of the switch to the reproductive stage. I determined the time of floral transition based on apical dominance release; thus, a vegetative stage is identified by an “edged” shaped SAM, whereas a reproductive meristem depicts a “dome shape”, showing that inhibition on lateral bud growth took place (Dun et al., 2006). The results from this analysis placed the time for floral transition in quinoa much like the time of bolting. I conclude that genes triggering floral transition should be already active at the early stages of quinoa development when the second pair of leaves appear. This is in agreement with reports in *Chenopodium rubrum* (*C. rubrum*), a species closely related to quinoa, where *CrFTL* genes respond to SD conditions to flower as early as five days after sowing (Cháb et al., 2008). Contrasting to quinoa, other crops transit to flower much later, as is the case of rapeseed and sugar beet. In winter rapeseed, the “dome” shaped apical meristem, which shows that flowering started, is observed six weeks after vernalization in LD (Matar et al., 2021). In *B. vulgaris*, the vegetative shoot apical meristem before vernalization has an “edged” shape. Later, no significant morphological changes in the apical meristem can be seen during vernalization and the characteristic “dome-shape” can be seen only 6-8 weeks after the end of vernalization (Chiurugwi et al., 2013).

Although floral transition starts early in quinoa, the time of flowering (i.e. when the first flower opens), also depends on adequate environmental signals, as demonstrated in several crops (e.g. tomato, rice, wheat), where nutrient supply, day length, light intensity, light quality, and ambient temperature, as well as endogenous signals transmitted by plant hormones, play an important role (Bäurle & Dean, 2006; Camejo et al., 2005; Hu et al., 2015). In winter rapeseed, for instance, the shortening of the pre-vernalization periods from three to one week derived in a flowering delay of up to 51 days after vernalization (Matar et al., 2021). In a study with *Brassica rapa*, plants grown at 28 °C instead of 21 °C, showed a delay of 7 days in floral transition, as seen by inspection of the morphology of the SAM (Del Olmo, Poza-Viejo, Piñeiro, Jarillo, & Crevillén, 2019).

After I had determined the time point when the SAM turned into a floral meristem, I searched for genes expressed in the leaves, which are likely involved in flowering time regulation under SD and LD conditions, as well as those which differentially respond to photoperiod. I reasoned that such genes are differentially regulated at W2 and W3. I am aware of the fact that there are flowering time regulators, whose constitutive expression is enough to flower, such as the *FTL* homolog in chrysanthemum (Higuchi & Hisamatsu, 2015). However, many known genes controlling flowering respond to photoperiod and other environmental cues, and its expression is correlated with the morphological changes at the SAM (Putterill & Varkonyi-Gasic, 2016).

I hypothesized different regulatory pathways in quinoa under SD and LD because, in a previous study, substantial differences in the expression patterns of six *FT*-like and six *COL* homologs under different photoperiodic conditions were found (Patiranage et al., 2021). It is reasonable to presume that different pathways and different genes may contribute to regulating the flowering time of quinoa under different photoperiod regimes, as is the case of rice or soybean. In rice, *HD3A* acts as a floral integrator under SD and *RICE FLOWERING LOCUS T 1 (RFT1)* under LD (Tsuji, Taoka, & Shimamoto, 2011). In soybean, the *FT* homologs *GmFT2a* and *GmFT5a* respond differently to photoperiod, and only

GmFT5a induces flowering under LD conditions (Kong et al., 2010). The hypothesis of two-pathway regulation of quinoa flowering time is supported by a study on *Chenopodium ficifolium*, which flowers under LD without showing differential expression of *FT* homologs, while under SD, *CrFTL1* is required for flowering (Štorchová et al., 2019). Moreover, when I compared gene expression profiles between accessions under SD and LD, I found only a few DEGs in common between these photoperiod regimes.

CqFT2B and *CqFT1A* constitute further evidence supporting the quinoa flowering time being differently regulated under SD and LD. *CqFT2B* and *CqFT1A* are the most likely functional *FT* homologs in quinoa. In my study, *CqFT2B* is suggested to be a functional ortholog of *AtFT* because of its significant up-regulation in PI-614886 under LD. However, it is more likely that this gene is a major player in flowering time regulation under SD than LD, because its expression pattern correlates with that of floral activator under SD only (Figure 6). The evidence of *CqFT2B*'s role as a flowering promoter under SD provided by my study adds to that of other studies (Patiranage et al., 2021; Wu et al., 2021). *CqFT2B* was presumed to be the quinoa floral activator under SD by Patiranage et al. (2021), which was evidenced also by Wu et al. (2021). They studied the effect of night break (NB) under SD in quinoa. They found that *CqFT2B* is a target of lncRNA involved in flowering time regulation. On the other hand, another *FT* ortholog, *CqFT1A*, was shown to influence flowering time under LD and not SD in a haplotype study of a quinoa core collection (Patiranage et al., 2021). Patiranage et al. (2021) found that a *CqFT1A* allele (named *CqFT1Aa*) is correlated with early flowering, while the allele named *CqFT1Ab* is correlated with late flowering under LD, exclusively. In correspondence with that study, I found that the day-neutral accession PI-614886, used in my study, carries the *CqFT1Aa* allele, while D-12082 carries the *CqFT1Ab* allele, which is in correlation with the observed flowering delay under LD. Accordingly, I observed a down-regulation of *CqFT1A* at W2 under LD in PI-614886, which flowered earlier than D-12082. Hence, according to my findings and in line with the previous observations in Patiranage et al. (2021), *CqFT1A* might work as a floral repressor under LD. The putative function of *CqFT1A* in LD conditions would be similar to that of *BvFT1*, which is the *FT* homolog in sugar beet, another member of the *Amaranthaceae* family. In *B. vulgaris*, *BvFT1* inhibits flowering under LD. Further evidence for the function of *CqFT2B* and *CqFT1A* in flowering time regulation is presented by the fact that they were not found in the SAM transcriptome data (Lee & Imaizumi, 2018). To conclude, *CqFT2B* might have a role in flowering promotion under SD but not under LD, while *CqFT1A* might be a flowering time repressor under LD but not in SD. Altogether, these findings support the hypothesis of two different pathways regulating flowering time in quinoa under different photoperiod regimes.

Interestingly, as a result of my transcriptome data dissection, I did not encounter the *TWIN SISTER of FT* (*TSF*) homologs mentioned by Wu et al. (2021), which are targets of differentially regulated lncRNA after NB under SD in quinoa. Besides, they mentioned genes down- or up-regulated after NB: *LHY*, *ELF3*, *HY5*, *PHYA*, and *CRY1*, which were not found as involved in flowering time regulation in my investigation, either. This might be because of the different experimental designs of the two studies, as I analyzed DEGs (mRNA) differentially responding to photoperiod regimes at the time points when flowering takes place based on histological analysis of the SAM.

How can I distinguish between genes directly involved in flowering time regulation and others? In my analysis, I examined differentially expressed genes between different conditions, accessions, and developmental stages to detect flowering time regulators and obtained hundreds of genes as a result. These numbers matched my predictions because flowering is regulated by a complex network comprising numerous pathways and hundreds of genes (Blümel, Dally, & Jung, 2015; Bouché, Lobet, Tocquin, & Périlleux, 2016).

I hypothesize that most of the DEGs between accessions under LD at W2 in leaf and SAM are involved in flowering time. Similarly, all DEGs between photoperiodic conditions at W2 are photoperiod-response candidates with a connection to the transition to flowering. There is increasing evidence that most phenotypes, which were previously thought to be simply inherited, are controlled by many genes, although with minor effects. The omnigenic paradigm even proposes that any trait is influenced by all genes that are expressed at the right moment (Boyle, Li, & Pritchard, 2017; Mathieson, 2021; Tautz, Reeves, & Pallares, 2020). Although major (Mendelian) genes with large effects had been found in many studies in maize, Buckler et al. (2009) found that, in fact, differences in flowering time among inbred lines are not determined by a few genes with large effects, but by the cumulative effects of many quantitative trait loci, each of which with small effects. Furthermore, according to the recently proposed omnigenic model, genes can be classified as core or peripheral in their association with a specific trait. Therefore, even if the candidate DEG does not have an obvious connection to flowering-related pathways, I could expect it to be integrated into non-core pathways (e.g., nitrogen metabolism) (Cho, Yoon, & An, 2017). Confirming this statement, my KEGG pathway mapping analysis showed a wide range of outcomes. Further evidence of the omnigenic nature of flowering time conveys the results of the GO annotations (Boyle et al., 2017). As expected, I found GO terms such as “Transcription regulator activity” and “Regulation of gene expression”, which are common for flowering-time-related genes (Song, Shim, Kinmonth-Schultz, & Imaizumi, 2015). However, several other functional categories, which also contribute to describing the function of the DEGs’ set, were displayed.

As a polyploid species, quinoa represents a challenge when finding functional orthologues of a particular gene family because of gene redundancy. Besides, neo- or sub-functionalization of paralogs causing alterations in gene function is expected, given quinoa’s evolutionary history (Jarvis et al., 2017). In light of this statement, out of 24 *PEBP* (*phosphatidyl ethanolamine-binding proteins*) family genes, five of which had been identified before as putative *FT* orthologues, only *CqFT2B* and *CqFT1A* were detected in my study in leaves (Jarvis et al., 2017; Patiranage et al., 2021), showing that transcriptome analysis is a useful tool to filter out those genes which are unlikely to be functional paralogs for regulating different traits in a polyploid species. Interestingly, in my study, I found genes, whose orthologues in other species, are previously reported to control flowering regulation through the vernalization pathway. These genes include *EARLY FLOWERING 5 (ELF5)*, *FRIGIDA INTERACTING PROTEIN (CqFIP1)*, and *FRIGIDA-like (CqFRL4A)* in the leaf transcriptome and *CqVRN1* in the SAM tissue. These genes function under the vernalization pathway in either Arabidopsis, rapeseed, or wheat (Johanson et al., 2000; Noh, Bizzell, Noh, Schomburg, & Amasino, 2004; Wang et al., 2011; Yan et al., 2003). Quinoa does not require vernalization to flower. Therefore, it is tempting to speculate that these quinoa orthologues possibly underwent neo-functionalization. Hence, during evolution, these genes distinctly adapted their function more than other *Amaranthaceae* species, such as sugar beet, which requires vernalization to flower under LD.

I found 57,820 newly-assembled transcripts in the leaf tissue as the outcome of the bioinformatics pipeline. The newly-assembled transcripts are those sequences assembled by the algorithm, which differed from the ones found in the reference genome. These include algorithmic artifacts, putative gene isoforms (e.g. splice variants), and, less probable, new genes transcribed specifically under my experimental conditions (Pertea et al., 2015). In my study, I assumed that interesting gene isoforms and new genes transcribed specifically under my experimental conditions, if any, are highly co-expressed with my main candidate floral regulators. In this way, I used co-expression networks to associate newly-assembled transcripts with biological processes and to prioritize candidate genes for splice variants (Van Dam, Vosa, van der Graaf, Franke, & de Magalhaes, 2018). Besides, I assumed that newly-assembled transcripts with a flowering time-related function have sequence similarity to genes belonging to known flowering time gene families. I assessed that similarity by blastx analysis.

As a result of co-expression and blastx analyses, no newly-transcribed gene was associated with a flowering function in my study. Importantly, in my transcriptome data dissection, I assumed that the same genes act as floral integrators in the two selected accessions, thus, possibly filtering out splicing variants of floral integrators. However, the lack of gene isoforms (e.g. splicing variants) as a result of co-expression and blastx analysis, would show that my assumption of the same flowering integrators in the two accessions under study is likely correct. More extensive analyses including modifications in the filtering criteria to dissect this data can be performed. Examples of such analysis are regulatory gene networks, interaction with target genes, sequence motif analysis, and integration of protein-protein interaction data.

Although transcriptome analysis provides a first insight into the function of a gene, it does not convey ultimate evidence of the role of the candidate genes in a biological process. Therefore, the information given by my study must be combined with other methods for functional characterization to confirm the role of the candidate genes as flowering time regulators in quinoa. In this scenario, overexpression and knockout of the candidate genes would provide the ultimate evidence of its function in quinoa. Up to date, there is no protocol for quinoa transformation; however, a recent Virus-Mediated transient expression technique (VIGS) protocol could enable functional studies in this crop (Ogata et al., 2021). Moreover, a combination of my transcriptome data with genetic mapping studies would provide stronger evidence of the suitability of candidate genes as flowering time regulators and haplotype analysis could unravel correlations between haplotypes of the candidate genes and/or their regulatory regions with days. Variants modifying *CqFT1A*'s expression could be of interest for growing quinoa in Northern latitudes given the putative function of this gene as a floral repressor acting only under long day conditions and in the early stages of quinoa development.

3.6. Acknowledgments

I thank Prof. Dr. Mark Tester and Dr. Elodie Rey (King Abdullah University of Science and Technology) for performing mapping and assembly of the raw reads and for continuous support. I thank Dr. Nazgol Emrani and Prof. Dr. Jung (Kiel University) for their evaluation of the experimental design, data treatment and results interpretation.

3.7. References

- Adamczyk, B. J., Lehti-Shiu, M. D., & Fernandez, D. E. (2007). The MADS domain factors *AGL15* and *AGL18* act redundantly as repressors of the floral transition in Arabidopsis. *The Plant Journal*, 50(6), 1007-1019.
- Alandia, G., Rodriguez, J., Jacobsen, S.-E., Bazile, D., & Condori, B. (2020). Global expansion of quinoa and challenges for the Andean region. *Global Food Security*, 26, 100429.
- Almeida-Silva, F., & Venancio, T. M. (2022). BioNERO: an all-in-one R/Bioconductor package for comprehensive and easy biological network reconstruction. *Functional Integrative Genomics*, 22(1), 131-136.
- Bäurle, I., & Dean, C. (2006). The timing of developmental transitions in plants. *Cell*, 125(4), 655-664.
- Blázquez, M. A., & Weigel, D. (1999). Independent regulation of flowering by phytochrome B and gibberellins in Arabidopsis. *Plant Physiology*, 120(4), 1025-1032.
- Blümel, M., Dally, N., & Jung, C. (2015). Flowering time regulation in crops—what did we learn from Arabidopsis? *Current Opinion in Biotechnology*, 32, 121-129.

- Bouché, F., Lobet, G., Tocquin, P., & Périlleux, C. (2016). FLOR-ID: an interactive database of flowering-time gene networks in *Arabidopsis thaliana*. *Nucleic Acids Research*, 44(D1), D1167-D1171.
- Boyle, E. A., Li, Y. I., & Pritchard, J. K. (2017). An expanded view of complex traits: from polygenic to omnigenic. *Cell*, 169(7), 1177-1186.
- Buckler, E. S., Holland, J. B., Bradbury, P. J., Acharya, C. B., Brown, P. J., Browne, C., . . . Glaubitz, J. C. (2009). The genetic architecture of maize flowering time. *Science*, 325(5941), 714-718.
- Camejo, D., Rodríguez, P., Morales, M. A., Dell’Amico, J. M., Torrecillas, A., & Alarcón, J. J. (2005). High temperature effects on photosynthetic activity of two tomato cultivars with different heat susceptibility. *Journal of Plant Physiology*, 162(3), 281-289.
- Cháb, D., Kolář, J., Olson, M. S., & Štorchová, H. (2008). Two *FLOWERING LOCUS T (FT)* homologs in *Chenopodium rubrum* differ in expression patterns. *Planta*, 228(6), 929.
- Chen, L., Nan, H., Kong, L., Yue, L., Yang, H., Zhao, Q., . . . Lu, S. (2020). Soybean *API* homologs control flowering time and plant height. *Journal of Integrative Plant Biology*, 62(12), 1868-1879.
- Chen, Y., McCarthy, D., Ritchie, M., Robinson, M., Smyth, G., & Hall, E. (2020). edgeR: differential analysis of sequence read count data User’s Guide. *Bioinformatics*, 26(1).
- Chiurugwi, T., Holmes, H. F., Qi, A., Chia, T. Y., Hedden, P., & Mutasa-Göttgens, E. (2013). Development of new quantitative physiological and molecular breeding parameters based on the sugar-beet vernalization intensity model. *J. Agric. Sci*, 151, 492-505.
- Cho, L. H., Yoon, J., & An, G. (2017). The control of flowering time by environmental factors. *The Plant Journal*, 90(4), 708-719.
- Choi, K., Kim, J., Hwang, H.-J., Kim, S., Park, C., Kim, S. Y., & Lee, I. (2011). The *FRIGIDA* complex activates transcription of *FLC*, a strong flowering repressor in *Arabidopsis*, by recruiting chromatin modification factors. *The Plant Cell*, 23(1), 289-303.
- Del Olmo, I., Poza-Viejo, L., Piñeiro, M., Jarillo, J. A., & Crevillén, P. (2019). High ambient temperature leads to reduced *FT* expression and delayed flowering in *Brassica rapa* via a mechanism associated with H2A. Z dynamics. *The Plant Journal*, 100(2), 343-356.
- Drabešová, J., Cháb, D., Kolář, J., Haškovcová, K., & Štorchová, H. (2014). A dark–light transition triggers expression of the floral promoter *CrFTL1* and downregulates *CONSTANS*-like genes in a short-day plant *Chenopodium rubrum*. *Journal of Experimental Botany*, 65(8), 2137-2146.
- Dun, E. A., Ferguson, B. J., & Beveridge, C. A. (2006). Apical dominance and shoot branching. Divergent opinions or divergent mechanisms? *Plant Physiology*, 142(3), 812-819.
- Fuller, H. J. (1949). Photoperiodic responses of *Chenopodium quinoa* Willd. and *Amaranthus caudatus* L. *American Journal of Botany*, 175-180.
- Gaudinier, A., & Blackman, B. K. (2020). Evolutionary processes from the perspective of flowering time diversity. *New Phytologist*, 225(5), 1883-1898.
- Golicz, A. A., Steinfort, U., Arya, H., Singh, M. B., & Bhalla, P. L. (2020). Analysis of the quinoa genome reveals conservation and divergence of the flowering pathways. *Functional Integrative Genomics*, 20(2), 245-258.

- Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., . . . Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*, 36(10), 3420-3435.
- Granado-Rodríguez, S., Aparicio, N., Matías, J., Pérez-Romero, L. F., Maestro, I., Gracés, I., . . . Del Hierro, J. N. (2021). Studying the impact of different field environmental conditions on seed quality of quinoa: The case of three different years changing seed nutritional traits in Southern Europe. *Frontiers in plant science*, 12.
- Higuchi, Y., & Hisamatsu, T. (2015). *CsTFL1*, a constitutive local repressor of flowering, modulates floral initiation by antagonising florigen complex activity in chrysanthemum. *Plant Science*, 237, 1-7.
- Hu, Y., Liang, W., Yin, C., Yang, X., Ping, B., Li, A., . . . Cai, Q. (2015). Interactions of *OsMADS1* with floral homeotic genes in rice flower development. *Molecular Plant*, 8(9), 1366-1384.
- Iqbal, S., Basra, S., Saddiq, M. S., Yang, A., Akhtar, S. S., & Jacobsen, S.-E. (2020). The extraordinary salt tolerance of quinoa. In *Emerging Research in Alternative Crops* (pp. 125-143): Springer.
- Jarvis, D. E., Ho, Y. S., Lightfoot, D. J., Schmöckel, S. M., Li, B., Borm, T. J., . . . Saber, N. (2017). The genome of *Chenopodium quinoa*. *Nature*, 542(7641), 307.
- Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R., & Dean, C. J. S. (2000). Molecular analysis of *FRIGIDA*, a major determinant of natural variation in Arabidopsis flowering time. 290(5490), 344-347.
- Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 28(1), 27-30.
- Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nature Methods*, 12(4), 357-360.
- Kim, S.-K., Yun, C.-H., Lee, J. H., Jang, Y. H., Park, H.-Y., & Kim, J.-K. (2008). *OsCO3*, a *CONSTANS-LIKE* gene, controls flowering by negatively regulating the expression of FT-like genes under SD conditions in rice. *Planta*, 228(2), 355-365.
- Komiya, R., Ikegami, A., Tamaki, S., Yokoi, S., & Shimamoto, K. (2008). *Hd3a* and *RFT1* are essential for flowering in rice. *Development*, 135(4).
- Kong, F., Liu, B., Xia, Z., Sato, S., Kim, B. M., Watanabe, S., . . . Harada, K. (2010). Two coordinately regulated homologs of *FLOWERING LOCUS T* are involved in the control of photoperiodic flowering in soybean. *Plant Physiology*, 154(3), 1220-1231.
- Langfelder, P., & Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, 9(1), 1-13.
- Lee, J., & Lee, I. (2010). Regulation and function of *SOC1*, a flowering pathway integrator. *Journal of Experimental Botany*, 61(9), 2247-2254.
- Lee, N., & Imaizumi, T. (2018). Uncoupling FT protein transport from its function. *Plant Cell Physiology*, 59(8), 1487.
- Liu, L., Li, C., Song, S., Teo, Z. W. N., Shen, L., Wang, Y., . . . Yu, H. (2018). *FTIP*-dependent *STM* trafficking regulates shoot meristem development in Arabidopsis. *Cell Reports*, 23(6), 1879-1890.
- Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta CT$ method. *Methods*, 25(4), 402-408.

- Maldonado-Taipe, N., Patirange, D. S., Schmöckel, S. M., Jung, C., & Emrani, N. (2021). Validation of suitable genes for normalization of diurnal gene expression studies in *Chenopodium quinoa*. *PloS one*, 16(3), e0233821.
- Más, P., Kim, W.-Y., Somers, D. E., & Kay, S. A. (2003). Targeted degradation of *TOC1* by *ZTL* modulates circadian function in *Arabidopsis thaliana*. *Nature*, 426(6966), 567-570.
- Matar, S., Kumar, A., Holtgräwe, D., Weisshaar, B., & Melzer, S. (2021). The transition to flowering in winter rapeseed during vernalization. *Plant, Cell & Environment*, 44(2), 506-518.
- Mathieson, I. (2021). The omnigenic model and polygenic prediction of complex traits. *The American Journal of Human Genetics*, 108(9), 1558-1563.
- Murai, K., Miyamae, M., Kato, H., Takumi, S., & Ogihara, Y. (2003). *WAP1*, a wheat *APETALA1* homolog, plays a central role in the phase transition from vegetative to reproductive growth. *Plant Cell Physiology*, 44(12), 1255-1265.
- Murphy, K. M., Matanguihan, J. B., Fuentes, F. F., Gómez-Pando, L. R., Jellen, E. N., Maughan, P. J., & Jarvis, D. (2018). Quinoa breeding and genomics. *Plant Breeding Reviews*, 42, 257-320.
- Noh, Y. S., Bizzell, C. M., Noh, B., Schomburg, F. M., & Amasino, R. M. (2004). *EARLY FLOWERING 5* acts as a floral repressor in *Arabidopsis*. *The Plant Journal*, 38(4), 664-672.
- Ogata, T., Toyoshima, M., Yamamizo-Oda, C., Kobayashi, Y., Fujii, K., Tanaka, K., . . . Nagatoshi, Y. (2021). Virus-Mediated Transient Expression Techniques Enable Functional Genomics Studies and Modulations of Betalain Biosynthesis and Plant Height in Quinoa. *Frontiers in plant science*, 12.
- Patiranage, D. S., Asare, E., Maldonado-Taipe, N., Rey, E., Emrani, N., Tester, M., & Jung, C. (2021). Haplotype variations of major flowering time genes in quinoa unveil their role in the adaptation to different environmental conditions. *Plant, Cell & Environment*.
- Patirange, D. S., Rey, E., Emrani, N., Wellman, G., Schmid, K., Schmöckel, S. M., . . . Jung, C. (2020). Genome-wide association study in the pseudocereal quinoa reveals selection pattern typical for crops with a short breeding history. *bioRxiv*.
- Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., & Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols*, 11(9), 1650-1667.
- Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, 33(3), 290-295.
- Pin, P. A., Benlloch, R., Bonnet, D., Wremmerth-Weich, E., Kraft, T., Gielen, J. J., & Nilsson, O. (2010). An antagonistic pair of FT homologs mediates the control of flowering time in sugar beet. *Science*, 330(6009), 1397-1400.
- Putterill, J., & Varkonyi-Gasic, E. (2016). *FT* and florigen long-distance flowering control in plants. *Current Opinion in Plant Biology*, 33, 77-82.
- Robinson, M. D., & Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology*, 11(3), 1-9.
- Shi, P., & Gu, M. (2020). Transcriptome analysis and differential gene expression profiling of two contrasting quinoa genotypes in response to salt stress. *BMC Plant Biology*, 20(1), 1-15.

- Song, Y. H., Shim, J. S., Kinmonth-Schultz, H. A., & Imaizumi, T. (2015). Photoperiodic flowering: time measurement mechanisms in leaves. *Annual Review of Plant Biology*, 66, 441-464.
- Sosa-Zuniga, V., Brito, V., Fuentes, F., & Steinfort, U. (2017). Phenological growth stages of quinoa (*Chenopodium quinoa*) based on the BBCH scale. *Annals of Applied Biology*, 171(1), 117-124.
- Štorchová, H., Hubáčková, H., Abeyawardana, O. A., Walterová, J., Vondráková, Z., Eliášová, K., & Mandák, B. (2019). *Chenopodium ficifolium* flowers under long days without upregulation of *FLOWERING LOCUS T (FT)* homologs. *Planta*, 250(6), 2111-2125.
- Suarez-Lopez, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F., & Coupland, G. (2001). *CONSTANS* mediates between the circadian clock and the control of flowering in *Arabidopsis*. *Nature*, 410(6832), 1116-1120.
- Subedi, M., Neff, E., & Davis, T. M. (2021). Developing *Chenopodium ficifolium* as a potential B genome diploid model system for genetic characterization and improvement of allotetraploid quinoa (*Chenopodium quinoa*). *BMC Plant Biology*, 21(1), 1-18.
- Tautz, D., Reeves, G., & Pallares, L. F. (2020). New experimental support for long standing concepts of polygenic genetics implies that the Mendelian genetic paradigm needs to be revised: The New (Old) Genetics, Version 1.0. *NAL-live*, 2020(1), 1-15.
- Tian, H., Li, Y., Wang, C., Xu, X., Zhang, Y., Zeb, Q., . . . Li, L. (2021). Photoperiod-responsive changes in chromatin accessibility in phloem companion and epidermis cells of *Arabidopsis* leaves. *The Plant Cell*, 33(3), 475-491.
- Tsuji, H., Taoka, K.-i., & Shimamoto, K. (2011). Regulation of flowering in rice: two florigen genes, a complex gene network, and natural variation. *Current Opinion in Plant Biology*, 14(1), 45-52.
- Van Dam, S., Vosa, U., van der Graaf, A., Franke, L., & de Magalhaes, J. P. (2018). Gene co-expression analysis for functional classification and gene–disease predictions. *Briefings in Bioinformatics*, 19(4), 575-592.
- Walsh, B. M., Adhikary, D., Maughan, P. J., Emshwiller, E., & Jellen, E. N. (2015). *Chenopodium* polyploidy inferences from Salt Overly Sensitive 1 (SOS1) data. *American Journal of Botany*, 102(4), 533-543.
- Wang, N., Qian, W., Suppanz, I., Wei, L., Mao, B., Long, Y., . . . Jung, C. (2011). Flowering time variation in oilseed rape (*Brassica napus* L.) is associated with allelic variation in the *FRIGIDA* homologue *BnaA. FRI. a*. *Journal of Experimental Botany*, 62(15), 5641-5658.
- Wang, Q., Zuo, Z., Wang, X., Gu, L., Yoshizumi, T., Yang, Z., . . . Han, Y.-J. (2016). Photoactivation and inactivation of *Arabidopsis* cryptochrome 2. *Science*, 354(6310), 343-347.
- Wu, Q., Bai, X., Wu, X., Xiang, D., Wan, Y., Luo, Y., . . . Qin, P. (2020). Transcriptome profiling identifies transcription factors and key homologs involved in seed dormancy and germination regulation of *Chenopodium quinoa*. *Plant Physiology Biochemistry*, 151, 443-456.
- Wu, Q., Bai, X., Zhao, W., Shi, X., Xiang, D., Wan, Y., . . . Peng, L. (2019). Investigation into the underlying regulatory mechanisms shaping inflorescence architecture in *Chenopodium quinoa*. *BMC genomics*, 20(1), 1-20.

Chapter 3

Wu, Q., Luo, Y., Wu, X., Bai, X., Ye, X., Liu, C., . . . Zou, L. (2021). Identification of the specific long-noncoding RNAs involved in night-break mediated flowering retardation in *Chenopodium quinoa*. BMC genomics, 22(1), 1-18.

Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T., & Dubcovsky, J. (2003). Positional cloning of the wheat vernalization gene *VRN1*. Proceedings of the National Academy of Sciences, 100(10), 6263-6268.

3.8. Supplementary data

Supplementary Table 9. Sampling stages and tissues from D-12082 and PI-614886 under two different photoperiod regimes. Samples were taken for RNA-sequencing (RNA-seq) and histological analysis. Plants were grown at 22 °C under long-day (LD) conditions: 16 h light; and short-day conditions (SD): 8 h light. Sequenced samples are written in red.

Supplementary Table 10. Quality filtering and raw read statistics for the transcriptome libraries. Three biological replicates of shoot apical meristem and leaf tissue from two accessions grown under two different conditions were sequenced. Samples were collected weekly. Long-day (LD) conditions: 16 h light, 22°C; short-day conditions (SD): 8 h light, 22°C. PE= pair-end.

Supplementary Table 11. List of candidate genes putatively regulating flowering time in quinoa with an associated flowering time function in quinoa and other species.

Supplementary Table 12. Gene Ontology (GO) annotation of differentially expressed genes (DEGs) between SD and LD in leaves of D-12082, which differentially respond to photoperiod to likely regulate flowering time. The number of genes in each node of different GO hierarchy levels is shown for three GO categories, Biological Process, Cellular Component, and Molecular Function.

Supplementary Table 13. KEGG classification of differentially expressed genes (DEGs) between SD and LD in leaves of D-12082, which differentially respond to photoperiod to likely regulate flowering time.

Supplementary Table 14. Gene Ontology (GO) annotation of differentially expressed genes (DEGs) between PI-614886 and D-12082 at W2 and W3 under long-days (LD) in leaves, which are putatively controlling flowering under LD (16 h light). Number of genes in each node of different GO hierarchy levels is shown for three GO categories, Biological Process, Cellular Component, and Molecular Function.

Supplementary Table 15. KEGG classification of differentially expressed genes (DEGs) between PI-614886 and D-12082 at W2 and W3 under long-days (LD) in leaves, which are putatively controlling flowering under LD (16 h light).

Supplementary Table 16. Gene Ontology (GO) annotation of differentially expressed genes (DEGs) with similar spatial expression patterns in leaves between PI-614886 and D-12082, which likely regulate flowering time under short-days (8 h light). Number of genes in each node of different GO hierarchy levels is shown for three GO categories: Biological Process, Cellular Component, and Molecular Function.

Supplementary Table 17. KEGG classification of differentially expressed genes (DEGs) with similar spatial expression patterns between PI-614886 and D-12082 under short-days (SD) in leaves, which likely regulate flowering time under short-days (8 h light).

Supplementary Table 18. Gene Ontology (GO) annotation of differentially expressed genes (DEGs) in the shoot apical meristem (SAM) between PI-614886 and D-12082 at W2 and W3, which are likely involved in the regulation of the time to flower at the SAM. Number of genes in each node of different GO hierarchy levels is shown for three GO categories: Biological Process, Cellular Component, and Molecular Function.

Supplementary Table 19. KEGG classification of differentially expressed genes (DEGs) in the shoot apical meristem (SAM) between PI-614886 and D-12082 at W2 and W3, which are likely involved in the regulation of the time to flower at the SAM.

Supplementary Table 20. Results of the blastx analysis for newly-assembled transcripts that are co-expressed with annotated DEGs that putatively regulate flowering in leaves. The top 10% of newly-assembled transcripts with the highest degree and module membership (MM) > 0.8 from each module in the co-expression hierarchical clustering were used for blastx. Newly-assembled transcripts are those sequences assembled as transcripts by StringTie, which are not found as transcripts in the reference genome QQ74_V2.

Supplementary Table 21. Gene Ontology (GO) annotation of the newly-assembled transcripts that are co-expressed with annotated DEGs that putatively regulate flowering in leaves. Number of genes in each node of different GO hierarchy levels is shown for three GO categories: Biological Process, Cellular Component, and Molecular Function. Newly-assembled transcripts are those sequences assembled as transcripts by StringTie, which are not found as transcripts in the reference genome QQ74_V2.

Supplementary Figure 14. Quinoa plants at bolting and the first flower opens stages. Plants from (A) PI-614886 and (B) D-12082 grown under long-days (16 h light) are shown. Red arrows show visible floral buds and open flowers. Scale bars=3cm.

Supplementary Figure 15. Unsupervised samples clustering with the filtered and normalized transcriptome data. (A) Leaf and shoot apical meristem (SAM) samples clustered by tissue and accession. SAM and leaf transcriptome values are located to the right and left of the dotted line, respectively. (B) Biological replicates of leaf transcriptomes clustered together. Biological replicates within accession are equally colored. PI-614886 and D-12082 data are located above and under the dotted line, respectively. logFC= Trimmed Mean log-ratios (\log_2 fold change of M values).

Supplementary Figure 16. Transcriptome data investigation to identify genes differentially responding to photoperiod (A) Heatmap and stacked column graph of differentially expressed genes (DEGs) between short-day (SD) and long-day (LD) in D-12082 at W2 and W3. (B) Heatmap and stacked column graph of the DEGs in (A), which were down- or up-regulated in W3. (C) Heatmap and stacked column graph of the DEGs in (B) whose differential expression differed between W2 and W3. (D, E) Heatmap and stacked column graph of DEGs between W2-W5, W3-W5, and W4-W5 for the short-day (D-12082) and the day-neutral (PI-614886) accessions. DEGs with different expression patterns between SD and LD conditions are shown. (F) Heatmap and stacked column graph of the DEGs resulting from removal of PI-614886 DEGs in (D) from those in D-12082. The total number of genes in each step is shown at the bottom of the heatmaps. In the heatmap, SD.W4vsW5 corresponds, for instance, to \log_2 (normalized reads at W4/ normalized reads at W5) for a specific gene model under SD. W=week.

Supplementary Figure 17. Transcriptome data investigation to identify genes regulating flowering time under long-day (LD) and short-day (SD) conditions in leaves. (A) Heatmap and stacked column graph of DEGs between PI-614886 and D-12082 at W2 and W3 under LD. (B) Heatmap and stacked column graph of DEGs having similar spatial expression pattern between D-12082 and PI-614886 under SD conditions. Total number of genes in the analysis is shown at the bottom right of the heatmaps. W4vsW5 corresponds, for instance, to \log_2 (normalized reads at W4/ normalized reads at W5) for a specific gene model under SD. W=week.

Supplementary Figure 18. Transcriptome data investigation of DEGs from the apical meristem. Heatmap and stacked column graph of DEGs between PI-614886 and D-12082 in shoot apical meristem (SAM) samples are shown. The total number of genes is shown at the bottom right of the heatmap. W2.SD corresponds, for instance, to \log_2 (normalized reads in PI-614886/ normalized reads in D-12082) for a specific gene model at W2 under SD. W=week.

Supplementary Figure 19. Number of genes per module resulting from the co-expression analysis in leaf transcriptome. Every color represents one module of the clustering trees in Figure 11. (A) Differentially expressed genes (DEGs) between short-day (SD) and long-day (LD) conditions in D-12082, which differentially respond to photoperiod to likely regulate flowering time. (B) DEGs between PI-614886 and D-12082 at W2 and W3 under LD, which are putatively controlling flowering under LD. (C) DEGs with similar spatial expression patterns between PI-614886 and D-12082, which likely control the time to flower under SD.

Supplementary Figure 20. Co-expression networks of annotated and newly-assembled transcripts showing differentially expressed genes between PI-614886 and D-12082 at W2 and W3 under LD, which are putatively controlling flowering under LD in leaves. Every color represents one module of the clustering tree in Figure 11C. (A) plum3, (B) royalblue, (C) darkolivegreen4, (D) black, (E) red, (F) darkgrey, (G) bisque4, (H) orange3, (I) grey, (J) lightpink4, (K) darkseagreen4, (L) thistle, (M) darkmagenta, (N) skyblue1, (O) darkslateblue, (P) darkgreen, (Q) grey60, (R) purple, (S) blue and (T) maroon modules. Degree: sum of connection weights of a gene to all other genes in the module by Weighted Gene Co-expression Network Analysis (WGCNA). Number of genes in each module is shown in Supplementary Figure 19.

Supplementary Figure 21. Co-expression networks of annotated and newly-assembled transcripts showing (DEGs) between short-day (SD) and long-day (LD) conditions in D-12082, which differentially respond to photoperiod to likely regulate flowering time in leaves (A to F) and DEGs with similar spatial expression patterns between PI-614886 and D-12082, which likely control the time to flower under SD (G to I). Every color represents one module of the clustering trees in Figure 11A and Figure 11B. Networks correspond to the (A) brown, (B) blue, (C) green, (D) yellow, (E) red, (F) grey, (G) blue, (H) brown, and (I) grey modules. Degree: sum of connection weights of a gene to all other genes in the module by Weighted Gene Co-expression Network Analysis (WGCNA). Number of genes in each module is shown in Supplementary Figure 19.

Supplementary Figure 22. Scattered plot showing the correlation between expression data from RNA-seq and RT-qPCR. RT-qPCR values correspond to the \log_2 of the fold change (FC) between day length conditions, between accessions, and between weeks. Data are shown for seven genes. R= Pearson's R. W = week.

4. *Agrobacterium*-mediated transformation experiments

4.1. Introduction

Crop transformation is a useful tool for scientists and breeders, which can be used with a wide range of aims. Besides supplying new and genetically diverse plant materials to breeding programs, genetic plant transformation permits the direct introduction of useful genes into important crops (Keshavareddy et al., 2018). Plant transformation represents a valuable tool for functional analysis to prove the role of (a) gene(s) in a biological process.

In quinoa, the only existing report of transformation dates from 1990. Komari (1990) reported on the transformation of cells in suspension culture. However, attempts to regenerate plants from the transgenic cells were unsuccessful. Another study in *Chenopodium rubrum* (*C. rubrum*), a close relative of quinoa, tested *Agrobacterium* cocultivation with callus from stem, cotyledon, and leaf. After transformation, regeneration could not be achieved (Solís, Mlejnek, Studená, & Procházka, 2003). On the other hand, there are several reports on successful *in vitro* regeneration of shoots from different quinoa explants (Hesami & Daneshvar, 2016; Hesami, Naderi, & Yoosefzadeh-Najafabadi, 2018; Telahigue & Toumi, 2017). These protocols show, for instance, direct organogenesis from cotyledon explant and indirect organogenesis through callus produced from the root, cotyledon, and hypocotyl tissues.

However, the development of an *Agrobacterium*-mediated transformation protocol for quinoa represents a series of challenges. First, *in vitro* regeneration has to be achieved. For this, the utility of existing protocols has to be proven and/or optimized to obtain a reproducible protocol (Keshavareddy et al., 2018). Second, the success of *Agrobacterium*-mediated transformation relies on a combination of many factors, which include genotype, *Agrobacterium* strain, a selectable marker, tissue culture parameters, etc. As seen for other crops like sorghum, finding a suitable combination of these factors is time-consuming. Moreover, even when the transformation is successful, it is possible that high efficiency is not achieved (Che et al., 2018).

Up to date, there is no established protocol for quinoa transformation; thus, limiting the options for functional analysis and transgenic plant development. In this context, I aimed to investigate the feasibility of *Agrobacterium*-mediated quinoa transformation. In this study, I first aimed to prove that *C. quinoa* can be regenerated *in vitro*; later, I tested *Agrobacterium*-mediated transformation by the GUS reporter system; finally, I intended to develop a protocol for *Agrobacterium*-mediated quinoa transformation, which still needs further optimization.

4.2. Materials and methods

4.2.1. Plant material

I used seventeen accessions in quinoa transformation experiments (Table 8). Accessions were used for developing an *in vitro* regeneration protocol (10 accessions) and/or for *Agrobacterium* transformation (10 accessions). Plants were grown in modified Murashige and Skoog medium media (mMS) under 16 h light at 24 °C.

Table 8. Selected accessions for *in vitro* regeneration.

Accession name	Origin	Days to flowering ¹
BO-13	Chile	65
BO-63*	Chile	59
EMBRAPA-Brazil*	NA	55
BO-32	Chile	60
PI-596498	Peru	75
Ames 13745*	USA	59
CHEN-427	Peru	66
CHEN-390	Peru	70
BO-60	Chile	68
PI-510537	Peru	66
PUC-mix-green*	Chile	63
Vikinga*	NA	54
D-12289*	Peru	80
NSL-91567*	NA	63
D-11889*	Peru	61
BO-11*	Chile	56
PI-478414*	Chile	68

NA: Information not available

*Accessions used for *Agrobacterium*-mediated transformation

¹Data collected in the field in 2019 as reported by Patiranage et al. (2020).

4.2.2. Plant transformation and *in vitro* cultivation

In the first stage, I tested *in vitro* regeneration in quinoa without transformation. For this purpose, seeds were disinfected in 1.0 or 3.0% sodium hypochlorite solution (NaOCl) and sown in mMS medium. Explants were cut from petiole, hypocotyl and leaf tissues. They were transferred to either light (16 h) or dark conditions for callus induction. Then, calli were transferred to shoot induction media. I tested genotype, tissue, light conditions, and age of calli at transfer to shooting media. After ca. six weeks, the shoots were transferred to rooting media (Figure 13). The media composition for every stage was selected by a literature survey from the protocols of Hesami and Daneshvar (2016), Hesami et al. (2018), Telahigue and Toumi (2017), and Zarhloul et al. (1999).

In the second stage, I performed *Agrobacterium*-mediated transformation in quinoa calli. I adopted the protocol for rapeseed by Dr. José Orsini (Saaten-Union Biotec GmbH, Germany) with the following modifications (Supplementary Table 22):

- Explants used: petioles
- Callus induction: one week in mMS + 0.25 mg BAP + 0.5 mg 2-4 D; pH: 5.8.
- *Agrobacterium* strain used: *A. tumefaciens* strain GV3101 pMP90RK.
- Inoculation: For the vector pCambia1305.1, 15 ml LB [+ 3 µl rifampicin (25 mg/ml) + 3 µl gentamycin (50 mg/ml) + selection: 3 µl kanamycin (100 mg/ml)] with 200 µl storage culture. For the vector pCas9_TPC_Red, 3 µl spectinomycin was used instead of kanamycin.

I used two vectors: pCambia1305.1 and pCas9_TPC_Red. The vector pCambia1305.1 contains the GUS reporter gene with a catalase intron and hygromycin resistance gene for selection of transgenic plantlets (Supplementary Figure 23A). The vector pCas9_TPC_Red, in which the DsRed cassette has been added into pCas9-TPC, allows red fluorescence selection (Supplementary Figure 23B). This vector was kindly provided by Dr. Hongli Dong

(Southwest University, Chongqing). The vectors were transformed in *A. tumefaciens* strain GV3101 pMP90RK by Dr. Yixin Cui and Dr. Tahmina Islam.

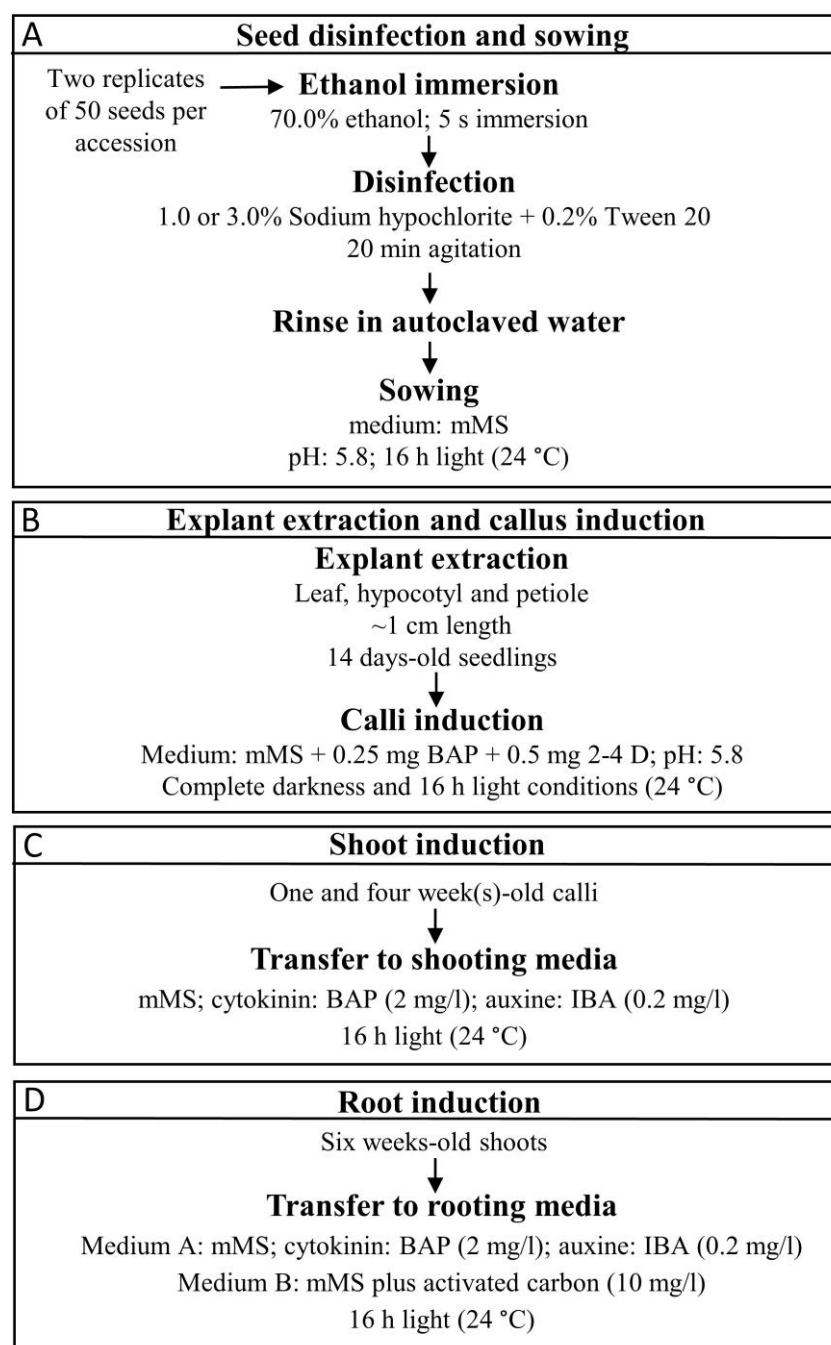


Figure 13. Experimental design for *in vitro* regeneration in quinoa without transformation. Different parameters were tested in four main experimental stages: (A) seed disinfection and sowing for production of sterile seedlings, (B) explant extraction and callus induction, (C) shoot induction, and (D) root induction. mMS: modified Murashige and Skoog, BAP: 6-Benzyl amino purine, IBA: Indole-3-butyric acid.

For *Agrobacterium*-mediated transformation with pCambia1305.1, *Agrobacterium* was cultivated with 3 µl kanamycin (kanamycin selection). Later, I used 100 petioles per accession (BO-63 and Ames 13745) and tested co-cultivation treatments in a factorial design. I tested different co-cultivation conditions plus a negative control for every accession: co-cultivation in liquid media for 25 and 45 min, co-cultivation in solid media for 3 and 5 days, and negative control by non-infected calli. Solid and liquid co-cultivation media were made

according to Zarhloul et al. (1999). I performed GUS staining for all the co-cultivation treatments and the control calli.

For *Agrobacterium*-mediated transformation with pCas9_TPC_Red, I used a total of ~3,300 petiole explants from BO-63 (including ~200 non-transformed calli as control). In this transformation experiment, I tested the effect of sonication, calli cutting and modification of silver nitrate concentration on the regeneration rate after transformation. The investigated parameters were: transformation with or without sonication (5min, 25 Hz); with or without calli cutting; and 100mg/l to 150 and 200 mg/l silver nitrate concentration in the shooting media. Calli cutting refers to a wound of approximately 2 mm in depth made on the surface of the callus before the transformation. I evaluated the regeneration efficiency after transformation:

$$\text{Regeneration efficiency} = \frac{\text{Number of regenerated shoots}}{\text{Total number of explants used}} * 100$$

I used red fluorescence for the selection of putative transformed calli and shoots. Fluorescence was examined under a stereomicroscope (Nikon SMZ25; Nikon Instruments Europe BV, Amsterdam, the Netherlands) using the NIS-ELEMENTS BR microscope imaging software (Nikon) with the next parameters; Zoom:0,63; Mag: 0,32; exposure time: 30 s.

To find the adequate concentration of phosphinotricin (PPT) for transformant shoots selection, I tested five concentrations (0.5, 1.0, 2.0, and 2.5 mg/l) in the shooting media. I used one replicate, 20 shoots each, for every concentration, and two controls (0 mg/l of PPT).

To evaluate the selection efficiency after transformation with pCas9_TPC_Red, I assessed the putative transformants by PCR.

4.2.3. GUS staining

I performed GUS staining of the infected calli resulting from all treatments by transformation with pCambia1305.1. The calli were incubated at 37 °C for 24 h in X-Gluc solution. Later, I performed optical clearing via ethanol (70.0%).

4.2.4. DNA isolation and PCR

DNA was isolated from putatively transformed calli by the standard CTAB method (Porebski, Bailey, & Baum, 1997). PCR was performed with primer combination Cas9F+Cas9R designed at the Cas9 fragment (Supplementary Table 3) for verification of cas9 insertion under the following cycling conditions: 94°C for 3 minutes, 40 cycles of 94 °C for 30 seconds, 65 °C for 30 seconds and 72 °C for 1 minute, followed by 72 °C for 5 minutes final elongation. PCR products were separated on 1.0% agarose gels at 100 V for 30 min.

4.3. Results

4.3.1. Regeneration of whole plants from quinoa callus

Ten accessions were selected based on their days to flowering, which was recorded in the summer of 2019 in the field (Kiel, Germany) (Patirange et al., 2020). Five accessions were early flowering accessions (up to 65 days after sowing) and five were late flowering (from 80 days after sowing). For *in vitro* regeneration, I designed an experiment with four main stages: seed disinfection for the production of sterile seedlings, explant extraction, callus induction, shoot induction, and root induction.

For seed disinfection, I tested different concentrations of NaOCl and evaluated the presence of contamination, germination rate, root length, and plant height seven days after sowing. As an outcome, I successfully generated sterile seedlings for eight out of ten accessions selected for this experiment (Table 9). Accessions PI-596498, CHEN-390, and PI-510537 required a higher NaOCl concentration (3.0%). Ames 13745 showed the highest values of plant height and root length seven days after sowing.

Table 9. Results from germination experiments with ten quinoa genotypes. Plants were grown in mMS media under 16 h light (24 °C). The germination rate, plant height, and root length of the seedlings were measured 7 days after sowing. When seedling contamination was observed, NaOCl concentration was increased to 3.0%.

Accession name	NaOCl concentration (%)	<i>in vitro</i> germination rate (%)	Disinfection	Plant height (cm)	Root length (cm)
BO-13	1.0	90.5	sterile	2.33±0.15	2.90±0.36
BO-63	1.0	90.9	sterile	2.07±0.40	2.67±0.60
EMBRAPA-Brazil	1.0	65	sterile	0.93±0.06	2.00±0.26
BO-32	1.0	100	sterile	1.63±0.06	1.57±0.06
PI-596498	1.0	100	contaminated	-	-
	3.0	75	sterile	1.67±0.76	2.13±0.45
Ames 13745	1.0	68.2	sterile	4.50±0.70	3.17±0.29
CHEN-427	1.0	100	contaminated	-	-
	3.0	0	sterile	-	-
CHEN-390	1.0	75	contaminated	-	-
	3.0	28	sterile	2.70±0.36	3.27±0.97
BO-60	1.0	60	contaminated	-	-
	3.0	10	contaminated	-	-
PI-510537	1.0	90	contaminated	-	-
	3.0	77.4	sterile	3.97±0.80	3.60±0.40

Then, I continued with the five accessions that showed high germination rates and were sterile after 1% NaOCl treatment to obtain callus from sterile seedlings (Ames 136745, BO-13, BO-63, BO-32, and EMBRAPA-Brazil). In the callogenesis stage, I tested light conditions and different tissues for explant extraction. The results of explant extraction and callus induction were evaluated by callus diameter, appearance (friability), color, and callogenesis percentage per genotype. Friability was defined as described by (Ribeiro, Gayer, de Castro, Coelho, & Albarello, 2015). As result, the explant showing the best callus generation percentage and callus quality (green color and friable) was the petiole (Multi-Factor ANOVA; $p=0.05$) (Table 10). Contrary to what is described in the literature (Hesami & Daneshvar, 2016; Hesami et al., 2018; Telahigue & Toumi, 2017), I proved that light conditions are better than darkness to promote callus growth in quinoa. Accession BO-13 could not generate callus and was discarded for the next stages of the experiment.

Table 10. Callus production from three different explants and shooting rates from petiole explants. Explants and calli were cultivated under 16 h light or complete darkness. Calli were evaluated 14 days after transference to callus induction media. Friability was defined as described by Ribeiro et al. (2015). Shooting rates were evaluated six weeks after transference to shooting media.

Accession	Light Conditions	Tissue	Diameter (cm)	Callogenesis (%)	Color	Friable	Age of callus when transferred to shooting media	Shooting rate (%)
Ames 13745	16 h light	Hypocotyl	2.1±0.3	100	Purple	Yes	-	-
		Leaf	0.4±0.2	100	Green	Yes	-	-
		Petiole	0.4±0.1	100	Green	Yes	one week one month	4.54 0.00
	Darkness	Hypocotyl	0.3±0.2	40	White	No	-	-
		Leaf	0.3±0.1	80	White	No	-	-
		Petiole	-	0	-	-	-	-
BO-13	16 h light	Hypocotyl	-	0	-	-	-	-
		Leaf	0.8±0.1	71.4	Green-Yellow	Yes	-	-
		Petiole	1±0.3	100	Green	Yes	-	-
	Darkness	Hypocotyl	-	0	-	-	-	-
		Leaf	0.7±0.0	40	White	No	-	-
		Petiole	-	0	-	-	-	-
BO-32	16 h light	Hypocotyl	0.5±0.1	100	Green	Yes	-	-
		Leaf	0.7±0.3	80	Green	Yes	-	-
		Petiole	0.9±0.2	100	Green	Yes	one week one month	0.00 0.00
	Darkness	Hypocotyl	-	0	-	-	-	-
		Leaf	0.2±0.0	50	White	No	-	-
		Petiole	-	0	-	-	-	-
BO-63	16 h light	Hypocotyl	1±0.2	100	Green-Yellow	Yes	-	-
		Leaf	1.3±0.5	100	Green	Yes	-	-
		Petiole	1.5±0.5	100	Green	Yes	one week one month	25.00 4.00
	Darkness	Hypocotyl	0.6±0.1	100	White	No	-	-
		Leaf	0.7±0.0	60	White	No	-	-

		Petiole	-	0	-	-	-	-
EMBRAPA- Brazil	16 h light	Hypocotyl	1.1±0.1	100	Purple	Yes	-	-
		Leaf	0.4±0.4	50	Green-Yellow	Yes	-	-
		Petiole	1.3±0.3	100	Green-Yellow	Yes	one week	0
	Darkness	Hypocotyl	0.5±0.0	80	White	No	-	-
		Leaf	0.3±0.1	60	White	No	-	-
		Petiole	-	0	-	-	-	-

Later, I transferred the calli obtained from petiole to shooting media because petiole was the best explant for callus generation, according to the results of the callus induction test. I used one- and four-week-old calli. As an outcome, after six weeks of cultivation, BO-63 and Ames 13745 produced shoots when one-week-old calli were transferred to shooting media. When one-month-old calli were transferred to shooting media, only BO-63 produced shoots. The highest shooting rate was obtained with one-week-old calli of the accession BO-63 (Table 10).

Altogether, 28 shoots were obtained after six weeks (Supplementary Table 22 and Supplementary Table 23). The shoots were transferred to two rooting media: eight shoots to mMS media and 20 to mMS media plus activated carbon. As result, after five days, the shoots placed in mMS with 10 mg/l activated carbon underwent vitrification and death. On the contrary, shoots placed in mMS media produced roots. The first roots became visible six weeks after transferring to rooting media. After four weeks of the first root appearance, plants started flowering (*in vitro*) and then, I transferred them to 6 cm pots filled with peat. Three plants survived after transference to soil and were further grown and conserved in a growth chamber under LD conditions (16 h light, 22°C) (Supplementary Figure 24). The plants were covered with plastic containers and grown until maturity. At maturity, I harvested seeds and sowed them in 3 × 3 cm 35-multiwell palettes to grow them under greenhouse conditions. After seven days, I measured the germination rate. The regenerated plants produced seeds from self-pollination, which had a 60.0% germination rate.

4.3.2. *Agrobacterium*-mediated transformation

For *Agrobacterium*-mediated transformation, I produced sterile seedlings, extracted the petioles, and produced one-week-old calli according to the best conditions observed in the regeneration experiment.

I carried out a transformation with the vector pCambia1305.1, which uses the GUS reporter system to test the feasibility of *Agrobacterium*-mediated transformation in quinoa. I used the two accessions that showed the ability to generate shoots from callus: BO-63 and Ames 13745. I tested four co-cultivation conditions plus a negative control for every accession: co-cultivation in liquid media for 25 and 45 min, co-cultivation in solid media for 72 and 120 h, and a negative control per accession (non-infected calli). As an outcome, GUS expression was observed in 4.0% of the *Agrobacterium*-infected calli of BO-63 (Table 11 and Supplementary Figure 25). I did not observe a difference between co-cultivation time treatments.

Table 11. Results overview of *Agrobacterium*-mediated transformation in quinoa. Two vectors, two solid and liquid co-cultivation conditions, and three selection methods: GUS, red fluorescence protein (RFP), and phosphinothricin (PPT). PCR was performed to verify the insertion of the Cas cassette.

Genotype	Vector	No. of calli	Solid co-cultivation (min)	Liquid co-cultivation (h)		No. of GUS positive calli	No. of RFP positive calli	PPT resistant calli at 2.5 mg/l	PCR positive calli	Regenerated shoots with 100 mg/l silver nitrate, without sonication (%)	Regenerated shoots with sonication and 200 mg/l silver nitrate (%)
Ames 13745	pCambia1305.1	100	25	72		0	-	-	-	-	-
			45	120		0	-	-	-	-	-
BO-63	pCambia1305.1	100	25	72		2	-	-	0	-	-
			45	120		2	-	-	0	-	-
	pCas9_TPC_Red	3,100	25	72		-	10	-	0	0.4%	4.0%
		200	0	0		-	-	20	-	40.0%	47.0%

To continue, I used the vector pCas9_TPC_Red for *Agrobacterium*-mediated transformation. This vector was chosen for future possibilities of the development of a CRISPR transformation protocol in quinoa and because it does not require a destructive selection method, such as GUS staining. For this purpose, I used the accession BO-63, which depicted *In vitro* regeneration capability and showed positive GUS staining in the experiment with the vector pCambia1305.1. Moreover, the vector pCas9_TPC_Red would allow testing two methods of selection of putatively transformed tissues: red fluorescence (red fluorescent protein (RFP) expression) and phosphinothricin (PPT) selection. Accordingly, I examined red inflorescence selection at the callus and shoot stages; and PPT selection at the shooting stage.

Preliminary results by using red fluorescence selection showed RFP expression for a few calli (Supplementary Figure 26). Nevertheless, these positive events could not be confirmed by standard PCR (Supplementary Figure 27). The selection of quinoa putative transformed calli by fluorescence requires a complex calibration because the callus constitutes an irregular tissue, which in quinoa depicts a variety of colors (Table 11). Thus, the use of red fluorescence selection at the calli stage may cause false positives and/or false negatives.

Regarding selection by PPT, I tested five concentrations (0.5, 1.0, 2.0, and 2.5 mg/l) added to the shooting media to find the concentration of PPT to distinguish transformant shoots. Given the low regeneration efficiency after transformation (Table 5), I used shoots regenerated from non-infected calli for this experiment. Besides, I used two controls (0 mg/l of PPT). As an outcome, I observed that the appropriate PPT concentration for transformant shoot selection is 2.5 mg/l because, at this concentration, leaf bleaching, as well as root browning and death had a higher incidence than at other concentrations (Supplementary Figure 28). However, the presented data is preliminary and further replicates of this experiment with shoots regenerated after *Agrobacterium* infection have to be carried out.

The attempt of *Agrobacterium*-mediated transformation of calli with pCas9_TPC_Red produced tissue browning and an important reduction of the regeneration rate after transformation when compared to the control calli, which were not infected (Table 12). While the negative control calli depicted the expected 40.0% shooting rate, the transformed calli tissue strongly reduced the shooting capability to 0.4%. Hence, I attempted to increase the regeneration efficiency after transformation. For this, I tested the effect of sonication, calli cutting and modification of silver nitrate concentration on the regeneration rate after transformation because these factors have been reported to improve the regeneration efficiency after *Agrobacterium* infection (Lentz et al., 2018; Pitzschke, 2013). The tested parameters were transformation with or without sonication (5min, 25 Hz), with or without calli cutting, and silver nitrate concentrations of 100 mg/l and 200 mg/l. As an outcome, I observed that calli cutting does not favor transformation and speeds up tissue browning. Interestingly, the combination of sonication and a higher silver nitrate concentration delayed tissue browning and improved the shooting rate of BO-63 from 0.4% to 4.0% after transformation (Table 12). However, further verification of transformants by standard PCR showed no transformation.

Because the shooting rate of BO-63 was not yet sufficient for establishing an efficient transformation protocol, I aimed to find an accession with a better shooting rate after transformation with *Agrobacterium*. Therefore, I tested the protocol with sonication of 5 min at 25 Hz and 200 mg/l of silver nitrate in nine genotypes. The target shooting rate after transformation should be at least 40.0% to establish an efficient transformation protocol, similar to the existing ones (Dovzhenko, Dal Bosco, Meurer, & Koop, 2003; Maheshwari, Selvaraj, & Kovalchuk, 2011). In this way, an initial number of ~2,500 explants, as used in this experiment, might be enough to obtain a transformant plant (assuming a transformation rate of 2.0%). Unfortunately, the results showed that none of the tested genotypes depicted the desired shooting rate after transformation. I observed the highest shooting rates for Vikinga

and EMBRAPA-Brazil (~7.0%) followed by D-12289, which depicted a similar regeneration rate to BO-63 (Table 12). The given shooting rates were measured six weeks after transference to shooting media and without selection restriction.

Table 12. The shooting rate of putatively transformed quinoa calli from different accessions was measured six weeks after transformation with pCas9_TPC_Red. Shooting rate equals the number of regenerated shoots/ total number of explants used.

Accession name	Shooting rate (%)
Ames-13745	0.00
PUC-mix-green	0.00
Vikinga	7.00
EMBRAPA-Brazil	6.67
D-12289	4.00
NSL-91567	0.02
D-11889	0.00
BO-11	0.00
PI-478414	0.00

4.4. Discussion

I investigated the feasibility of *Agrobacterium*-mediated transformation in quinoa. I proved that *C. quinoa* can be regenerated *in vitro* and provided a reproducible protocol. However, I encountered the problem of tissue browning as a consequence of *Agrobacterium* infection, which strongly reduced the shooting rate. Although varying the silver nitrate concentration and adding sonication improved shooting rates, these modifications did not result in an efficient transformation protocol.

To date, robust protocols for transformation and plant regeneration have been established for a limited number of crops and varieties. Transformation remains a bottleneck in many plant species and crop genotypes because of poor transformability or recalcitrancy to established regeneration conditions. Usually, recalcitrant crops depict browning of the explant tissues, which is triggered by the accumulation and oxidation of phenolic compounds produced by wounding stress (Dixon & Paiva, 1995; Lee & Whitaker, 1995). Besides, *A. tumefaciens* interacts with host factors, which adjust the production of reactive oxygen species (ROS) and hormonal levels, disallowing regeneration after transformation (Pitzschke, 2013). For instance, a low regeneration rate due to tissue browning is one of the major limitations in sorghum, considered one of the most resistant species to genetic transformation (Dreger et al., 2019). For the callus tissue of such species to be amenable to transformation, a significant experimental effort is required to optimize tissue culture protocols (Anjanappa & Gruissem, 2021).

In the first instance, *C. quinoa* appeared recalcitrant to regeneration after *Agrobacterium* transformation. I showed that getting callus from explants does not represent a serious bottleneck, as it is for other crops (Anjanappa & Gruissem, 2021). Hence, for *C. quinoa*, changing the concentration or composition of hormonal growth regulators at the shooting stage might be enough for achieving the transformation of recalcitrant genotypes. This strategy was successful when establishing a transformation protocol for the recalcitrant cassava cultivar Verdinha (Lentz et al., 2018). Alternatively, vectors that include sequences coding for growth regulators could be tested. For instance, expressing the genes of the maize *BABY BOOM* (*BBM*) and *WUSCHEL2* (*WUS2*) transcription factors was successful in recalcitrant sorghum varieties (Nelson-Vasilchik, Hague, Mookkan, Zhang, & Kausch, 2018).

Tissue culture techniques are complex, time-consuming, and labor-intensive. Moreover, regenerated plants can exhibit trait variations that arise from mutations or genome instability during tissue culture (Fossi, Amundson, Kuppu, Britt, & Comai, 2019). Therefore, testing

transformation methods that avoid tissue culture in *C. quinoa* would be pertinent. The floral dip protocol, developed for *Arabidopsis thaliana* (Clough & Bent, 1998), would represent an alternative, for instance. Other more recent tissue culture-independent transformation strategies, which would require more significant effort than the floral dip method, could be tested. These methods include pollen and nanoparticles to deliver DNA or RNA and the *de novo* meristems induction (Lv, Jiang, Chen, & Chen, 2020; Maher et al., 2020).

Despite the lack of transformation protocols for quinoa, a recent Virus-Mediated transient expression technique (VIGS) protocol could enable functional studies in quinoa (Ogata et al., 2021). However, the achieved transformation would be transient and the reproducibility of this method is yet to be tested.

In summary, I made important advances toward the development of a protocol for *Agrobacterium*-mediated transformation in quinoa. All parameters were optimized for a whole-plant *in vitro* regeneration and several factors were tested for the improvement of the transformation step. Nevertheless, the protocol is still limited by tissue browning after transformation and significant experimental effort is required to optimize it. Future experiments might focus on two key strategies to improve the response of callus and its regenerative capacity: genotype selection of explant donors and media modifications. However, if the purpose of the study is to perform only functional analysis, it is more practical to test the reproducibility of the available VIGS protocol (Ogata et al., 2021), which will enable the exploration of molecular mechanisms governing important traits in quinoa.

4.5. References

- Anjanappa, R. B., & Gruissem, W. (2021). Current progress and challenges in crop genetic transformation. *Journal of Plant Physiology*, 261, 153411.
- Che, P., Anand, A., Wu, E., Sander, J. D., Simon, M. K., Zhu, W., . . . Liu, D. (2018). Developing a flexible, high-efficiency *Agrobacterium*-mediated sorghum transformation system with broad application. *Plant Biotechnology Journal*, 16(7), 1388-1395.
- Clough, S. J., & Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal*, 16(6), 735-743.
- Dixon, R. A., & Paiva, N. L. (1995). Stress-induced phenylpropanoid metabolism. *The Plant Cell*, 7(7), 1085.
- Dovzhenko, A., Dal Bosco, C., Meurer, J., & Koop, H. (2003). Efficient regeneration from cotyledon protoplasts in *Arabidopsis thaliana*. *Protoplasma*, 222(1), 107-111.
- Dreger, M., Mól, R., Deja, A., Raj, E., Mańkowska, G., & Wielgus, K. (2019). Improved plant regeneration in callus cultures of *Sorghum bicolor* (L.) Moench. *In Vitro Cellular Developmental Biology-Plant*, 55(2), 190-198.
- Fossi, M., Amundson, K., Kuppu, S., Britt, A., & Comai, L. (2019). Regeneration of *Solanum tuberosum* plants from protoplasts induces widespread genome instability. *Plant Physiology*, 180(1), 78-86.
- Hesami, M., & Daneshvar, M. H. (2016). Development of a regeneration protocol through indirect organogenesis in *Chenopodium quinoa* willd. *Indo Am J Agric Vet Sci*, 4(2), 25-32.
- Hesami, M., Naderi, R., & Yoosefzadeh-Najafabadi, M. (2018). Optimizing sterilization conditions and growth regulator effects on *in vitro* shoot regeneration through direct organogenesis in *Chenopodium quinoa*. *BioTechnologia. Journal of Biotechnology Computational Biology and Bionanotechnology*, 99(1).
- Keshavareddy, G., Kumar, A., & Ramu, V. S. (2018). Methods of plant transformation-a review. *International Journal of Current Microbiology and Applied Sciences*, 7(7), 2656-2668.

- Komari, T. (1990). Transformation of cultured cells of *Chenopodium quinoa* by binary vectors that carry a fragment of DNA from the virulence region of pTiBo542. *Plant Cell Reports*, 9(6), 303-306.
- Lee, C. Y., & Whitaker, J. R. (1995). *Enzymatic browning and its prevention*: ACS Publications.
- Lentz, E. M., Eisner, S., McCallum, E. J., Schlegel, K., Campos, F. d. A. d. P., Gruissem, W., & Vanderschuren, H. (2018). Genetic transformation of recalcitrant cassava by embryo selection and increased hormone levels. *Methods Protocols*, 1(4), 42.
- Lv, Z., Jiang, R., Chen, J., & Chen, W. (2020). Nanoparticle-mediated gene transformation strategies for plant genetic engineering. *The Plant Journal*, 104(4), 880-891.
- Maher, M. F., Nasti, R. A., Vollbrecht, M., Starker, C. G., Clark, M. D., & Voytas, D. F. (2020). Plant gene editing through de novo induction of meristems. *Nature Biotechnology*, 38(1), 84-89.
- Maheshwari, P., Selvaraj, G., & Kovalchuk, I. (2011). Optimization of *Brassica napus* (canola) explant regeneration for genetic transformation. *New Biotechnology*, 29(1), 144-155.
- Nelson-Vasilchik, K., Hague, J., Mookkan, M., Zhang, Z. J., & Kausch, A. (2018). Transformation of recalcitrant sorghum varieties facilitated by *BABY BOOM* and *WUSCHEL2*. *Current Protocols in Plant Biology*, 3(4), e20076.
- Ogata, T., Toyoshima, M., Yamamizo-Oda, C., Kobayashi, Y., Fujii, K., Tanaka, K., . . . Nagatoshi, Y. (2021). Virus-Mediated Transient Expression Techniques Enable Functional Genomics Studies and Modulations of Betalain Biosynthesis and Plant Height in Quinoa. *Frontiers in plant science*, 12.
- Patirange, D. S., Rey, E., Emrani, N., Wellman, G., Schmid, K., Schmöckel, S. M., . . . Jung, C. (2020). Genome-wide association study in the pseudocereal quinoa reveals selection pattern typical for crops with a short breeding history. *bioRxiv*.
- Pitzschke, A. (2013). *Agrobacterium* infection and plant defense—transformation success hangs by a thread. *Frontiers in plant science*, 4, 519.
- Porebski, S., Bailey, L. G., & Baum, B. R. (1997). Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Molecular Biology Reporter*, 15(1), 8-15.
- Ribeiro, I. G., Gayer, C. R. M., de Castro, T. C., Coelho, M. G. P., & Albarello, N. J. J. o. M. P. R. (2015). Compact callus cultures and evaluation of the antioxidant activity of *Hovenia dulcis* Thunb.(*Rhamnaceae*) under *in vivo* and *in vitro* culture conditions. 9(1), 8-15.
- Solís, J. F., Mlejnek, P., Studená, K., & Procházka, S. (2003). Application of sonication-assisted *Agrobacterium*-mediated transformation in *Chenopodium rubrum* L. *Plant Soil Environ*, 49(6), 255-260.
- Telahigue, D., & Toumi, L. (2017). Influence of medium and growth regulators on callogenesis of quinoa (*Chenopodium quinoa* Willd.) and effect of hydrous stress induced by PEG 6000 on the callus. *Horticultural Biotechnology Research*, 01-09.
- Zarhloul, M. K., Friedt, W., Khoschkhoy-Yazdi, M. R., & Liihs, W. (1999). Genetic transformation and shoot regeneration ability of resynthesised *Brassica napus* line 'RS 306'. *EUCARPIA*.

4.6. Supplementary Data

Supplementary Table 22. *Agrobacterium*-mediated transformation protocol for quinoa.

Supplementary Table 23. Summary of the dimensions and timeline of the *in vitro* regeneration experiment.

Supplementary Figure 23. Vectors used in this study. (A) pCambia1305.1: kanamycin selection against *Agrobacteria*, GUS reporter gene with catalase intron and hygromycin resistance gene. (B) pCas9_TPC_Red vector. Spectinomycin selection against *Agrobacteria* with a red fluorescence marker gene inserted in the DsRed cassette.

Supplementary Figure 24. Overview of quinoa regeneration protocol for accession BO-63's petiole explants. (A) Sterile seedlings are grown for two weeks; later, (B) petiole explants are cultured under 16 h light for one week. After six weeks, (C) shoots are obtained. Next, two weeks-old shoots are transferred to rooting media, where the (D) first roots are visible after six weeks. (E) Regenerated plants two weeks after transference to soil. (F) The regenerated plant seven weeks after transference to soil. (G) Regenerated plant reached maturity. Scale bars= 1cm.

Supplementary Figure 25. Calli from BO-63 after GUS staining. (A) Non-infected callus (control). (B) Callus showing GUS expression. *Agrobacterium*-mediated transformation was performed with the vector pCambia1305.1. Calli were grown from petiole explants under 16 h light; mMS media, 0.25 mg/l BAP + 0.50 mg/l 2,4-D. Liquid and solid co-cultivation were performed for 25 min and 72 hours, respectively. Scale bar= 5 mm.

Supplementary Figure 26. Red fluorescence selection of putatively transformed calli. From left to right: plain view of exemplary calli under white light; view under red fluorescence exposure; and a negative control constituted by non-transformed calli. The white arrow shows red fluorescence, which indicates a putatively transformed callus; exposure time: 30 s. Scale bar= 5,000 μ m.

Supplementary Figure 27. Agarose gel electrophoresis of PCR products with primers Cas9F+Cas9R for verification of Cas9 insertion. DNA was isolated from 14 putative transformed calli, which show blue staining or red fluorescence. Lane M=high range DNA ladder; lanes 1-14=calli DNA; PC=bacterial culture used for inoculation; NC=water. Expected band size of the PC: 1,735 bp. Agarose gel conditions: 1.0%, 100 V, 30 min.

Supplementary Figure 28. Non-*Agrobacterium* infected regenerated shoots from the accession BO-63 grown in mMS media + 2000 μ l BAP + 200 μ l IBA, with different concentrations of phosphinothricin: (A) 0.5 (B) 1.0 (C) 2.0 and (D) 2.5 mg/l. Photos were taken six weeks after transference to shooting media.

5. Closing discussion

Quinoa is a crop with tolerance to abiotic and biotic stresses. In addition, it exhibits unique nutritional characteristics. Because of these characteristics, this crop is considered an option to improve world food security (Alandia, Odone, Rodriguez, Bazile, & Condori, 2021). Despite the increased interest in this species in recent years, quinoa breeding is still in the initial stages. One of the major goals of quinoa breeding is to change the seasonal timing of reproduction to create unique varieties that are more suited to local conditions. Particularly, quinoa must be adapted to long-day conditions when grown in temperate climates and high latitudes such as Europe, North America, and China because this crop shows primarily short-day behavior (Patirange et al., 2020). Therefore, a better understanding of the genetic mechanisms behind photoperiod response and the timing to flower could aid in adapting quinoa to these new environments. In the present study, the solution to this problem is addressed by three strategies: QTL mapping of agronomically important traits; transcriptome analysis to identify genes that respond differentially to photoperiod and regulate flowering time; and an early insight into *Agrobacterium*-mediated quinoa transformation.

As an outcome of the first strategy, fifteen QTL contributing to the phenotypic variation of days to flowering (DTF), days to maturity (DTM), plant height (PH), panicle length (PL), panicle density (PD), seed number per plant (SN), seed weight per plant (SW), thousand kernel weight (TKW), saponin content (SC), and mildew susceptibility (MS) were mapped. This goal was achieved by phenotyping segregating F₂ plants and F₃ families in the greenhouse and field, respectively. Skim sequencing, a resourceful cost-effective method, was used to genotype the F₂ population. With the whole genome sequences of 334 F₂ plants with coverage of ~1x per plant on average, a high-density linkage map with 133,923 SNPs was effectively constructed. Therefore, this study presents the first high-density genetic map of quinoa (average density of one marker per ~8,97 Kb) that incorporates QTL for several important agronomical traits. Furthermore, two significant QTL, common in F₂ and F₃ generations, depicted pleiotropy for three highly correlated traits: DTF, PH, and TKW. Interestingly, the pleiotropic QTL harbored several putative candidate genes involved in photoperiod response and flowering time regulation.

In the second approach, where I aimed to identify candidate genes for flowering time regulation and photoperiod response by transcriptome analysis, I could find four groups of differentially expressed genes (DEGs): DEGs differentially responding to photoperiod to likely control flowering time; DEGs putatively regulating flowering time under short- and long-day conditions; and DEGs probably timing flowering at the shoot apical meristem (SAM). To investigate these groups of DEGs, I studied leaf and SAM transcriptomes of two accessions with different life cycles at different stages of development. Additionally, I investigated quinoa floral transition at the histological level. The results of the morphological evaluation precisely showed the time of transition to flowering and aided to clarify the expected expression patterns of putative candidate genes. Interestingly, the overall results of this experiment show that flowering time is likely regulated by two distinct pathways under SD and LD conditions and that two quinoa *FT* orthologues are probably important for regulating these pathways.

In the third approach, I aimed to investigate the feasibility of *Agrobacterium*-mediated quinoa transformation. In this experiment, I first successfully regenerated quinoa *in vitro* and provided a reproducible protocol for regenerating full plants from the callus of BO-63 accession. Following this, I wanted to develop a protocol for *Agrobacterium*-mediated quinoa transformation. Nonetheless, I encountered the problem of tissue browning as a consequence of *Agrobacterium* infection, which strongly reduced the shooting rate. To address this problem, I modified silver nitrate concentrations in the shooting media and tested sonication during calli transformation. Although the modification of these parameters led to a slight increase in the shooting rate, the rate was not sufficient yet to allow an efficient transformation protocol.

5.1. Two pleiotropic QTL were found in two different environments

At the beginning of this study, several questions were asked to gain a better understanding of the genetic mechanisms regulating important agronomical traits in quinoa. First, I wanted to know which QTL regulate DTF, DTM, PH, PL, TKW, and seed yield in quinoa. In the present study, the construction of a high-density genetic map allowed the detection of 15 QTL controlling ten investigated traits. However, QTL for SN and SW were mapped solely in the F_2 population; thus, the statistical power of the QTL detection might be weakened. For these traits, the lack of phenotyping in the F_3 generation prevented the measurement of environmental variances and thus significantly decreased the power and precision of QTL detection (Zhang & Xu, 2004). A similar case is given for MS, which was phenotyped only in the F_3 generation. More accurate detection of the QTL underlying MS in quinoa would require experimental replicates to calculate environmental variances. Furthermore, the power of QTL mapping for DTM in the F_3 generation may have been affected by the number of phenotyped individuals in the field. About ~60% of the individuals could not be phenotyped in the field for DTM since they did not reach maturity before the time of harvest in autumn. As environment directly influences quantitative traits, different locations and repetitions could modify the resulting QTL of my study, either in their position, or, more likely, in their confidence intervals and effects (Boopathi, 2020). The QTL found for DTM, SN, SW, and MS are more likely to change in position, R^2 and p-values, owing to the lack of replicates or individuals to phenotype. On the other hand, two pleiotropic QTL, which were common in F_2 and F_3 generations, depicted pleiotropy for three traits with high breeding relevance: DTF, PH, and TKW. This is a remarkable finding because the same QTL was found in two different environments and explained the highest phenotypic variation among all QTL (22.01%). These QTL (*pleio4.1* and *pleio14.1*) had the highest additive and dominant effects together with *pleio 4.2*. These results can be applied in practical quinoa breeding by designing molecular markers linked to these loci (*pleio4.1* and *pleio14.1*), which can guide breeders into selecting early, short, and high-yielding quinoa genotypes, simultaneously.

5.2. Histologic and transcriptomic changes towards elucidating quinoa flowering regulation

One question stated at the beginning of this study was about the time of flowering initiation stage in quinoa accessions under short-day and long-day conditions. The importance of this question lies in the critical differences observed when phenotyping days to flowering at the histological level, as compared to phenotyping plants for bolting and opening of the first flower. Studying the transition to the reproductive stage at the SAM is important to disclose the regulation of flowering time at the molecular level (Matar, Kumar, Holtgräwe, Weisshaar, & Melzer, 2021). Accordingly, to answer this question, I carried out the histological analysis of the apices of the short-day accession D-12082 and the day-neutral accession PI-614886 under SD and LD conditions, which showed that both accessions had a vegetative SAM 14 days after sowing (das) and transition to flowering happened 21 days after sowing (das) for PI-614886. I observed a 25 to 46 days difference between the initiation of the reproductive phase based on histological analysis of the SAM and days to the opening of the first flower. Hence, by analysis of the morphological changes at the SAM, I could accurately identify the time of transition to the reproductive stage in quinoa. Surprisingly, the transition to the reproductive stage in quinoa was completed 21 das, when the plants were at an early stage of development (first or second pair of leaves were visible). This case has been shown for *Bidens radiata*, which enters floral transition 72 hours after sowing with the initiation of the second pair of leaves (Halevy, 2018). Besides, it has been shown that the *CrFTL* genes of *Chenopodium rubrum* respond to flowering inductive conditions (12 h light) as early as 5 days after sowing (Cháb et al., 2008). On the other hand, as expected, my results differ from what is known from sugar beet, another member of the *Amaranthaceae* family, which requires vernalization to flower. In *B. vulgaris*, the SAM before vernalization has an “edged” shape, indicating a vegetative stage. Later, no significant morphological changes in the apical

Closing discussion

meristem can be seen during vernalization and the characteristic “dome-shape”, indicating floral transition, can be seen only 6-8 weeks after the end of vernalization (Chiurugwi et al., 2013). Importantly, I found floral transition in quinoa based on apical dominance release (“edged” shape indicates a vegetative SAM while “dome-shape” reproductive), a phenotype widely used for this purpose (Dun, Ferguson, & Beveridge, 2006). However, a more thorough histological analysis of the time of transition to the reproductive stage would be possible by comparing SAM of flowering quinoa plants against a non-flowering control, which at the moment, does not exist for quinoa.

One key question at the start of this study was if the quinoa flowering time is regulated differently under SD and LD conditions. In a previous study by our group, it was hypothesized that flowering time regulation in quinoa follows different pathways under SD and LD because we found substantial differences in the expression patterns of six *FT*-like and six *COL* homologs under different photoperiodic conditions (Patiranage et al., 2021). To confirm this hypothesis, I performed transcriptome analysis. I used a short-day accession and a day-neutral accession to generate leaf and SAM transcriptomes at different developmental stages under SD and LD conditions. As an outcome, I found 1,812 genes putatively regulating flowering time under LD (differentially expressed between accessions) and 57 genes under SD (differentially expressed between weeks of development). Remarkably, when we compared gene expression profiles between accessions under SD and LD, we found only a few DEGs in common between these photoperiod regimes. Hence, my study supports that floral transition in quinoa is regulated differently under LD and SD. Furthermore, I found that two quinoa *FT* orthologues are probably important for regulating these pathways. One of these orthologues, *CqFT1A*, might play a role as a floral repressor under LD. This finding contradicts the results reported by Patiranage et al. (2021). This difference is probably because the putative repressive function of *CqFT1A* is shown at an early stage of development (14 das) and this stage was not considered for investigation by Patiranage et al. (2021). Interestingly, a regulation of flowering time that differs between different photoperiod regimes has been observed in other crops. In rice, for instance, *HEADING DATE 3A (HD3A)* and *RICE FLOWERING LOCUS T 1 (RFT1)* act as floral integrators under SD and LD, respectively. *HD3A* is suppressed under LD while the same happens to *RFT1* under SD (Tsuji, Taoka, & Shimamoto, 2011). A further example of the independent regulation of the time to flower under different photoperiod regimes is given for soybean. In this crop, the *FT* homologs *GmFT2a* and *GmFT5a* respond differently to photoperiod, and only *GmFT5a* induces flowering under LD conditions (Kong et al., 2010). Therefore, it is reasonable to presume that different pathways and different genes may also contribute to regulating flowering time in quinoa. It is also tempting to speculate that an upstream gene differentially responding to photoperiod mediates the bifurcation of the two pathways controlling flowering time under SD and LD conditions, as of soybean. In soybean, *GmFT2a* and *GmFT5a* expression under LD is mainly controlled upstream by *PHYTOCHROME A (PHYA)* genes.

Furthermore, I inquired about the question of whether the candidate genes for photoperiod sensitivity express differently between photoperiod-sensitive and insensitive genotypes. In my study, I carried out transcriptome analysis using one day-neutral (photoperiod-insensitive) and one short-day (photoperiod-sensitive) accession. To dissect the transcriptome data, I obtained the DEGs between SD and LD for the photoperiod-sensitive accession (D-12082) followed by a comparison between both accessions in each condition (LD and SD) separately. However, detection of genes differentially responding to photoperiod was possible only in the photoperiod-sensitive accession. Differential gene expression analysis between LD and SD conditions in the photoperiod-insensitive accession would not have led to any conclusive result because PI-614886 transits to flowering at the same time under LD and SD conditions. Hence, with the experimental settings that I used in my transcriptome study, the question about differential expression of photoperiod-sensitivity candidate genes between genotypes could not be answered. It is relevant to answer this question since it is known that genotype-

specific gene isoforms have an important role in flowering time regulation. Over the past decade, investigations of key genes have identified only a few genotype-specific gene isoforms as being involved in floral transition; however, those with a role in this process are essential for post-transcriptional regulation of flowering time genes (Pose et al., 2013; Y.-Y. Wang, Xiong, Ren, & Wang, 2019). For instance, two main FLM protein splice variants, coded by *FLOWERING LOCUS M* (*FLM*), play different roles in the regulation of the time to flower in *A. thaliana*. FLM- β splicing variant acts at low temperatures and prevents early flowering. By contrast, the FLM- δ variant acts at higher temperatures as a flowering activator (Pose et al., 2013). These natural isoforms cause accelerated flowering in many *A. thaliana* accessions (Lutz et al., 2015). In my study, gene isoforms were studied by using newly-assembled transcripts, which are gene sequences assembled by the algorithm from raw reads. Such transcripts differ from the ones found in the reference genome. I assumed that interesting gene isoforms, if any, are highly co-expressed with my main candidate floral regulators. In this way, I used co-expression networks to associate newly-assembled genes with biological processes and find genotype-specific gene variants involved in flowering time regulation (Van Dam, Vosa, van der Graaf, Franke, & de Magalhaes, 2018). However, the outcome of the co-expression analysis did not show gene isoforms with sequence similarity to flowering time gene families (assessed by blastx). Thus, I concluded that no gene isoforms were involved in the photoperiod-response and flowering time regulation of the two quinoa accessions I investigated (D-12082 and PI-614886) under the analyzed conditions. Nonetheless, further experiments to address the primary question of whether the candidate genes for photoperiod sensitivity express differently between photoperiod-sensitive and insensitive genotypes might include a shift in the photoperiod regime. Consequently, plants could be grown under LD and SD conditions and then shifted to a different photoperiod regime (e.g. first under LD and later shifted to SD). Besides, it is known that changes in ambient temperature affect the splicing pattern of the genes (Park, Lee, Kim, & Park, 2019). Thus, for a deeper understanding of this matter, the experimental design might include a change in the temperature and sampling before and after this treatment.

5.3. Quinoa appears recalcitrant to *Agrobacterium*-mediated transformation

Resulting from my first two strategies (QTL mapping and transcriptome analysis), I provided lists of candidate genes regulating important agronomical processes, highlighting those which likely regulate DTF, PH, and TKW (*TK11*, *MET1b*, and *RICESLEEPER3*) pleiotropically, and two main genes that might act in independent regulation of quinoa's time to flower under SD and LD conditions (*CqFT2B* and *CqFT1A*). However, investigation of the function of these genes is essential since it will provide the ultimate evidence of their role in these biological processes. Therefore, the last question I addressed in my study was if quinoa can be transformed by the *Agrobacterium*-mediated transformation system. Although I successfully regenerated full plants *in vitro* from calli, there was a problem with extremely low shooting rates because of the *Agrobacterium* infection. To overcome this problem, I tested a different concentration of silver nitrate in the shooting media, which was increased from 100 mg/l to 200 mg/l. Simultaneously, I infected the calli with *Agrobacteria* by sonication (5min, 25 Hz) and tested calli wounding. Sonication-Assisted-*Agrobacterium*-mediated gene Transfer system (SAAT) increases transformation efficiency by introducing uniform and small cavities in the plant tissue, which allows the access of bacteria into inner plant tissue (Sedaghati, Haddad, & Bandehpour, 2021). The same principle applies to tissue wounding. On the other hand, silver nitrate restrains *Agrobacterium* growth and enables plant cell recovery, which results in increased efficiency of transformation (Opabode & Reviews, 2006). Several researchers claimed these methods were successful when transforming recalcitrant genotypes of wheat, watermelon, purslane, and cotton (Gurusaravanan, Vinoth, & Jayabalan, 2020; Opabode & Reviews, 2006; Sedaghati et al., 2021; Vasudevan, Siva, Krishnan, & Manickavasagam, 2020). In my experiments, calli wounding worsen tissue browning and the variation of silver nitrate and addition of sonication did not lead to a significant improvement in shooting rates, which would be sufficient for an efficient transformation protocol. These

results were in line with what is previously observed for recalcitrant crops, such as *B. vulgaris*. Although there are some studies related to the gene transfer in sugar beet, it is still known to be a recalcitrant species for *Agrobacterium*-mediated transformation (Gurel, 2021). Similarly, several African sorghum varieties show recalcitrant behavior and standard transformation methods that have been tested failed at improving regeneration and transformation rates (Che et al., 2018). However, answering the initial question of whether quinoa can be transformed by *Agrobacterium*-mediated transformation system requires further experiments to modify other factors that have yet to be tested. For instance, microinjection, vacuum infiltration, modifications in the acetosyringone dose, and importantly, an examination of *Agrobacterium* strains. In any case, the process of developing a transformation protocol entails extremely complex experimental designs (Kavipriya, Yuvaraja, & Senthil, 2019).

5.4. QTL mapping and transcriptome analysis provided breeding perspectives for quinoa

To conclude, my research offered several perspectives on bringing quinoa breeding into a more advanced stage. These perspectives can be mainly applied towards the necessity of adaptation of the crop to Northern latitudes. For instance, the results of QTL mapping facilitate MAS in quinoa breeding programs. Molecular markers linked to the pleiotropic QTL can be directly used in quinoa breeding programs for the simultaneous selection of different traits. These traits (DTF, PH, and TKW) are important targets in quinoa breeding, and information about the pleiotropic QTL effects could guide breeders toward the selection of early, short, and high-yielding quinoa genotypes. Knowledge of parental lines containing the markers linked to these major QTL and their effect provides the basis for gene pyramiding and might speed up genotype-building programs.

The putative floral repressive action of *CqFT1A* under LD provides a clue to facilitate quinoa adaptation to LD. If, as speculated, *CqFT1A* represses flowering in northern latitudes (LD conditions) and acts in early developmental stages (two-true-leaves stage), it might be adequate to set quinoa sowing dates as early as possible. Sowing dates as early March would avoid LD conditions at the two-leaves-visible stage; and thus, avoid *CqFT1A* repressing action. If sowing dates are to be set earlier than what is performed nowadays (late April, beginning of May), the frost tolerance of quinoa might be further exploited. The frost tolerance potential of quinoa has been confirmed for several cultivars, which did not depict frost damage at temperatures as low as -5 °C (Jacobsen et al., 2005; Jacobsen et al., 2007). Moreover, variants modifying *CqFT1A*'s expression could be of interest for growing quinoa in Northern latitudes.

Finally, I believe that there is a need to collect all reports on genetic mechanisms governing important agronomical traits in quinoa. Up to date, few but relevant efforts have been made towards the understanding of the regulation of these biological processes in this crop. In this context, a joint critical review could integrate GWAS, QTL, transcriptome, haplotype analysis, and RT-qPCR expression studies (Patiranage et al., 2021; Patiranage et al., 2020). As an early step into this extensive work, I screened candidate genes at the pleiotropic QTL *pleio4.1* and *pleio14.1* found in my study. From these genes, I selected those which were also found as candidate genes for flowering time regulation by transcriptome analysis. Later, I searched the candidate genes at the pleiotropic region identified by GWAS for days to flowering, days to maturity, plant height, and panicle length (Patiranage et al., 2020) for those obtained in my transcriptome study. As an outcome, no genes were found in common between GWAS and the transcriptome study, while 12 genes were in common between QTL and transcriptome studies (Supplementary Table 24). Among these genes, I found *VQ22*, coding for a VQ protein, which plays an important role in plant growth, development, and response to environmental conditions in Arabidopsis, most likely by acting as cofactors of WRKY transcription factors. *vq22* Arabidopsis mutants showed stunted growth (Cheng et al., 2012).

Closing discussion

The fact that these 12 genes were found as candidates in over one study provides stronger evidence of their role in controlling flowering time in quinoa. However, in-depth analyses to integrate all available data must yet be performed. For instance, raw reads from the biparental population used in my study and from the transcriptome study could be *de novo* mapped to the genomic regions associated with important agronomical traits found by GWAS (Patirange et al., 2020). This would allow a thorough study of colocalization of markers and the detection of interesting haplotypes in the biparental population and genes that are differentially expressed between a short-day and a day-neutral accession within the GWAS significant regions. Moreover, novel algorithms carry out fine integration of omics studies (e.g. summary statistics, samples correlation) (Gleason, Yang, Pierce, He, & Chen, 2020; Munz et al., 2020). In this manner, integration of the available studies would provide us with information to prioritize the selection of candidate genes for future functional studies. These functional studies could include VIGS as described by a recent publication (Ogata et al., 2021). Alternatively, screening mutants and assessing their phenotypic effects seems to be another feasible approach for determining the function of candidate genes in quinoa.

6. Summary

Chenopodium quinoa is a species native to the Andean region of South America, has unique nutritional attributes, and is well-adapted to abiotic and biotic stresses. Due to these attributes, the interest in this crop increased in the past years. However, quinoa breeding is still in its infancy. One of the major obstacles to expanding quinoa cultivation to temperate climates and high latitudes is its short-day nature. A better knowledge of photoperiod response and flowering time can aid quinoa's adaption to these new environments. In my study, I aimed to solve this problem by three approaches: QTL mapping of agronomically important traits; transcriptome analysis to identify genes differentially responding to photoperiod and involved in flowering time; and by contributing to the development of an *Agrobacterium*-mediated transformation protocol in quinoa.

In my first approach, I aimed to localize QTL contributing to the phenotypic variation of agronomically important traits. For this purpose, I used F₂ plants and F₃ families grown in the greenhouse and the field, respectively. These populations, which originated from a cross between the parental lines PI-614889 and CHEN-109, depicted large phenotypic variation in days to flowering, days to maturity, plant height, panicle length, thousand kernel weight, saponin content, and mildew susceptibility. Later, I used Skim sequencing to genotype the F₂ population and constructed a high-density genetic map with 133,923 SNPs. As result, I detected 15 QTL for ten traits. Importantly, two significant QTL, common in F₂ and F₃ generations, depicted pleiotropy for days to flowering, plant height, and TKW. The pleiotropic QTL harbored several putative candidate genes involved in photoperiod response and flowering time regulation. This study presents the first high-density genetic map of quinoa that incorporates QTL for several important agronomical traits.

In my second approach, I aimed to identify candidate genes for flowering time regulation and photoperiod response by transcriptome analysis. In this experiment, I used a short-day accession D-12082 and a day-neutral accession PI-614886. To begin, I studied quinoa floral transition at the histological level. Then, I generated the transcriptomes of the investigated accessions under two photoperiod regimes (short- and long-days) and from two tissues (leaf and shoot apical meristem) at different stages of development. As an outcome, I identified differentially expressed genes (DEGs) responding to photoperiod to likely control flowering time; DEGs putatively regulating flowering time under short- and long-day conditions; and DEGs likely timing flowering at the shoot apical meristem. Besides, the results show that flowering time is likely regulated by two distinct pathways under short- and long-day conditions and that two quinoa *FT* homologs are probably important for regulating these pathways.

As an outcome of my first two approaches, I provided an insight into the genetic mechanisms of photoperiod response and flowering time regulation in quinoa. However, functional characterization of the resulting candidate genes is essential. Thus, in my third approach, I aimed to investigate the feasibility of *Agrobacterium*-mediated quinoa transformation. In this experiment, I proved that *C. quinoa* can be regenerated *in vitro* and provided a reproducible protocol. Later, I constructed a protocol for *Agrobacterium*-mediated quinoa transformation, which still needs further optimization. Importantly, I encountered the problem of tissue browning because of *Agrobacterium* infection. This strongly reduced the shooting rates, which did not allow an efficient transformation protocol.

My study offers several future breeding perspectives. First, the pleiotropic locus identified by QTL mapping can facilitate marker-assisted selection in quinoa breeding programs. The provided information about QTL effects could guide breeders toward the selection of early, short, and high-yielding quinoa genotypes. Furthermore, given the putative function of *CqFT1A* as a floral repressor acting only under long day conditions and in the early stages of quinoa development, variants modifying *CqFT1A*'s expression could be of interest for growing quinoa in Northern latitudes.

7. Zusammenfassung

Quinoa ist eine in Südamerika beheimatete Art, die über einzigartige Ernährungseigenschaften verfügt und gut an abiotische und biotische Stressfaktoren angepasst ist. Aufgrund dieser Eigenschaften hat das Interesse an dieser Kulturpflanze in den letzten Jahren zugenommen. Die Quinoa-Züchtung steckt jedoch noch in den Anfängen. Eines der Haupthindernisse für die Ausdehnung des Quinoa-Anbaus auf gemäßigte Klimazonen und hohe Breitengrade ist das Kurztagverhalten der Pflanze. Eine bessere Kenntnis der Photoperiodenreaktion und der Blütezeit kann die Anpassung von Quinoa an diese neuen Umgebungen erleichtern. In meiner Studie habe ich mit drei Ansätzen zur Lösung dieses Problems beigetragen: QTL-Kartierung agronomisch wichtiger Merkmale; Transkriptomanalyse zur Identifizierung von Genen, die unterschiedlich auf Photoperioden reagieren und an der Blütezeit beteiligt sind; und durch einen Beitrag zur Entwicklung eines *Agrobacterium*-vermittelten Transformationsprotokolls für Quinoa.

Im ersten Ansatz ging es darum, QTL zu lokalisieren, die zur phänotypischen Variation agronomisch wichtiger Merkmale beitragen. Dazu verwendete ich F₂-Pflanzen und F₃-Familien, die im Gewächshaus bzw. auf dem Feld gezogen wurden. Diese Populationen, die von den Elternlinien PI-614889 und CHEN-109 abstammten, wiesen eine große phänotypische Variation bei den Tagen bis zur Blüte und bis zur Reife, der Pflanzenhöhe, der Rispenlänge, dem Tausendkorngewicht (TKG), dem Saponingehalt und der Mehltauanfälligkeit auf. Später verwendete ich die Skim-Sequenzierung zur Genotypisierung der F₂-Population und erstellte eine hochdichte genetische Karte mit 133,923 SNPs. Als Ergebnis konnte ich 15 QTL für zehn Merkmale nachweisen. Von diesen wiesen zwei QTL, die in den F₂- und F₃-Generationen vorkommen, eine Pleiotropie für die Tage bis zur Blüte, die Pflanzenhöhe und das TKG auf. Die pleiotropen QTL beherbergten mehrere mutmaßliche Kandidatengene, die an der Photoperiodenreaktion und der Regulierung der Blütezeit beteiligt sind.

Im zweiten Ansatz wollte ich durch Transkriptomanalyse Kandidatengene für die Regulierung der Blütezeit und die Photoperiodenreaktion identifizieren. Für diesen Versuch verwendete ich den Kurztag-Genotyp D-12082 und den tagneutralen Genotyp PI-614886. Zunächst untersuchte ich die Entwicklung der Quinoa-Blüte auf histologischer Ebene. Dann erstellte ich die Transkriptome der untersuchten Akzessionen unter zwei Photoperiodenregimen und aus zwei Geweben (Blatt und Sprossapikalmeristem) in verschiedenen Entwicklungsstadien. Als Ergebnis fand ich differentiell exprimierte Gene (DEGs), die unterschiedlich auf die Photoperiode reagieren, um den Blühzeitpunkt zu steuern; DEGs, die den Blühzeitpunkt unter Kurz- und Langtagbedingungen regulieren; und DEGs, die den Blühzeitpunkt am sprossapikalen Meristem steuern. Die Ergebnisse zeigen, dass die Blütezeit unter Kurz- und Langtagbedingungen wahrscheinlich durch zwei verschiedene Wege reguliert wird und dass zwei *FT*-Orthologe in Quinoa wahrscheinlich für die Regulierung dieser Wege wichtig sind.

Eine funktionelle Charakterisierung der aus meinen ersten beiden Ansätzen resultierenden Kandidatengene ist jedoch unerlässlich. Im dritten Ansatz wollte ich daher die Machbarkeit der *Agrobacterium*-vermittelten Quinoa-Transformation untersuchen. In diesem Experiment konnte ich nachweisen, dass *C. quinoa in vitro* regeneriert werden kann. Dafür habe ich ein reproduzierbares Protokoll erstellt. Später erstellte ich ein Protokoll für die *Agrobacterium*-vermittelte Quinoa-Transformation, das noch weiter optimiert werden muss.

Meine Studie bietet Perspektiven für die zukünftige Züchtung. Erstens können die Ergebnisse der QTL-Kartierung die markerunterstützte Selektion in Zuchtprogrammen erleichtern. Darüber hinaus könnten die Informationen über QTL-Effekte die Züchter bei der Selektion von frühen, kurzen und ertragreichen Quinoa-Genotypen unterstützen. Angesichts der vermuteten Funktion von *CqFT1A* als Blütenrepressor, der nur unter Langtagbedingungen und in frühen Entwicklungsstadien von Quinoa wirkt, könnten Sequenzvarianten mit veränderter Genaktivität für den Anbau in nördlichen Breitengraden nützlich sein.

8. Appendix

8.1. Supplementary Tables

Supplementary Table 1. Plant material used in QTL mapping. DTF: days to flowering, DTM: days to maturity, PH: plant height, PL: panicle length, PD: panicle density, TKW: thousand kernel weight, SW: seed weight per plant, SN: seed number per plant, SC: Saponin content, MS: Mildew susceptibility.

Seed code	Generation	No. of plants/families	No. of plants genotyped	No. of plants phenotyped	
				Greenhouse	Field
171115	PI-614889 (seed parent)	-	-	10	20
170876	CHEN-109 (pollen parent)	-	-	10	20 (DTM, SW and SN were not recorded)
190031	F ₂	336	336	336 except for MS	-
191203-191562	F ₃	334	-	-	DTF: 5,891 DTM: 2,343 PH: 5,860 PL: 5,860 PD: 5,860 MS: 6,346 SC: 330 families were bulked together (out of 334 sown families, four were lost due to biotic stress after germination) TKW: 330 families were bulked together (out of 334 sown families, four were lost due to biotic stress after germination) SW: 0 SN: 0

Supplementary Table 2. Methods for phenotypic evaluation (modified from Stanschewski et al., 2021).

Trait	Acronym	Description
Days to flowering	DTF	Number of days, from sowing until the first flower opens. Recorded thrice a week in the F ₂ and twice a week in the F ₃ population.
Days to maturity	DTM	Number of days, from sowing until the panicle is completely brown. Recorded twice a week.
Plant height	PH	Distance in cm from root collar to panicle apex recorded once at BBCH-81 to BBCH-89.
Panicle length	PL	Distance in cm from panicle base to tip recorded once at BBCH-81 to BBCH-89.
Panicle density	PD	Visual score based on the observed range of panicle density recorded at BBCH-81 to BBCH-89. The score ranges from 1 to 7: (1) “Low”: loose panicle with low number of spaced glomerules and panicle axes clearly visible; (3) “Intermediate”: high number of glomerules tightly arranged with panicle axis often visible; (5) “High”: high number of glomerules tightly packed and scarcely seen panicle axes; and (7) “Very high”: very high number of glomerules compactly arranged, panicle axes not visible. Recorded once at BBCH-81 to BBCH-89. An example of panicle density scoring in the F ₂ population can be observed in Supplementary Figure 3.
Mildew susceptibility	MS	Performed only in the F ₃ population (incidence only in the field). Performed as a visual score ranging from 1 to 3: (1) “High”: symptoms observed in the whole plant or at least up to the upper third of the plant, (2) “Intermediate”: symptoms observed in the lowest two-thirds of the plant, (3) “Low”: no visible symptoms or only present in the lower third of the plant. Recorded once at BBCH-81 to BBCH-89.
Saponin content	SC	Recorded after harvest by foam test. Twenty seeds were placed in a 2 ml epi, shaken for 1 min at 1100 g in a GenoGrinder and incubated at room temperature for 5 min; later, the height of the formed foam (cm) was recorded (Jarvis et al., 2017).
Seed weight per plant	SW	Weight of the total harvest (whole panicle) per plant. Recorded once at BBCH-99.
Number of seeds per plant	SN	Number of seeds contained in the whole panicle per plant. Recorded once at BBCH-99.
Thousand Kernel Weight	TKW	Recorded after harvest.

Supplementary Table 3. Primers used in this study.

Purpose	Gene	Gene ID	Primer ID	Primer sequence 5'-3'	Amplification product size (bp)	Annealing temperature (°C)
RT-qPCR	<i>CqHD3AB-1</i>	CQ000922	NMQ121+NMQ122	F: CGATCCTACGGCAACTTATTAC CTGATGCTTCCAGTCGCCG	127	60
	<i>CqHD3AB-2</i>	CQ019992	NMQ125+NMQ126	F: GCCCTCTCAGGTTGTTAATCAAC R: GGATTACTTGGGCTAGGAGC	106	58
	<i>CqFIP1</i>	CQ039979	NMQ129+NMQ130	F: GCAGCTGCAAAGAAGTGGTG R: CATTAGCTGCACATCGGAAGC	145	58
	<i>CqELF5</i>	CQ038819	NMQ133+NMQ134	F: GAAGGAAAAGGGTGAGGATCCAG R: GAGCTCCTGTGGGATTCAACG	144	60
	<i>CqSOC1</i>	CQ047524	NMQ101+NMQ102	F: GCTAGAGAGCCAGTTGG R: GCCTAGTATTTTCTGCAGC	120	57
	<i>CqADO1</i>	CQ012973	NMQ111+NMQ112	F: GCTTGTGCAGTCGGTAATCGGG R: CGACCTGGTGGTGGAGAGC	143	60
	<i>CqTSP9</i>	CQ000664	NMQ113+NMQ114	F: GCTGTCGTTGCAAGGGGTC R: CAGATCTGTCTGTCCTCCCAATGC	144	61
	<i>CqIDH-A</i>	CQ044820	NMQ057+NMQ058	F: GGACGGACTATTGAAGCTGAAGC R: GCATGGACACGTGGGCTTGC	125	60
	<i>CqPTB</i>	CQ013990	NMQ070+NMQ071	F: GCAGTGCAACAGGCTCCAAG R: GGCACCTGCACCCTCATG	143	57
	Transgene verification	pCas9-TPC <i>bar</i> cassette	Cas9F+ Cas9R	F: CAGTCTTTCACCTCTCTTTGG R: CCATCTTTGGGACCACTGTC	1,735 bp	65
				F: AGCCATTGCACTATGCCCTCTC R: TGGCCCAACACCTAAGTGACG		
	Genetic segregation verification	InDel marker JASS5 (Zhang et al., 2017)	DSQ140+DSQ141	F: AGCCATTGCACTATGCCCTCTC R: TGGCCCAACACCTAAGTGACG	164 or 189 bp	60

Supplementary Table 4. Polymerase chain reaction (PCR) details.

Protocol	Description
Master mix preparation	For a final volume of 20 µl per reaction: 15.9 µl of water, 2.0 µl of PCR buffer (10x), 0.4 µl of dNTP mix (10 Mm), 0.3 µl of forward and reverse primer (10 µM), 0.1 µl of Taq Polymerase (%U/µl), 1 µl of DNA template.
PCR amplification regime	5 min at 94°C, 35x (30 s at 94°C, 30 s at annealing temperature, 60 s at 72°C) and 5 min at 72°C. PCRs were performed in a LifeTouch Thermal Cycler (Biozym Scientific GmbH, Hess. Oldendorf, Germany).

Supplementary Table 5. Allele and genotype nomenclature used in this study.

Gene name	Location (chromosome_ physical position)	Allele name		Allele sequence		Variant type	Amino acid change
		PI-614889	CHEN-109	PI-614889	CHEN-109		
<i>HEN2</i>	chr4_56438962	A_1	A_2	A	T	Intronic SNP	-
<i>SUF4</i>	chr12_80423460	B_1	B_2	A	G	3' prime UTR SNP	-
<i>RUG3</i>	chr12_80551460	C_1	C_2	C	T	3' prime UTR SNP	-
<i>RUG3</i>	chr12_80552842	D_1	D_2	C	T	Missense SNP	p.Ala388Thr
<i>RUG3</i>	chr12_80553379	E_1	E_2	A	G	Intronic SNP	-
<i>TKII</i>	chr12_81633247	F_1	F_2	A	T	Intronic SNP	-
<i>WRKY13</i>	chr12_81728382	G_1	G_2	C	A	Missense SNP	p.Glu117Asp
<i>ATHB-15/CORONA</i>	chr4_51023634	H_1	H_2	G	A	Missense SNP	p.Asp22Asn
<i>HEN2</i>	chr4_56441565	I_1	I_2	T	C	Missense SNP	p.Ile724Met
<i>NPF2.6</i>	chr4_53557989	J_1	J_2	C	T	Missense SNP	p.Ala93Val
<i>ERF113</i>	chr12_81516821	K_1	K_2	G	A	Missense SNP	p.Pro326Leu
<i>FLD</i>	chr4_56844370	L_1	L_2	A	AT	3' UTR InDel	-
<i>FLD</i>	chr4_56847562	M_1	M_2	$TGAAC$	T	Intronic InDel	-
<i>TKII</i>	chr12_81633685	N_1	N_2	T	A	Missense SNP	p.Gln445Leu
<i>MET1b</i>	chr4_56534732	O_1	O_2	ATT	A	Frameshift	p.Gln222fs
<i>MET1b</i>	chr4_56534915	P_1	P_2	$AGTT$	A	Disruptive Inframe Deletion	p.Lys161_Leu162delinsMet
<i>RICESLEEPER3</i>	chr4_55091902	Q_1	Q_2	A	$AATTCCT$	Disruptive Inframe Insertion	p.Ile281_Thr282insProLeu

Supplementary Table 6. Genetic and phenotypic segregation for two traits in the F₂ and F₃ populations. Red axil pigmentation was determined five weeks after sowing. The InDel marker JAASS5 was described by Zhang et al. (2017). *R*₁ and *R*₂ represent the 189 bp and 164 bp alleles, respectively.

Plants	F ₂ population							F ₃ population			
	Red axil pigmentation		χ^2 ^a	JASS5 genotype			χ^2 ^b	JASS5 genotype			χ^2 ^c
	Red	Green		<i>R</i> ₁ <i>R</i> ₁	<i>R</i> ₁ <i>R</i> ₂	<i>R</i> ₂ <i>R</i> ₂		<i>R</i> ₁ <i>R</i> ₁	<i>R</i> ₁ <i>R</i> ₂	<i>R</i> ₂ <i>R</i> ₂	
O	254	82	0.06	8	28	12	2.00	58	57	79	5.02
E	252	84		12	24	12		72.75	48.50	72.75	

E: expected, O: observed.

^a3:1 segregation, $\chi^2_{(0.95;1)} = 3.84$

^b1:2:1 segregation, $\chi^2_{(0.95;2)} = 5.99$

^c3:2:3 segregation, $\chi^2_{(0.95;2)} = 5.99$

Appendix

Supplementary Table 7. Complete list of genes located within the non-overlapping QTL confidence intervals (*This table is provided with the supplementary data DVD, due to the excessive length*).

Supplementary Table 8. Genes within *pleio4.1* and *pleio14.1* confidence intervals with a previously described function related to flowering time, photoperiod response or yield regulation and their gene variants. Gene variants between CHEN-109 (pollen parent) and PI-614889 (seed parent) and the reference genome QQ74_V2 are shown (*This table is provided with the supplementary data DVD, due to the excessive length*).

Appendix

Supplementary Table 9. Sampling stages and tissues from D-12082 and PI-614886 under two different photoperiod regimes. Samples were taken for RNA-sequencing (RNA-seq) and histological analysis. Plants were grown at 22 °C under long-day (LD) conditions: 16 h light; and short-day conditions (SD): 8 h light. Sequenced samples are written in red.

Sampling date[§]	Accession	Photoperiod condition	Purpose	Tissue	Number of samples per week[§]
W2, W3, W4, W5, W7, W8	D-12082	SD	RNA-seq	Leaf	14
W2, W3, W4, W5, W7, W8, W9, W10	D-12082	LD	RNA-seq	Leaf	14
W2, W3, W4	D-12082	SD	RNA-seq	Apical meristem	40 to 56
W2, W3, W4, W5, W6	D-12082	LD	RNA-seq	Apical meristem	40 to 56
W2, W3, W4	D-12082	SD	Histology analysis	Apex	7
W2, W3, W4, W5, W6	D-12082	LD	Histology analysis	Apex	7
W2, W3, W4, W5, W7	PI-614886	SD	RNA-seq	Leaf	14
W2, W3, W4, W5, W7	PI-614886	LD	RNA-seq	Leaf	14
W2, W3, W4	PI-614886	SD	RNA-seq	Apical meristem	40 to 56
W2, W3, W4	PI-614886	LD	RNA-seq	Apical meristem	40 to 56
W2, W3, W4	PI-614886	SD	Histology analysis	Apex	7
W2, W3, W4	PI-614886	LD	Histology analysis	Apex	7

[§] W=weeks after sowing

[§] including biological replicates

Supplementary Table 10. Quality filtering and raw read statistics for the transcriptome libraries. Three biological replicates of shoot apical meristem and leaf tissue from two accessions grown under two different conditions were sequenced. Samples were collected weekly. Long-day (LD) conditions: 16 h light, 22°C; short-day conditions (SD): 8 h light, 22°C. PE= pair-end.

RNA-seq Library	Accession	Tissue	Growth condition	Week	Biological replicate	No. of raw reads (millions)	Clean reads/Raw reads (%)	No. of uniquely-mapped PE reads (millions)	No. of multiple-mapped PE reads (millions)
R101	D-12082	Leaf	SD	2	1	70.22	98.35	31.49	1.69
R102	D-12082	Leaf	SD	2	2	58.71	98.72	26.32	1.46
R103	D-12082	Leaf	SD	2	3	76.16	98.43	34.10	1.90
R104	D-12082	Leaf	LD	2	1	62.11	97.60	27.12	1.68
R105	D-12082	Leaf	LD	2	2	66.65	98.74	29.86	1.74
R106	D-12082	Leaf	LD	2	3	67.90	97.32	29.62	1.75
R107	PI-614886	Leaf	SD	2	1	60.82	98.72	27.48	1.42
R108	PI-614886	Leaf	SD	2	2	62.39	98.47	36.98	1.92
R109	PI-614886	Leaf	SD	2	3	62.22	98.58	28.25	1.43
R110	PI-614886	Leaf	LD	2	1	60.68	98.46	34.44	1.80
R111	PI-614886	Leaf	LD	2	2	80.09	97.24	41.89	2.09
R112	PI-614886	Leaf	LD	2	3	81.34	98.95	45.09	2.25
R201	D-12082	Leaf	SD	3	1	90.45	98.90	40.53	2.18
R202	D-12082	Leaf	SD	3	2	84.11	98.68	37.64	2.05
R203	D-12082	Leaf	SD	3	3	61.23	99.04	34.17	1.84
R204	D-12082	Leaf	LD	3	1	64.60	98.90	29.09	1.63
R205	D-12082	Leaf	LD	3	2	63.40	99.45	28.28	1.93
R206	D-12082	Leaf	LD	3	3	61.94	98.56	35.87	2.53

R207	PI-614886	Leaf	SD	3	1	62.71	99.22	28.27	1.79
R208	PI-614886	Leaf	SD	3	2	68.46	99.00	38.65	2.90
R209	PI-614886	Leaf	SD	3	3	78.85	98.94	38.38	1.91
R210	PI-614886	Leaf	LD	3	1	65.13	99.07	29.42	1.64
R211	PI-614886	Leaf	LD	3	2	76.17	98.58	34.41	1.75
R212	PI-614886	Leaf	LD	3	3	81.66	98.65	47.61	2.99
R301	D-12082	Leaf	SD	4	1	82.81	98.95	37.23	1.88
R302	D-12082	Leaf	SD	4	2	74.49	98.88	33.55	1.79
R303	D-12082	Leaf	SD	4	3	73.62	98.57	32.67	1.65
R304	D-12082	Leaf	LD	4	1	71.79	98.18	40.83	2.05
R305	D-12082	Leaf	LD	4	2	78.11	98.07	38.73	2.38
R306	D-12082	Leaf	LD	4	3	71.73	98.65	33.53	1.73
R307	PI-614886	Leaf	SD	4	1	73.83	98.44	33.02	1.79
R308	PI-614886	Leaf	SD	4	2	66.17	98.83	35.77	1.78
R309	PI-614886	Leaf	SD	4	3	65.90	98.64	29.67	1.51
R310	PI-614886	Leaf	LD	4	1	59.96	98.54	38.96	1.97
R311	PI-614886	Leaf	LD	4	2	65.70	98.06	29.39	1.48
R312	PI-614886	Leaf	LD	4	3	72.24	97.56	31.79	1.84
R401	D-12082	Leaf	SD	5	1	66.56	98.99	30.29	1.48
R402	D-12082	Leaf	SD	5	2	61.98	98.77	28.10	1.36
R403	D-12082	Leaf	SD	5	3	68.11	98.60	30.77	1.56
R404	D-12082	Leaf	LD	5	1	74.19	98.00	33.09	1.67

R405	D-12082	Leaf	LD	5	2	64.69	98.24	43.00	2.35
R406	D-12082	Leaf	LD	5	3	66.27	98.66	29.82	1.61
R407	PI-614886	Leaf	SD	5	1	62.32	98.53	28.30	1.47
R408	PI-614886	Leaf	SD	5	2	59.64	98.75	37.66	1.80
R409	PI-614886	Leaf	SD	5	3	68.40	98.88	41.32	2.01
R410	PI-614886	Leaf	LD	5	1	71.91	98.59	32.81	1.63
R411	PI-614886	Leaf	LD	5	2	58.95	98.75	26.91	1.34
R412	PI-614886	Leaf	LD	5	3	63.77	98.81	29.14	1.42
Rm101	D-12082	SAM	SD	2	1	34.20	98.71	40.02	1.84
Rm102	D-12082	SAM	SD	2	2	33.11	98.82	29.59	1.30
Rm103	D-12082	SAM	SD	2	3	31.44	96.05	27.06	1.22
Rm104	D-12082	SAM	LD	2	1	34.88	97.39	30.84	1.42
Rm105	D-12082	SAM	LD	2	2	39.68	96.66	34.26	1.63
Rm106	D-12082	SAM	LD	2	3	32.34	97.41	28.15	1.29
Rm107	PI-614886	SAM	SD	2	1	34.03	96.99	30.12	1.44
Rm108	PI-614886	SAM	SD	2	2	37.05	97.66	33.04	1.49
Rm109	PI-614886	SAM	SD	2	3	35.08	97.45	30.74	1.42
Rm110	PI-614886	SAM	LD	2	1	33.45	98.12	29.87	1.37
Rm111	PI-614886	SAM	LD	2	2	34.37	98.59	31.03	1.47
Rm112	PI-614886	SAM	LD	2	3	35.08	97.46	31.09	1.33
Rm201	D-12082	SAM	SD	3	1	31.62	98.85	31.12	1.40
Rm202	D-12082	SAM	SD	3	2	34.41	97.69	29.92	1.39

Rm203	D-12082	SAM	SD	3	3	32.33	98.60	33.48	1.56
Rm204	D-12082	SAM	LD	3	1	38.29	98.52	29.61	1.37
Rm205	D-12082	SAM	LD	3	2	34.90	97.47	30.11	1.39
Rm206	D-12082	SAM	LD	3	3	34.93	98.00	30.48	1.46
Rm207	PI-614886	SAM	SD	3	1	36.74	98.09	32.75	1.69
Rm208	PI-614886	SAM	SD	3	2	30.11	98.80	27.06	1.29
Rm209	PI-614886	SAM	SD	3	3	31.46	98.56	28.00	1.29
Rm210	PI-614886	SAM	LD	3	1	30.46	98.68	27.82	1.42
Rm211	PI-614886	SAM	LD	3	2	34.91	98.77	30.25	1.50
Rm212	PI-614886	SAM	LD	3	3	31.46	99.08	27.52	1.31

Supplementary Table 11. List of candidate genes putatively regulating flowering time in quinoa with an associated flowering time function in quinoa and other species.

Gene name	Gene ID[§]	Chromosome	Start	End
<i>CqFT1A</i>	CQ052447	Cq9A	51593659	51597564
<i>CqHD3AB</i>	CQ000922	Cq6B	14735351	14740133
<i>CqFT2B</i>	CQ019992	Cq4B	3497140	3503985
<i>CqCOL4A-1</i>	CQ037524	Cq4A	54461001	54462646
<i>CqCOL4A-2</i>	CQ052851	Cq2A	1028973	1031191
<i>CqCOL16A-1</i>	CQ027262	Cq6A	49409802	49412567
<i>CqCOL16A-2</i>	CQ030058	Cq5A	14795389	14799196
<i>CqFIP1</i>	CQ039979	Cq8A	48505065	48517240
<i>CqFRL4AA</i>	CQ045556	Cq3A	23432860	23438577
<i>CqFRL4AB</i>	CQ018308	Cq3B	49944175	49954824
<i>CqELF5</i>	CQ038819	Cq8A	10430182	10443713
<i>CqCO3A</i>	CQ044372	Cq3A	5402299	5406332
<i>CqCO3B</i>	CQ017032	Cq3B	55714022	5575425
<i>CqTPS9</i>	CQ000664	Cq6B	8195253	8203248
<i>CqPHYB</i>	CQ043833	Cq3A	559497	568173
<i>CqZTL</i>	CQ012973	Cq8B	74028153	74035097
<i>CqBIC1</i>	CQ043414	Cq7A	52591665	52597118
<i>CqSOC1</i>	CQ047524	Cq1A	6162395	6169396

[§] According to the reference genome QQ74_V2 (CoGe Genome ID: id60716)

Supplementary Table 12. Gene Ontology (GO) annotation of differentially expressed genes (DEGs) between SD and LD in leaves of D-12082, which differentially respond to photoperiod to likely regulate flowering time. Number of genes in each node of different GO hierarchy levels are shown for three GO categories, Biological Process, Cellular Component, and Molecular Function (*Supplementary Tables 12 to 21 are provided in .xlsx with the supplementary data DVD*).

Supplementary Table 13. KEGG classification of differentially expressed genes (DEGs) between SD and LD in leaves of D-12082, which differentially respond to photoperiod to likely regulate flowering time (*Supplementary Tables 12 to 21 are provided in .xlsx with the supplementary data DVD*).

Supplementary Table 14. Gene Ontology (GO) annotation of differentially expressed genes (DEGs) between PI-614886 and D-12082 at W2 and W3 under long-days (LD) in leaves, which are putatively controlling flowering under LD (16 h light). Number of genes in each node of different GO hierarchy levels are shown for three GO categories, Biological Process, Cellular Component, and Molecular Function (*Supplementary Tables 12 to 21 are provided in .xlsx with the supplementary data DVD*).

Supplementary Table 15. KEGG classification of differentially expressed genes (DEGs) between PI-614886 and D-12082 at W2 and W3 under long-days (LD) in leaves, which are putatively controlling flowering under LD (16 h light) (*Supplementary Tables 12 to 21 are provided in .xlsx with the supplementary data DVD*).

Supplementary Table 16. Gene Ontology (GO) annotation of differentially expressed genes (DEGs) with similar spatial expression patterns in leaves between PI-614886 and D-12082, which likely regulate flowering time under short-days (8 h light). Number of genes in each node of different GO hierarchy levels are shown for three GO categories: Biological Process, Cellular Component, and Molecular Function (*Supplementary Tables 12 to 21 are provided in .xlsx with the supplementary data DVD*).

Supplementary Table 17. KEGG classification of differentially expressed genes (DEGs) with similar spatial expression patterns between PI-614886 and D-12082 under short-days (SD) in leaves, which likely regulate flowering time under short-days (8 h light) (*Supplementary Tables 12 to 21 are provided in .xlsx with the supplementary data DVD*).

Supplementary Table 18. Gene Ontology (GO) annotation of differentially expressed genes (DEGs) in the shoot apical meristem (SAM) between PI-614886 and D-12082 at W2 and W3, which are likely involved in the regulation of the time to flower at the SAM. Number of genes in each node of different GO hierarchy levels are shown for three GO categories: Biological Process, Cellular Component, and Molecular Function (*Supplementary Tables 12 to 21 are provided in .xlsx with the supplementary data DVD*).

Supplementary Table 19. KEGG classification of differentially expressed genes (DEGs) in the shoot apical meristem (SAM) between PI-614886 and D-12082 at W2 and W3, which are likely involved in the regulation of the time to flower at the SAM (*Supplementary Tables 12 to 21 are provided in .xlsx with the supplementary data DVD*).

Supplementary Table 20. Results of the blastx analysis for newly-assembled transcripts that are co-expressed with annotated DEGs that putatively regulate flowering in leaves. The top 10% of newly-assembled transcripts with the highest degree and module membership (MM) > 0.8 from each module in the co-expression hierarchical clustering were used for blastx. Newly-assembled transcripts are those sequences assembled as transcripts by StringTie, which are not found as transcripts in the reference genome QQ74_V2 (*Supplementary Tables 12 to 21 are provided in .xlsx with the supplementary data DVD*).

Supplementary Table 21. Gene Ontology (GO) annotation of the newly-assembled transcripts that are co-expressed with annotated DEGs that putatively regulate flowering in leaves. Number of genes in each node of different GO hierarchy levels are shown for three GO categories: Biological Process, Cellular Component, and Molecular Function. Newly-assembled transcripts are those sequences assembled as transcripts by StringTie, which are not found as transcripts in the reference genome QQ74_V2 (*Supplementary Tables 12 to 21 are provided in .xlsx with the supplementary data DVD*).

Supplementary Table 22. *Agrobacterium*-mediated transformation protocol for quinoa.

Step	Description
Seed disinfection and sowing	<ol style="list-style-type: none"> 1. Place the seeds in >99% ethanol and shake for 5 seconds 2. Replace the ethanol with 3% Sodium hypochlorite and two drops of Tween 20 (two drops per 50 seeds) 3. Shake at room temperature for 20 min (do not exceed this time) 4. Replace the sodium hypochlorite with water, shake for 5 seconds 5. Replace the water with fresh water, after that shake for 5 seconds and rinse 6. Place the seeds in mMS medium 7. Wait for 21 to 28 days for the seedlings. Climate room at 24°C, 16h light/ 8h dark.
Callogenesis	<ol style="list-style-type: none"> 1. After 29 to 32 days obtain the petiole explants. Cut the petioles in segments of ~1 cm each. Avoid including leaf fragments in the segments. 2. Place the segments on callogenesis medium. 3. Wait for 21 to 30 days until the calli appearance. Climate room at 24°C, 16h light/ 8h dark.
Agrobacterium cultivation	<ol style="list-style-type: none"> 1. One day prior to transformation: inoculate 15 ml LB [+ 3 µl rifampicin (25 mg/ml) + 3 µl gentamycin (50 mg/ml) + selection of pCas9: 3 µl spectinomycin (100 mg/ml)] with 200 µl storage culture. 2. Incubate for 16 h at 28°C, 200 rpm on a rotary shaker. 3. Measure OD₆₀₀ (ideally it equals 1.2). Centrifuge the <i>Agrobacteria</i> 3000rpm, 10 min. Discard supernatant and resuspend in 40 ml co-cultivation medium O2 to reach an OD₆₀₀ = 0.3-0.5. Shake at 28°C for up to 2 h to dissolve the pellet. Use this solution directly for transformation.
Transformation	<ol style="list-style-type: none"> 1. Immerse the petioles with calli in liquid co-cultivation medium (Dr. José Orsini, Saaten-Union Biotec GmbH, Germany) including dissolved <i>Agrobacteria</i>. 2. Co-cultivate for 25 min in liquid medium. 3. Place the infected segments on solidified co-cultivation medium and incubate for further 3 days in the climate room at 24°C, 16h light/ 8h dark.
Shooting and selection against agrobacteria	<ol style="list-style-type: none"> 1. Place the petioles with calli on a shooting medium containing antibiotics for selection against agrobacteria. 2. Check two times per week for generation of shoots (and up to six weeks). 3. Once a shoot is detected it should be submitted to selection for transgenic plantlets and transferred to rooting media within 4-7 days after appearance otherwise the shoots go to vitrification-callus process. The shoots will not grow bigger than 2 cm in SM media.
Rooting	<ol style="list-style-type: none"> 1. After visual verification of the presence of the reporter gene, place the shoots on an mMS medium. 2. In the first two weeks, the shoot will depict slow growth and new leaves will be generated. Wait for ~6 weeks until root regeneration.

Supplementary Table 23. Summary of the dimensions and timeline of the *in vitro* regeneration experiment.

Accession	Sowing date	Sown seeds	Date of transfer to callogenesis medium	No. of explants ¹	Date of transfer to shooting medium	No. of shoots	Date of transfer to rooting medium
BO-13	29.03.2019	~50	A: 12.04.19	~150	14.05.2019	0	
	30.04.2019	~50	B: 03.05.19	~100	03.06.2019	0	
BO-63	29.03.2019	~50	A: 12.04.19	~150	14.05.2019	3	23.05.2019
	30.04.2019	~50	B: 03.05.19	~100	03.06.2019	2	23.05.2019
					13.05.2019	21	23.08.2019
Embrapa	29.03.2019	~50	A: 12.04.19	~150	14.05.2019	0	
	30.04.2019	~50	B: 03.05.19	~100	03.06.2019	0	
BO-32	29.03.2019	~50	A: 12.04.19	~150	14.05.2019	0	
	30.04.2019	~50	B: 03.05.19	~100	03.06.2019	0	
PI-596498	29.03.2019	~50	B: 03.06.19	~100	11.06.2019	0	
	15.04.2019	~50					
	23.05.2019	~50					
Ames 13745	25.01.2019	~50	A: 12.04.19	~150	14.05.2019	0	23.08.2019
	30.04.2019	~50	B: 03.05.19	~100	03.06.2019	2	
CHEN-427	29.03.2019	~50	B: 03.06.19	~100	11.06.2019	0	
	15.04.2019	~50					
	23.05.2019	~50					
CHEN-390	29.03.2019	~50	B: 14.05.19	~100	C		
	15.04.2019	~50					
BO-60	29.03.2019	~50	B: 03.06.19	~100	11.06.2019	0	
	15.04.2019	~50					
	23.05.2019	~50					
PI-510537	29.03.2019	~50	B: 03.06.19	~100	11.06.2019	0	
	15.04.2019	~50					
	23.05.2019	~50					

A: 2 weeks-old seedlings; B: 1-month-old seedlings, C: contamination.

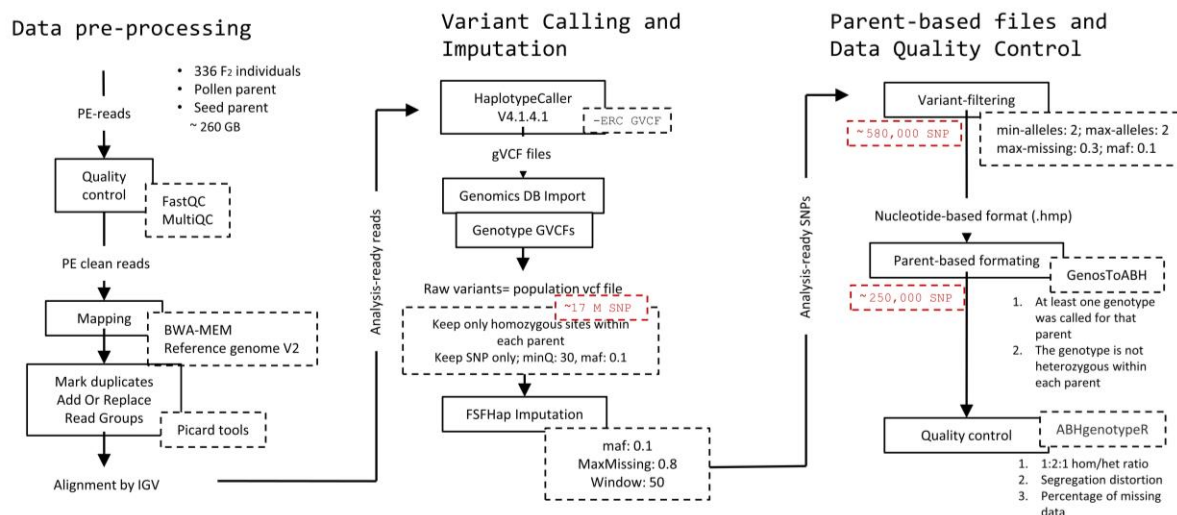
¹From petiole, leaves, and hypocotyl

Supplementary Table 24. Genes found as candidates by both, QTL and transcriptome studies.

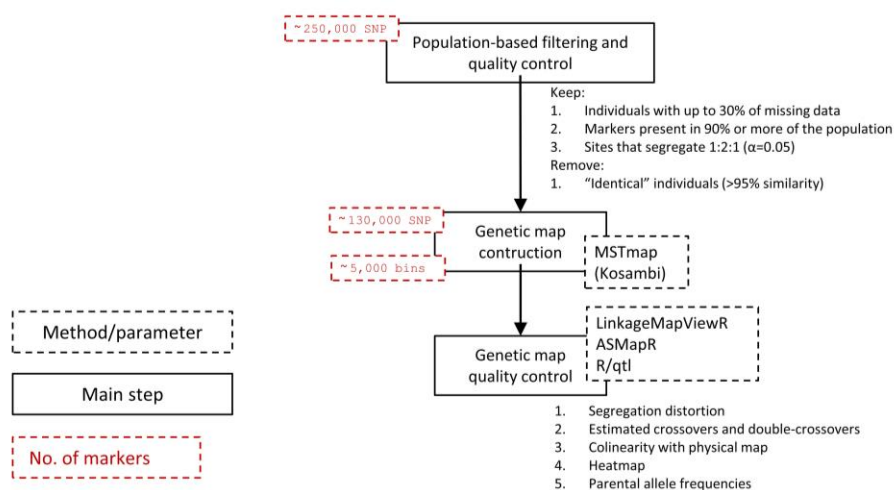
QTL locus	Gene ID	Gene description
<i>pleio4.1</i>	CQ014621	Protein of unknown function
<i>pleio14.1</i>	CQ002756	Similar to VQ22: VQ motif-containing protein 22
<i>pleio14.1</i>	CQ002781	Protein of unknown function
<i>pleio14.1</i>	CQ002783	Protein of unknown function
<i>pleio14.1</i>	CQ002784	Similar to mak16: Protein mak16
<i>pleio14.1</i>	CQ002806	Similar to PCMP-A1: Pentatricopeptide repeat-containing protein At1g31790
<i>pleio14.1</i>	CQ002823	Protein of unknown function
<i>pleio14.1</i>	CQ002827	Similar to FTSHI4: Probable inactive ATP-dependent zinc metalloprotease FTSHI
<i>pleio14.1</i>	CQ002845	Protein of unknown function
<i>pleio14.1</i>	CQ002856	Similar to At1g65750: Putative ribonuclease H protein At1g65750
<i>pleio14.1</i>	CQ002867	Similar to PAP7: Probable plastid-lipid-associated protein
<i>pleio14.1</i>	CQ002885	Similar to Rpp25l: Ribonuclease P protein subunit p25-like protein

8.2. Supplementary Figures

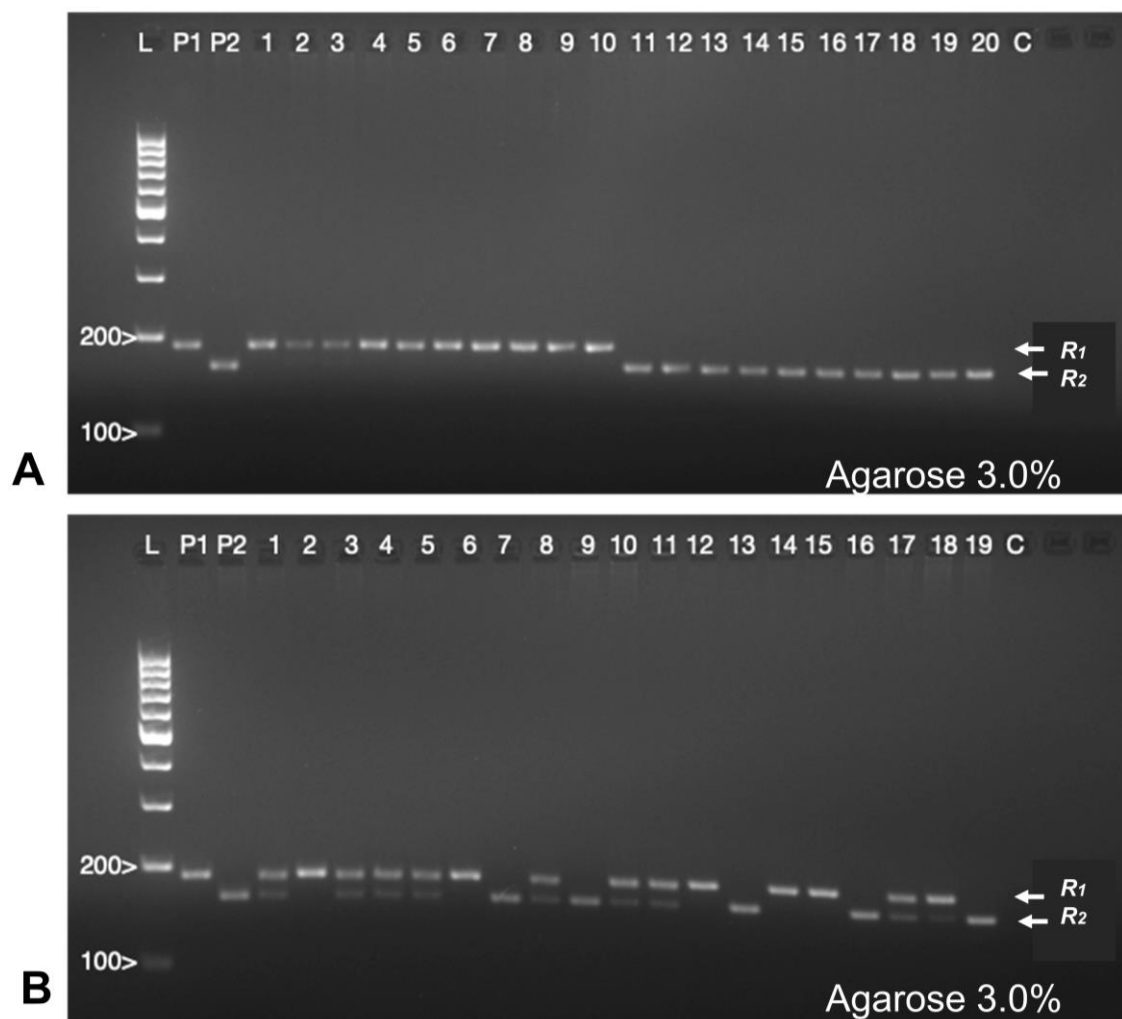
A



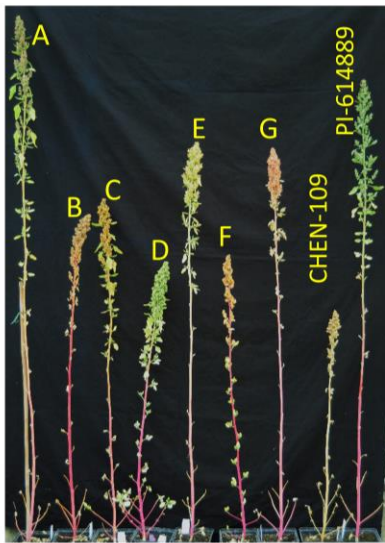
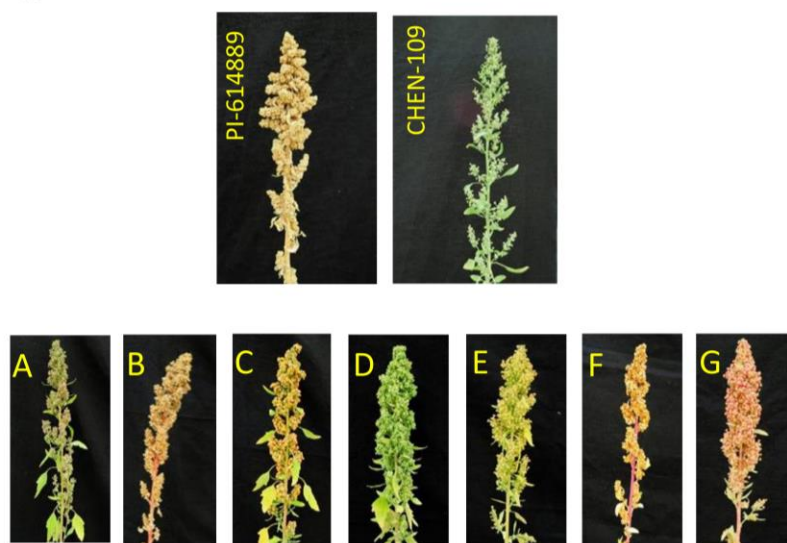
B



Supplementary Figure 1. Flow charts depicting the (A) bioinformatics pipeline and (B) the genetic map construction pipeline. The main steps of the pipelines are shown in solid-line boxes, method/parameters are shown in dashed boxes and the number of markers is shown in red dashed boxes. PE= paired-end.

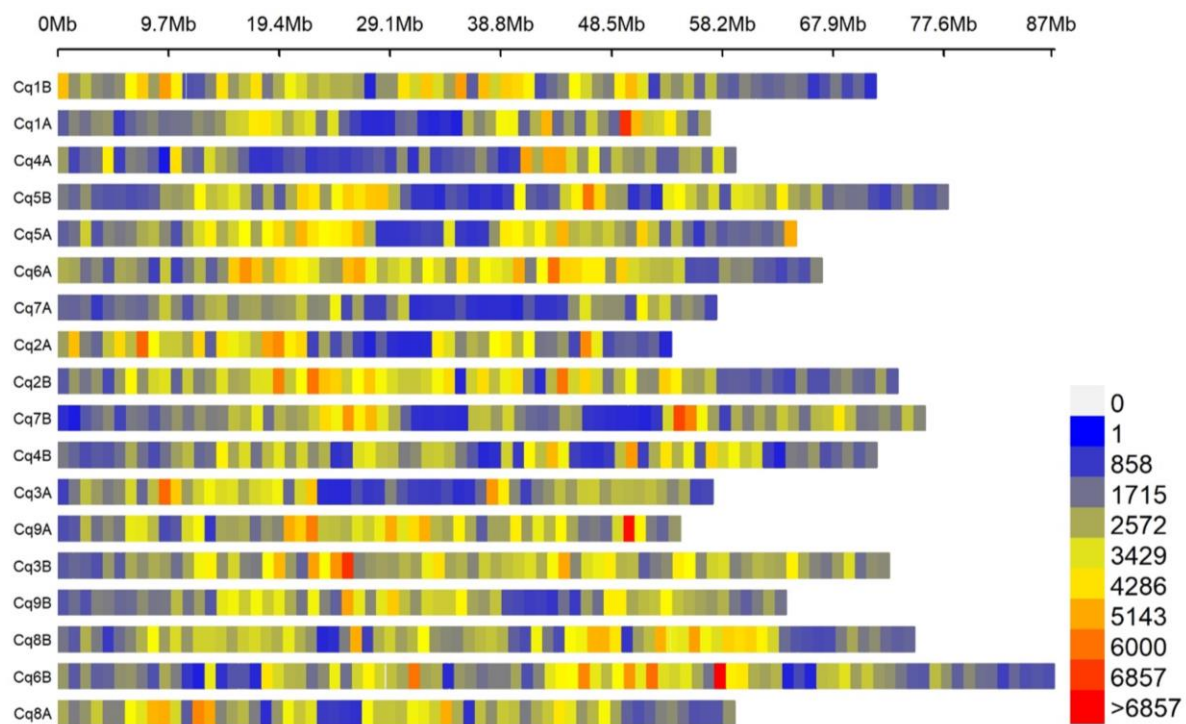


Supplementary Figure 2. Agarose gel electrophoresis of PCR products from the InDel marker JASS5. (A) Homozygous F_3 plants. From lane 1 to 10, plants are homozygous for allele R_1 (parent CHEN-109; 189 bp); and from lane 11 to 20, plants are homozygous for allele R_2 (parent PI-614889; 164 bp). (B) Plants from segregating F_3 families. L = middle range DNA ladder, P1 = parent CHEN-109, P2 = parent PI-614889, C = negative control. Agarose gels were run for 60 min at 100 V.

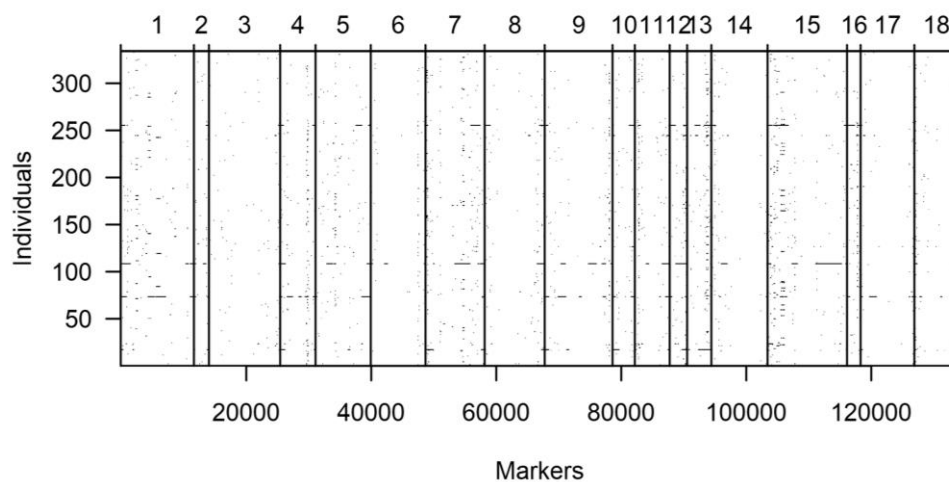
A**B****C**

Supplementary Figure 3. Phenotypic variation in the F_2 population. (A) Variation of panicle density. The numbers in yellow represent the scoring scale used for phenotyping. (B) Variation of plant height illustrated by seven F_2 individuals (A to G) and the parental lines. (C) Variation in days to maturity is illustrated by the different colors of the panicle. PI-614889 had reached maturity stage and was ready to harvest. CHEN-109 had not reached the seed filling stage. Individual C is at seed filling stage. All pictures were taken 16 weeks after sowing.

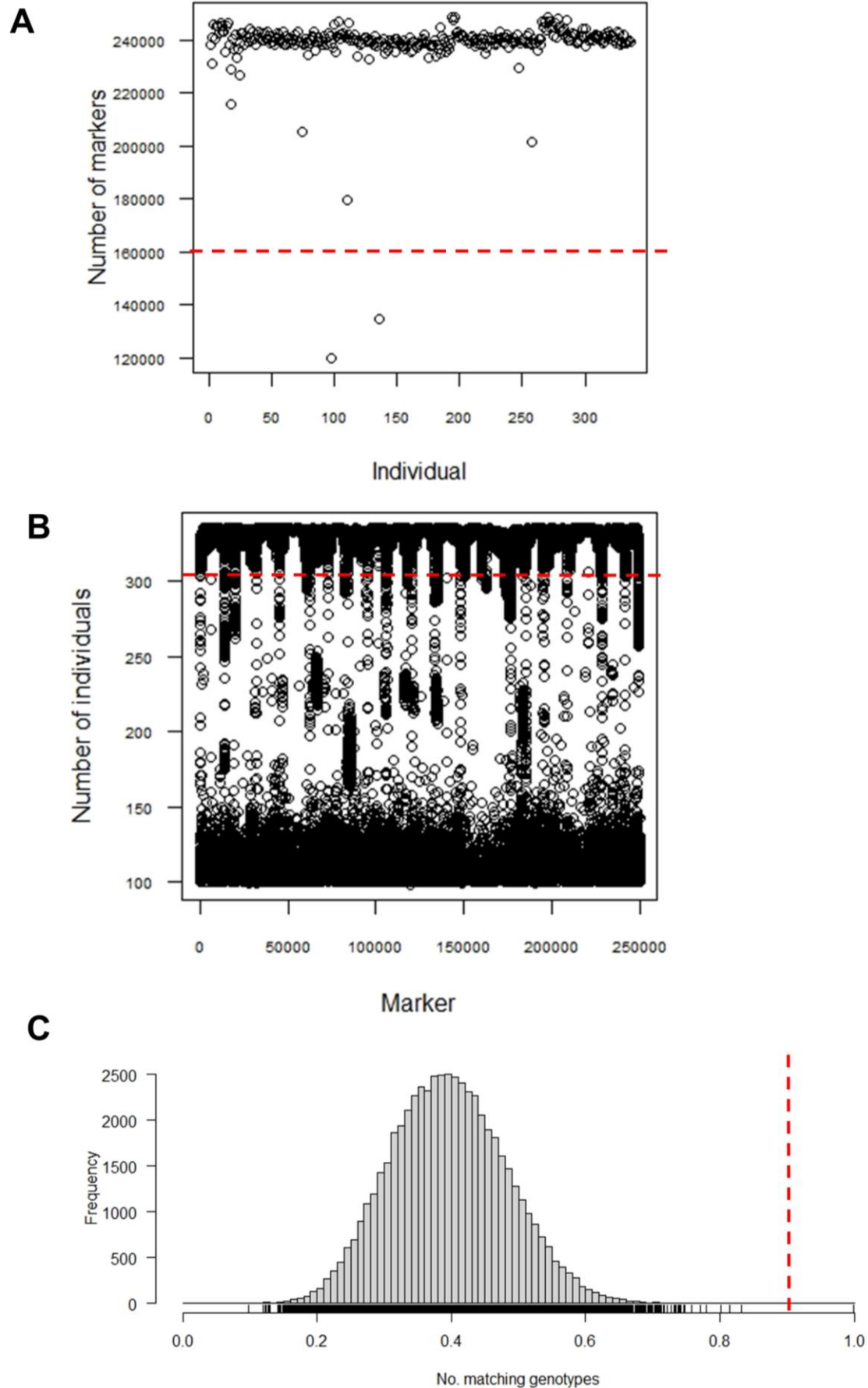
Appendix



Supplementary Figure 4. SNP densities across 18 quinoa chromosomes. The number of SNPs within a 1 Mb window size is shown by different colors. Densities were calculated by the CMplot R package using the raw data (~17 million SNP).



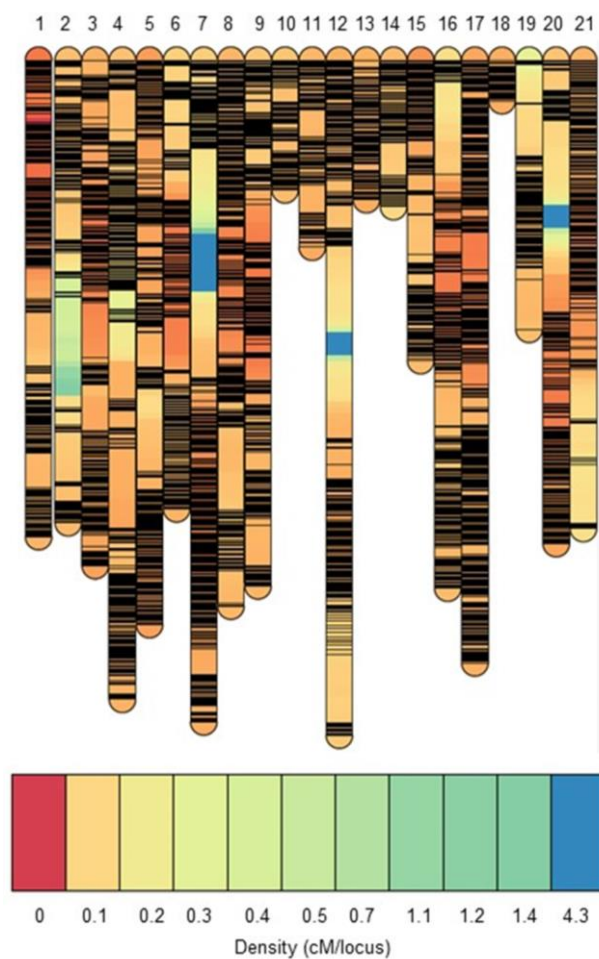
Supplementary Figure 5. Genome-wide missing data. Black dots show missing markers per individual. Black horizontal lines represent several continuous missing markers. Chromosomes are separated by black vertical lines. The numbers above correspond to the quinoa chromosomes (even numbers: Cq1A to Cq9A; uneven numbers: Cq1B to Cq9B).



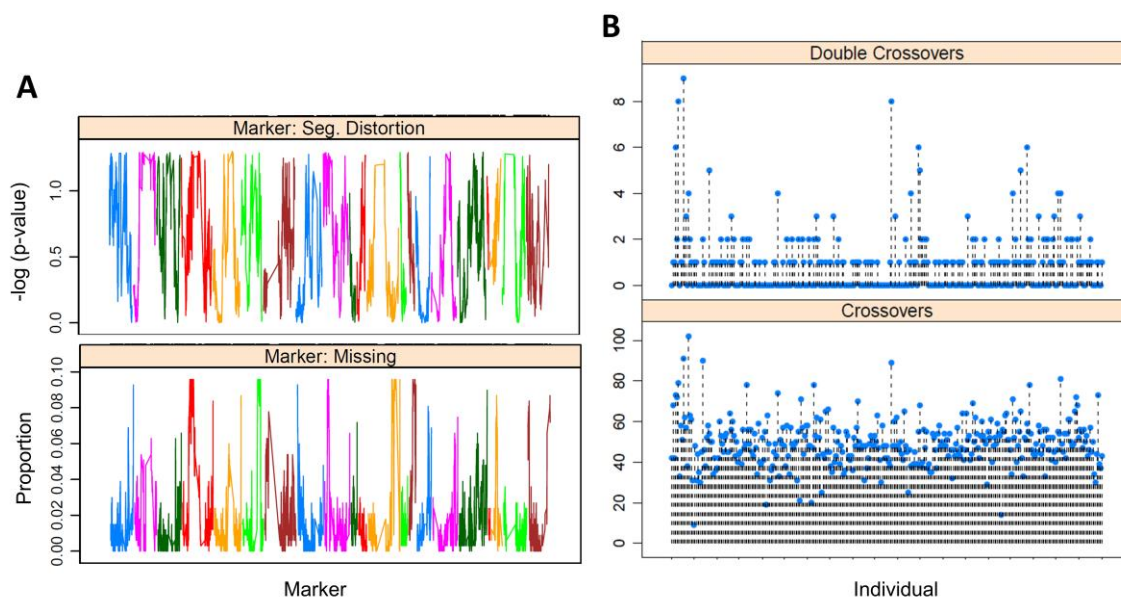
Supplementary Figure 6. Data filtering before map construction. (A) Number of markers per individual. (B) Number of genotyped individuals for each marker. (C) Histogram of the proportion of markers for which pairs of individuals have matching genotypes. Dashed red lines show the filtering thresholds. Data below the dashed red line in A and B, and to the right of the dashed line in C was removed.

Appendix

Supplementary Figure 7. Quinoa linkage map based on 334 plants from an F₂ population derived from a cross between CHEN-109 and PI-614889. The map consists of 133,913 markers arranged in 5,219 bins and it was drawn with the LinkageMapView R package. The numbers above indicate linkage groups (LG) and chromosome numbers. Horizontal blue lines show the location of the first marker of each bin followed by the number of markers in each bin in parenthesis. Marker names are coded as “S” + Chromosome number + “_” + physical position of the marker. Genetic distances in cM are written to the left of each LG (*This figure is provided with the supplementary data DVD, due to the excessive length*).

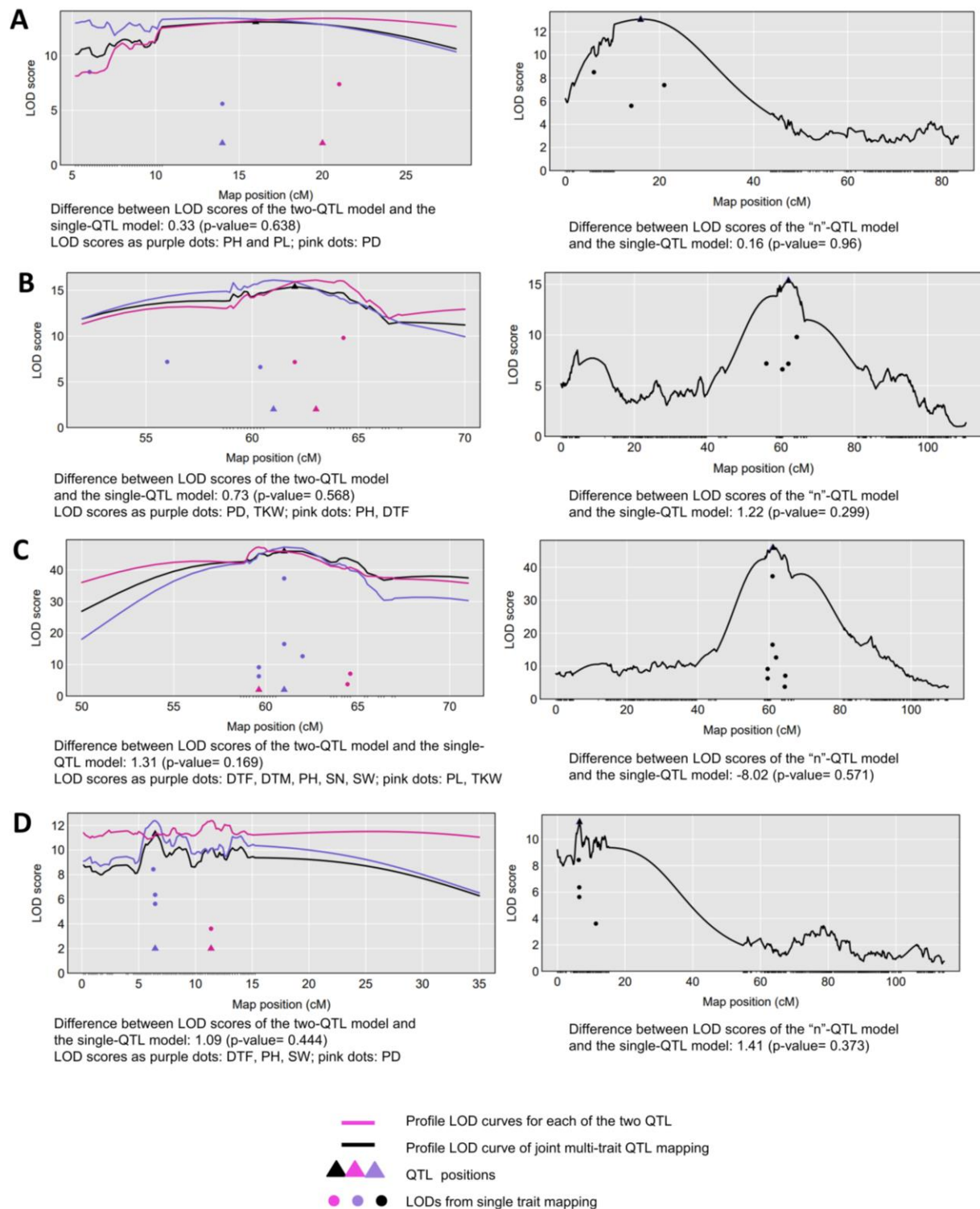


Supplementary Figure 8. Map density across 21 linkage groups from the F_2 population derived from a cross between CHEN-109 and PI-614889. Density given as cM/locus equates to $(\text{markers/cM})^{-1}$. Densities were recorded by the LinkageMapView R package. Binned markers are shown by horizontal black lines.

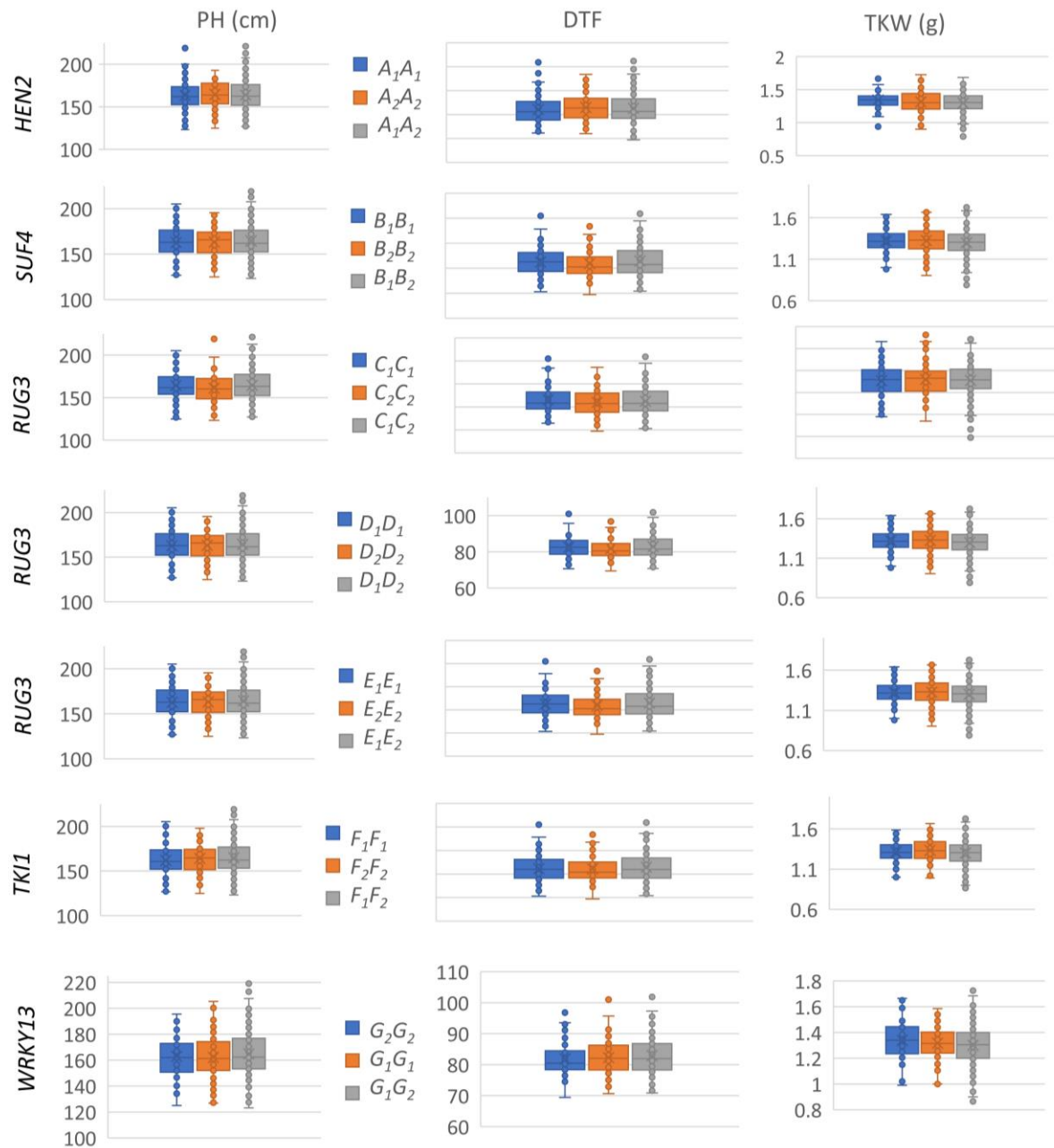


Supplementary Figure 9. Quality control of the marker data used for linkage map construction. (A) Marker's segregation distortion (log of p-value obtained after a Chi-square test) and missing proportion (number of individuals that are missing a marker in a specific locus / 334 individuals) calculated by the ASMap R package; different colors correspond to each linkage group. (B) Estimated number of crossovers and double crossovers calculated by ASMap R package; blue dots represent the estimated values for each F₂ individual.

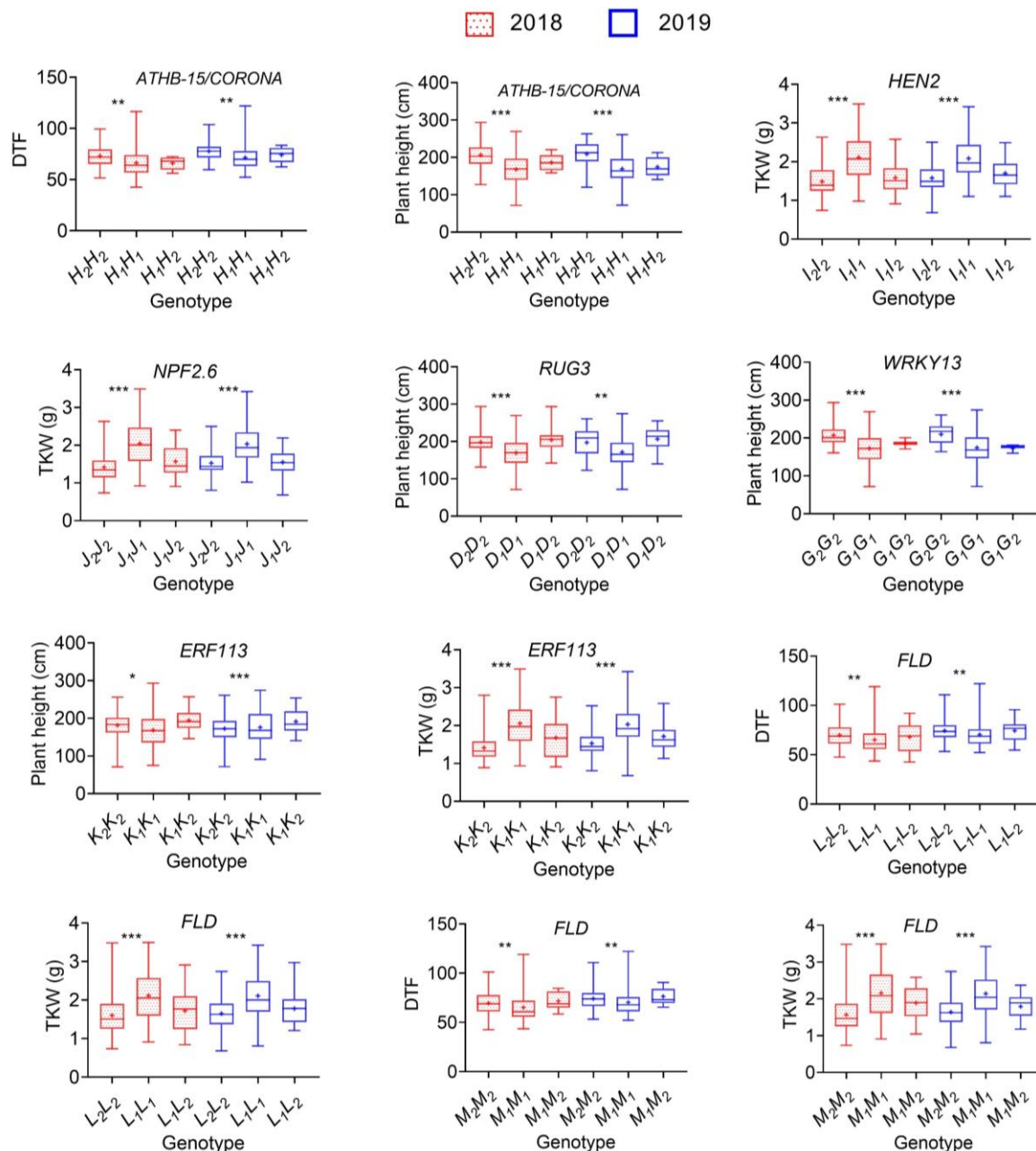
Supplementary Figure 10. Frequencies of the parental alleles calculated from the F₂ population using the program ABHgenotype R package in each of the linkage groups. Different alleles are shown by different colors. Linkage group numbers are shown to the right. The x-axis shows the physical positions of the 133,913 markers according to the reference genome QQ74_V2. Hetero: heterozygous genotype (*This figure is provided with the supplementary data DVD, due to the excessive length*).



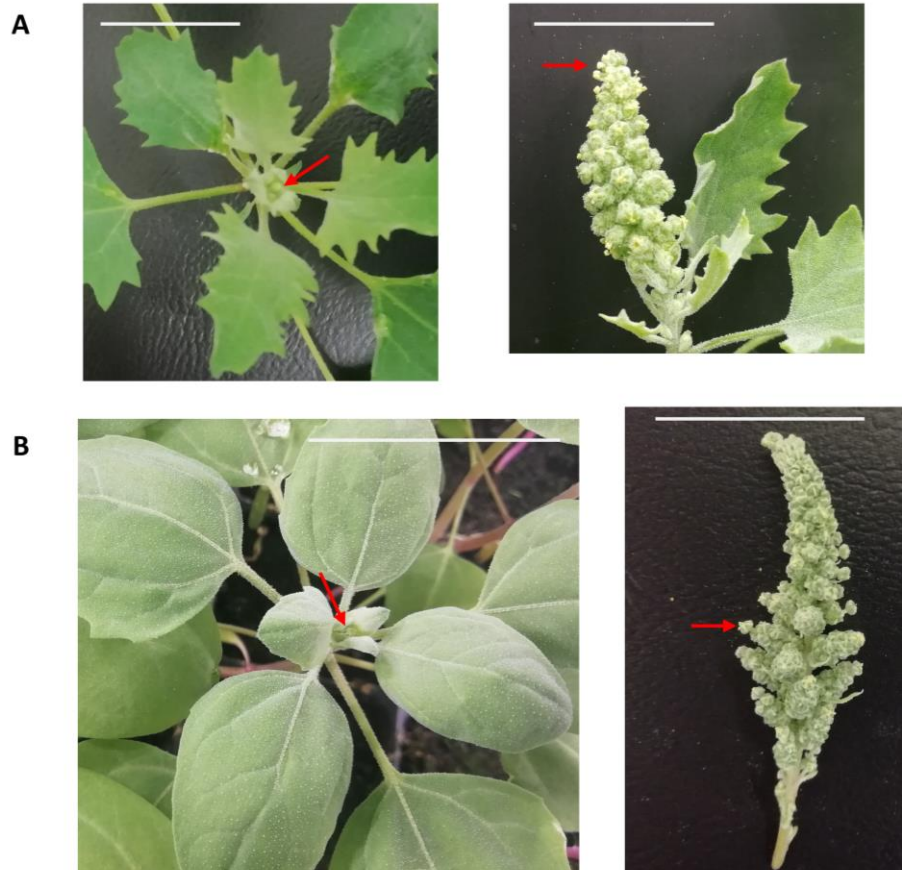
Supplementary Figure 11. Comparative QTL analysis to detect pleiotropy for (A) *pleio20.1*, (B) *pleio4.2*, (C) *pleio4.3* and (D) *pleio7.1*. Two tests were performed: one vs. two QTL (to the left) and one vs. "n" QTL (to the right). The black curve is the LOD score curve for the single-QTL model, with the estimated QTL location indicated by a black triangle. The blue and pink curves are profile LOD score curves for the two-QTL model. Dots indicate the LOD score for the traits considering a single-QTL model. DTF: days to flowering, DTM: days to maturity, PH: plant height, PL: panicle length, PD: panicle density, SN: seed number per plant, SW: seed weight per plant, TKW: thousand kernel weight.



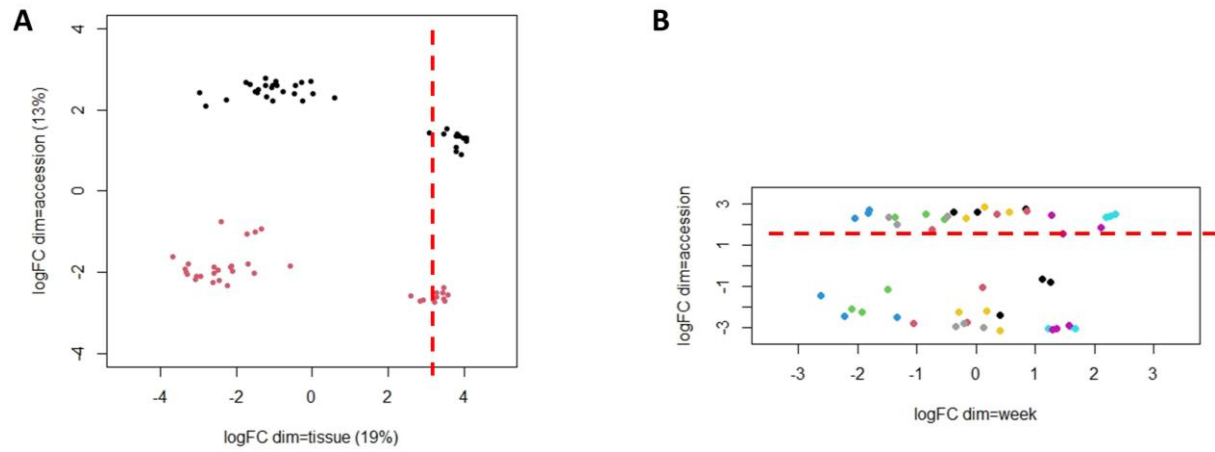
Supplementary Figure 12. Evaluation of variant haplotypes using the sequences of 328 F₂ individuals and the corresponding phenotypic data of 328 F₃ families. Phenotypic effects of haplotype variations within several candidate genes are shown: *DEXH-BOX ATP-DEPENDENT RNA HELICASE DEXH10* (*HEN2*), *SUPPRESSOR OF FRI 4* (*SUF4*), *RCC1 DOMAIN-CONTAINING PROTEIN 3* (*RUG3*), *WRKY TRANSCRIPTION FACTOR 13* (*WRKY13*) and *TSL-KINASE INTERACTING PROTEIN 1* (*TKI1*). The variants genotypes correspond to, for instance, A₁A₁ (our homozygous parent PI-614889), A₁A₂ (heterozygous), A₂A₂ (our homozygous parent CHEN-109) and are described in Table S4. Significant differences between genotypes are shown by asterisks (t-test, $\alpha < 0.05 = **$, $\alpha < 0.01 = *$, $\alpha < 0.001 = ***$). DTF: days to flowering, TKW: thousand kernel weight, PH: plant height.



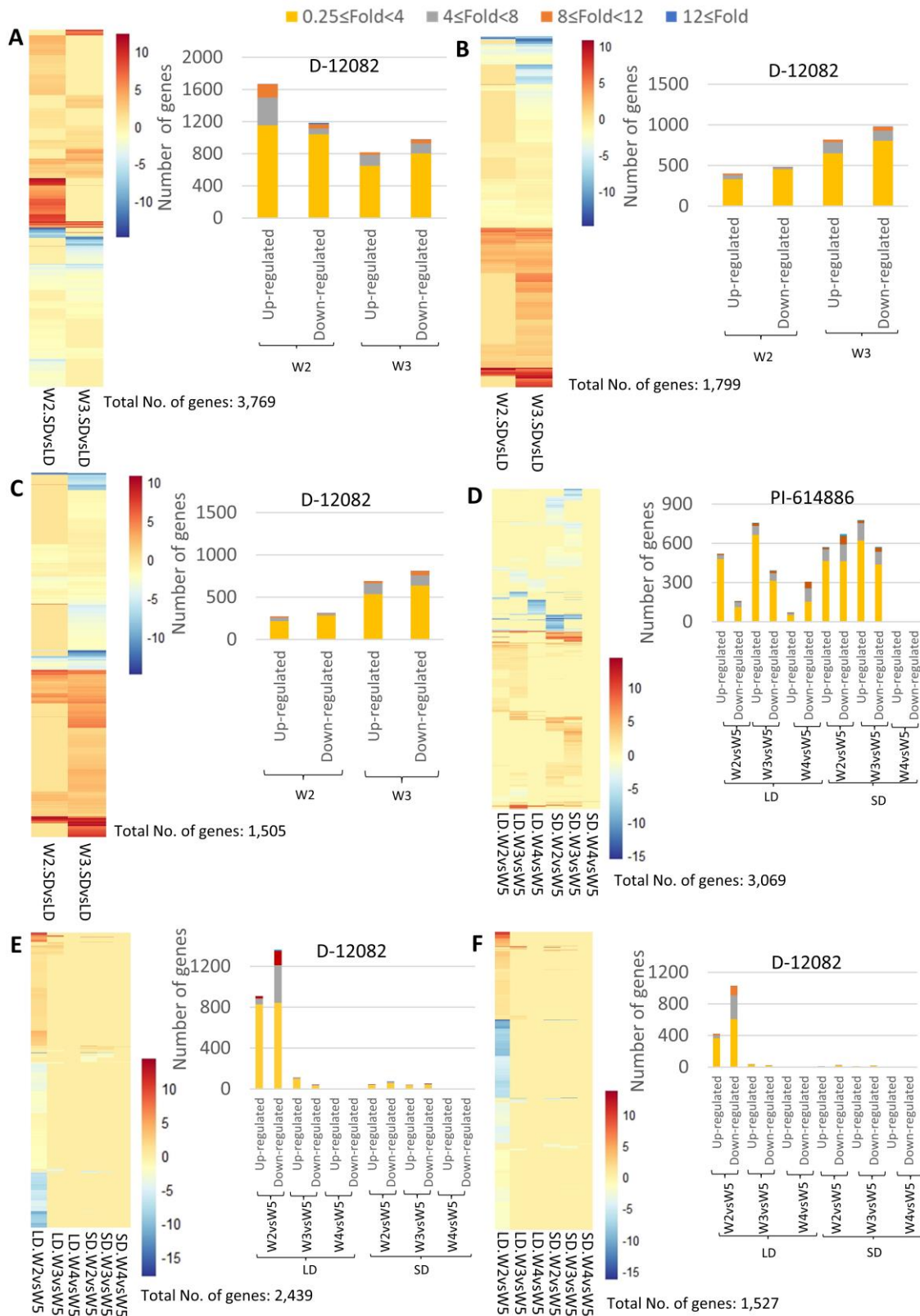
Supplementary Figure 13. Evaluation of variant haplotypes using available whole-genome sequencing and phenotypic data of 310 quinoa accessions (Patiranage et al., 2020). Phenotypic effects of haplotype variations within several candidate genes are shown: *ATHB-15/CORONA*, *DEXH-BOX ATP-DEPENDENT RNA HELICASE DEXH10 (HEN2)*, *NRT1/ PTR 2.6 (NPF2.6)*, *RCC1 DOMAIN-CONTAINING PROTEIN 3 (RUG3)*, *WRKY TRANSCRIPTION FACTOR 13 (WRKY13)*, *ETHYLENE-RESPONSIVE TRANSCRIPTION FACTOR 113 (ERF113)* and *FLOWERING LOCUS D (FLD)*. The variants genotypes correspond to, for instance, H_1H_1 (our homozygous parent PI-614889), H_1H_2 (heterozygous), H_2H_2 (our homozygous parent CHEN-109) and are described in Table S4. Significant differences between genotypes are shown by asterisks (t-test, $\alpha<0.05$ =**, $\alpha<0.01$ =***). Phenotypic data of different years are shown in different colors. DTF: days to flowering, TKW: thousand kernel weight.



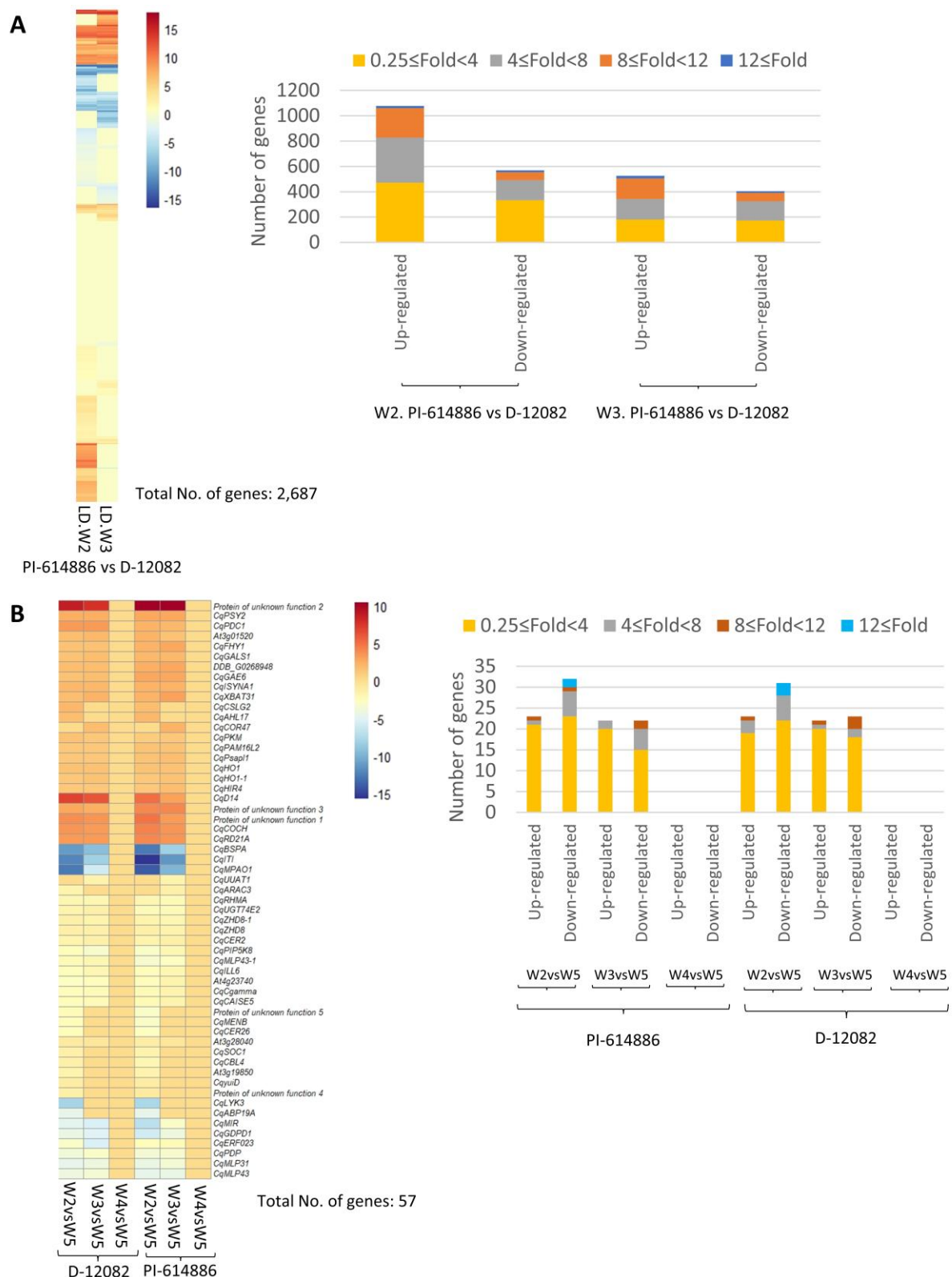
Supplementary Figure 14. Quinoa plants at bolting and the first flower opens stages. Plants from (A) PI-614886 and (B) D-12082 grown under long-days (16 h light) are shown. Red arrows show visible floral buds and open flowers. Scale bars=3cm.

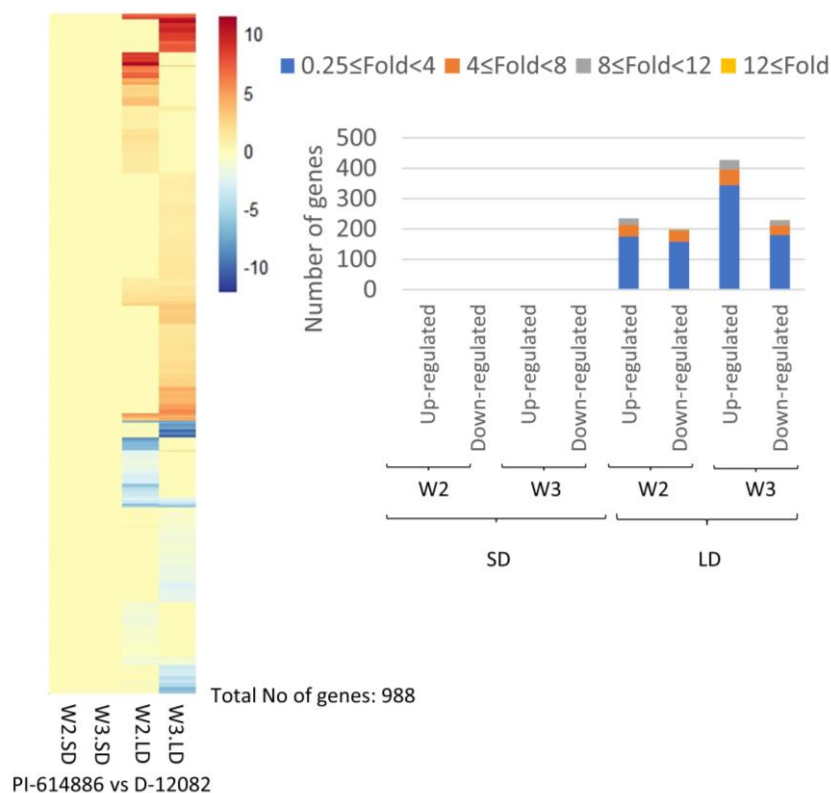


Supplementary Figure 15. Unsupervised samples clustering with the filtered and normalized transcriptome data. (A) Leaf and shoot apical meristem (SAM) samples clustered by tissue and accession. SAM and leaf transcriptome values are located to the right and left of the dotted line, respectively. (B) Biological replicates of leaf transcriptomes clustered together. Biological replicates are equally colored for PI-614886 (above the dotted line) and D-12082 (under the dotted line). logFC= Trimmed Mean log-ratios (\log_2 fold change of M values).

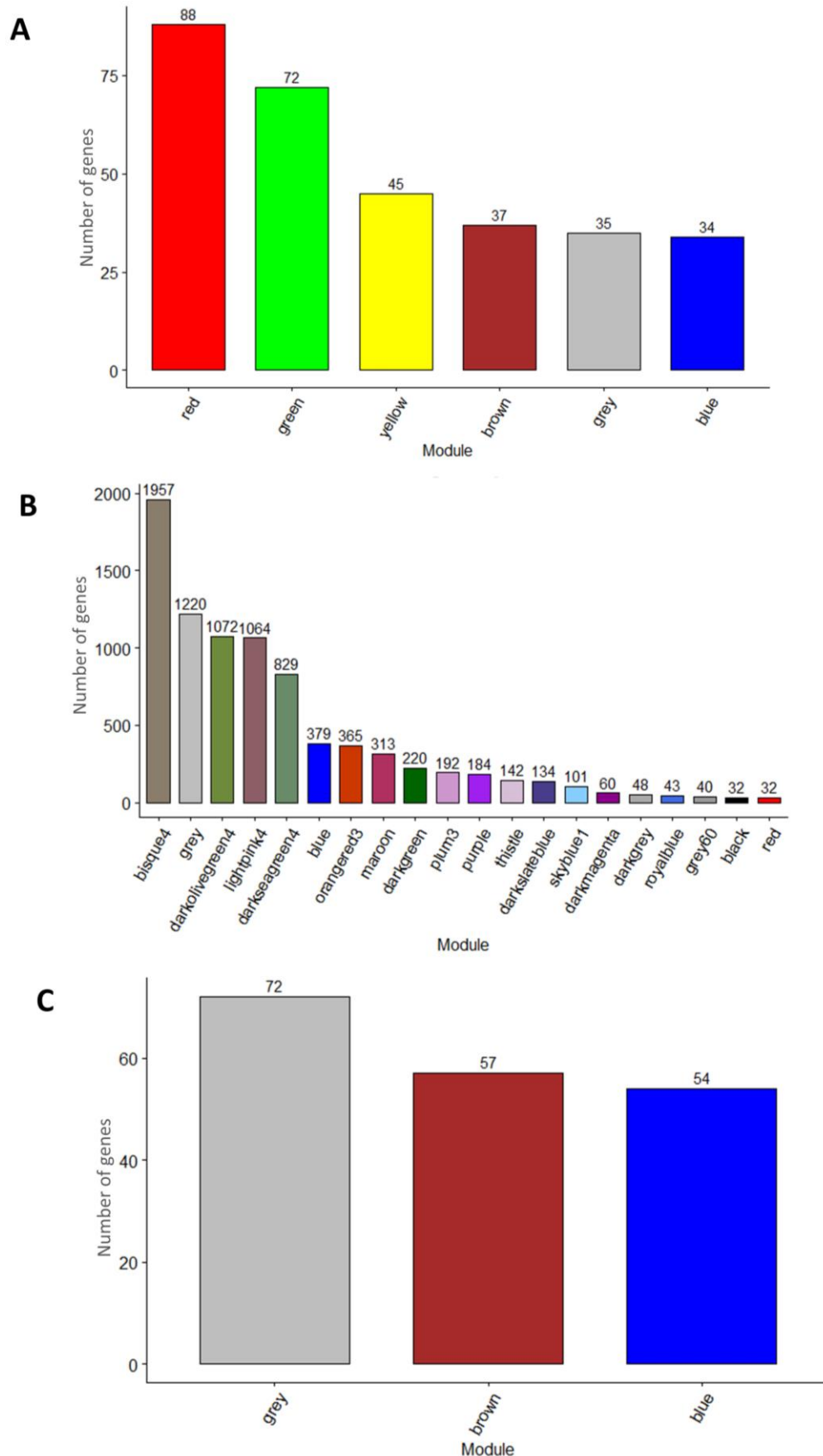


Supplementary Figure 16. Transcriptome data investigation to identify genes differentially responding to photoperiod (A) Heatmap and stacked column graph of differentially expressed genes (DEGs) between short-day (SD) and long-day (LD) in D-12082 at W2 and W3. (B) Heatmap and stacked column graph of the DEGs in (A), which were down- or up-regulated in W3. (C) Heatmap and stacked column graph of the DEGs in (B) whose differential expression differed between W2 and W3. (D, E) Heatmap and stacked column graph of DEGs between W2-W5, W3-W5, and W4-W5 for the short-day (D-12082) and the day-neutral (PI-614886) accessions. DEGs with different expression patterns between SD and LD conditions are shown. (F) Heatmap and stacked column graph of the DEGs resulting from removal of PI-614886 DEGs in (D) from those in D-12082. Total number of genes in each step is shown at the bottom of the heatmaps. In the heatmap, SD.W4vsW5 corresponds, for instance, to \log_2 (normalized reads at W4/ normalized reads at W5) for a specific gene model under SD. W=week.

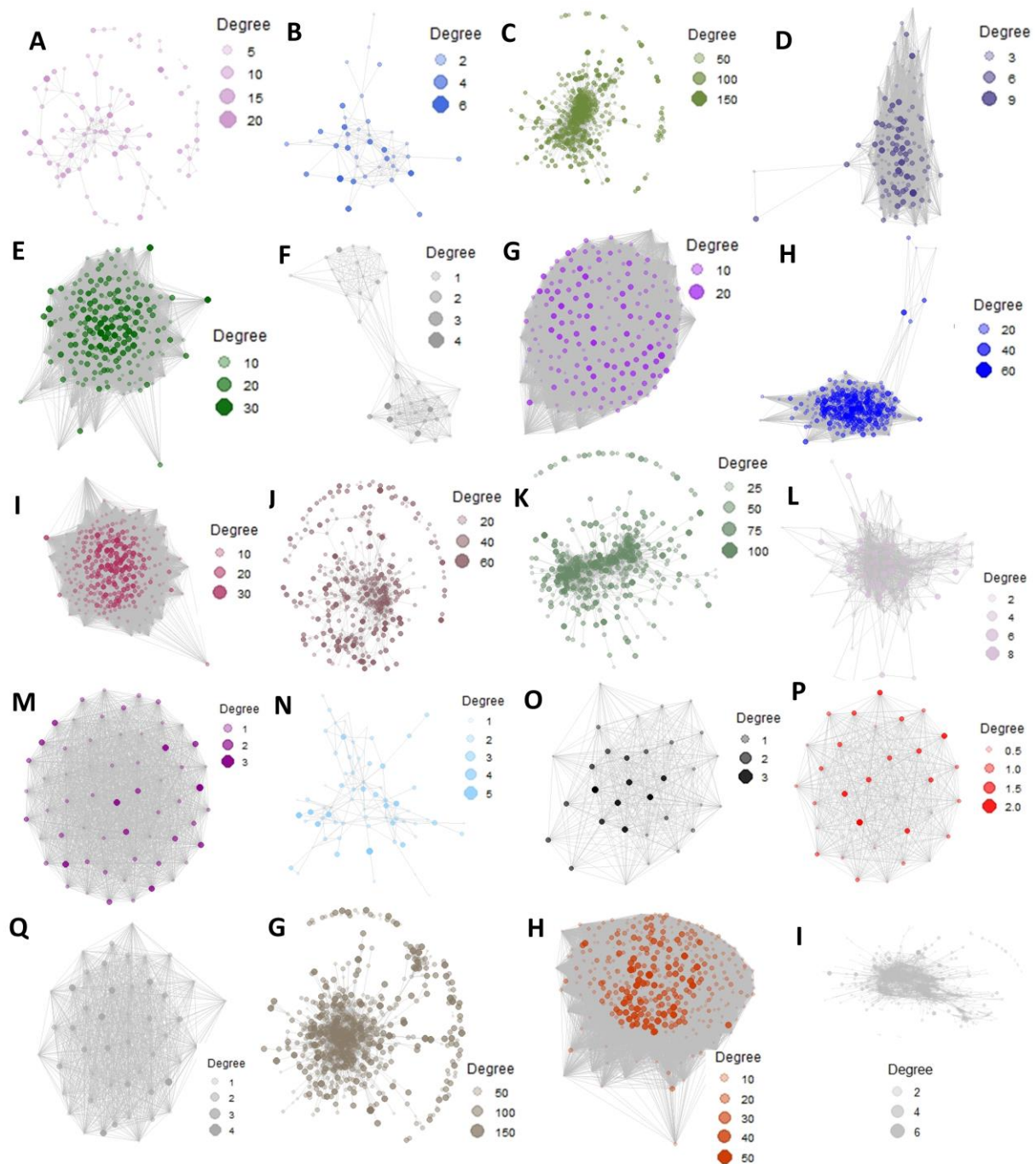




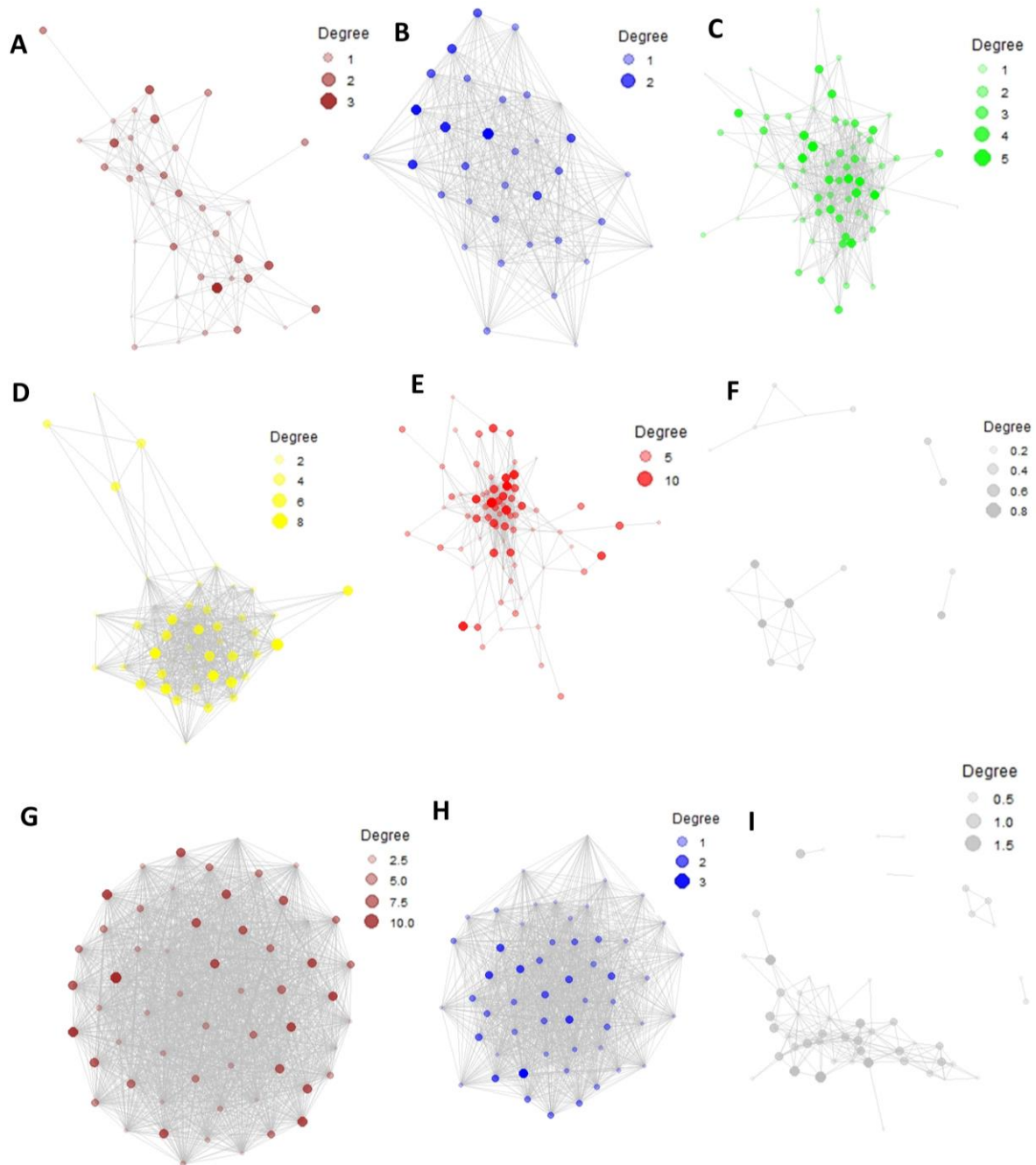
Supplementary Figure 18. Transcriptome data investigation of DEGs from the apical meristem. Heatmap and stacked column graph of DEGs between PI-614886 and D-12082 in shoot apical meristem (SAM) samples are shown. The total number of genes is shown at the bottom right of the heatmap. W2.SD corresponds, for instance, to \log_2 (normalized reads in PI-614886/ normalized reads in D-12082) for a specific gene model at W2 under SD. W=week.



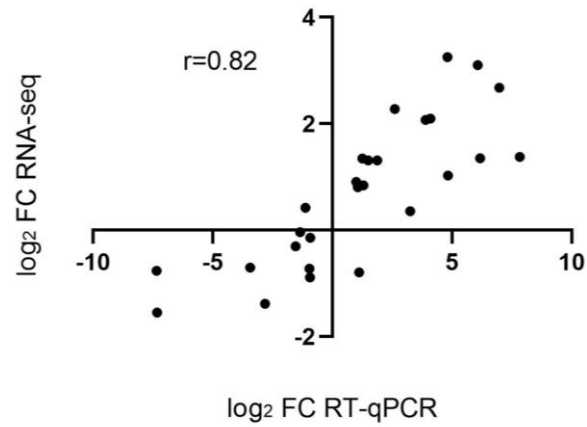
Supplementary Figure 19. Number of genes per module resulting from the co-expression analysis in leaf transcriptome. Every color represents one module of the clustering trees in Figure 11. (A) Differentially expressed genes (DEGs) between short-day (SD) and long-day (LD) conditions in D-12082, which differentially respond to photoperiod to likely regulate flowering time. (B) DEGs between PI-614886 and D-12082 at W2 and W3 under LD, which are putatively controlling flowering under LD. (C) DEGs with similar spatial expression patterns between PI-614886 and D-12082, which likely control the time to flower under SD.



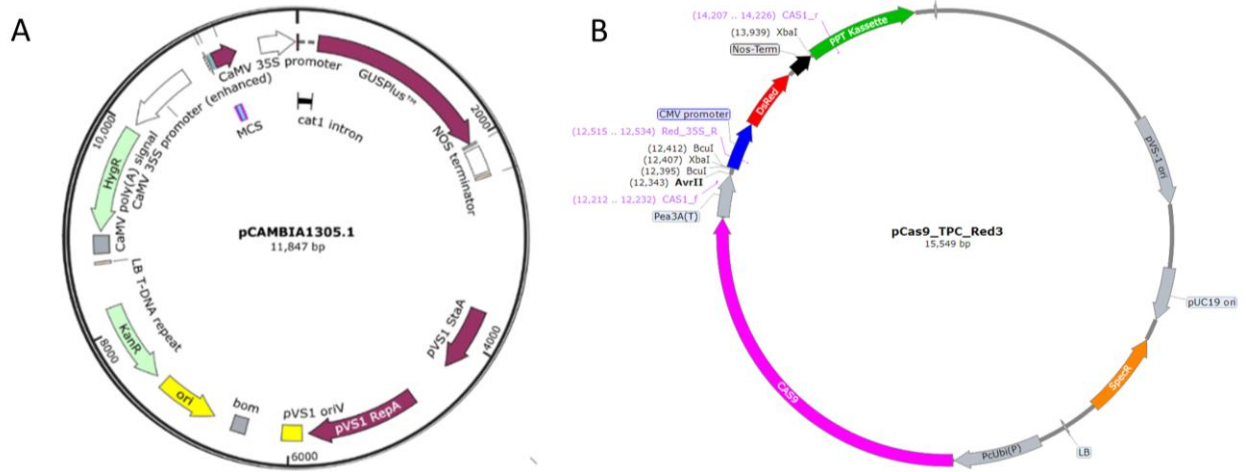
Supplementary Figure 20. Co-expression networks of annotated and newly-assembled transcripts showing differentially expressed genes between PI-614886 and D-12082 at W2 and W3 under LD, which are putatively controlling flowering under LD in leaves. Every color represents one module of the clustering tree in Figure 11C. (A) plum3, (B) royalblue, (C) darkolivegreen4, (D) black, (E) red, (F) darkgrey, (G) bisque4, (H) orange3, (I) grey, (J) lightpink4, (K) darkseagreen4, (L) thistle, (M) darkmagenta, (N) skyblue1, (O) darkslateblue, (P) darkgreen, (Q) grey60, (R) purple, (S) blue and (T) maroon modules. Degree: sum of connection weights of a gene to all other genes in the module by Weighted Gene Co-expression Network Analysis (WGCNA). Number of genes in each module is shown in Supplementary Figure 19.



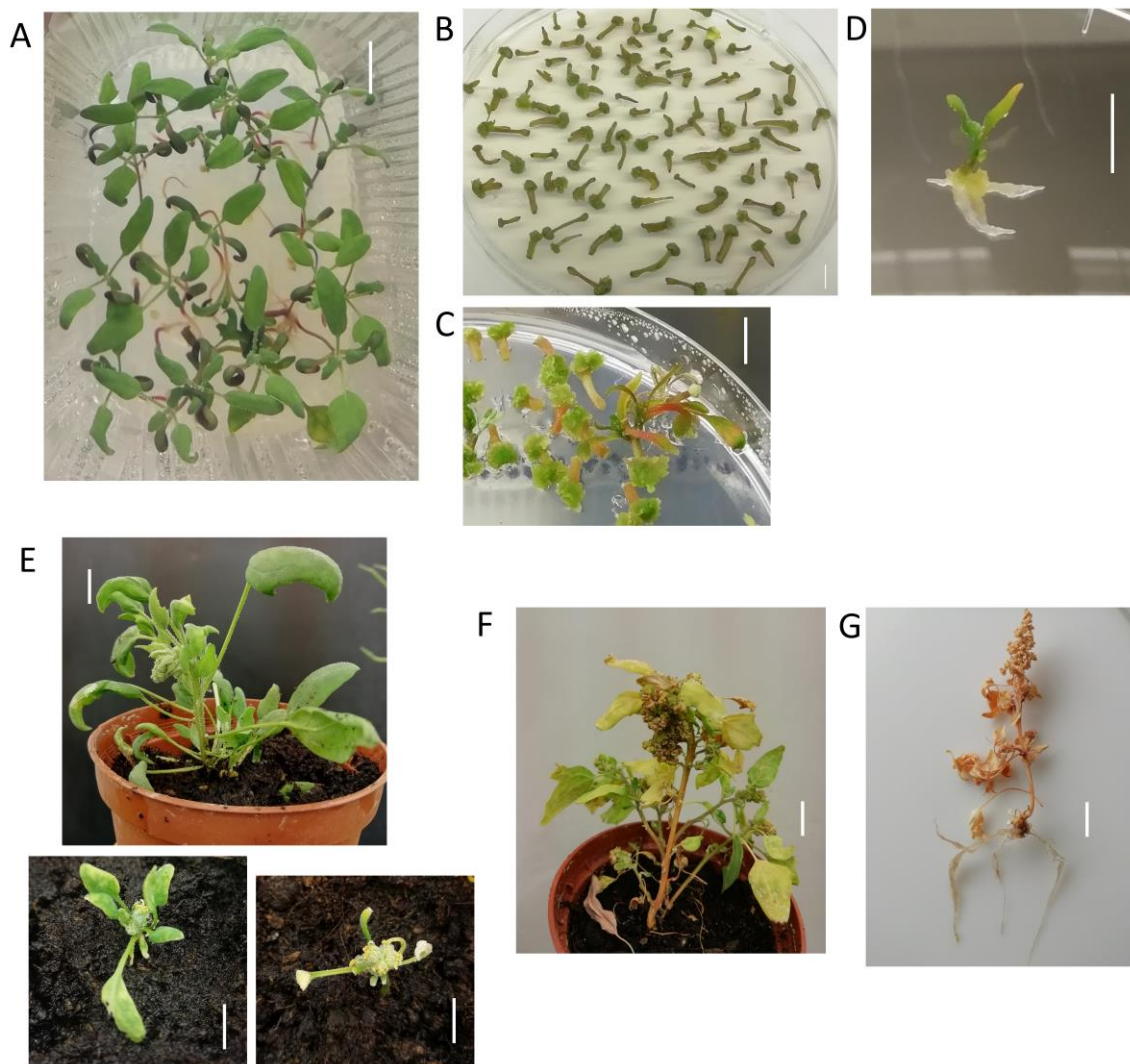
Supplementary Figure 21. Co-expression networks of annotated and newly-assembled transcripts showing (DEGs) between short-day (SD) and long-day (LD) conditions in D-12082, which differentially respond to photoperiod to likely regulate flowering time in leaves (A to F) and DEGs with similar spatial expression patterns between PI-614886 and D-12082, which likely control the time to flower under SD (G to I). Every color represents one module of the clustering trees in Figure 11A and Figure 11B. Networks correspond to the (A) brown, (B) blue, (C) green, (D) yellow, (E) red, (F) grey, (G) blue, (H) brown, and (I) grey modules. Degree: sum of connection weights of a gene to all other genes in the module by Weighted Gene Co-expression Network Analysis (WGCNA). Number of genes in each module is shown in Supplementary Figure 19.



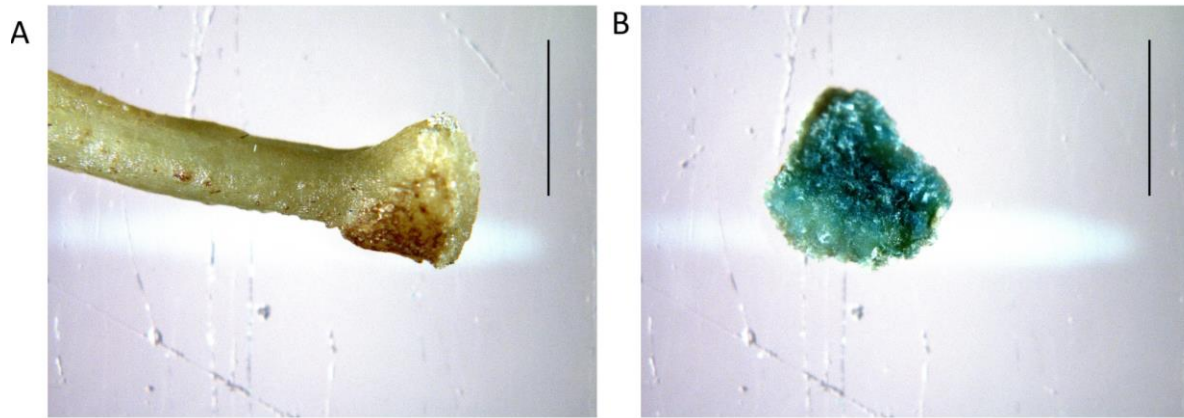
Supplementary Figure 22. Scattered plot showing the correlation between expression data from RNA-seq and RT-qPCR. RT-qPCR values correspond to the \log_2 of the fold change (FC) between day length conditions, between accessions, and between weeks. Data are shown for seven genes. R= Pearson's R. W = week.



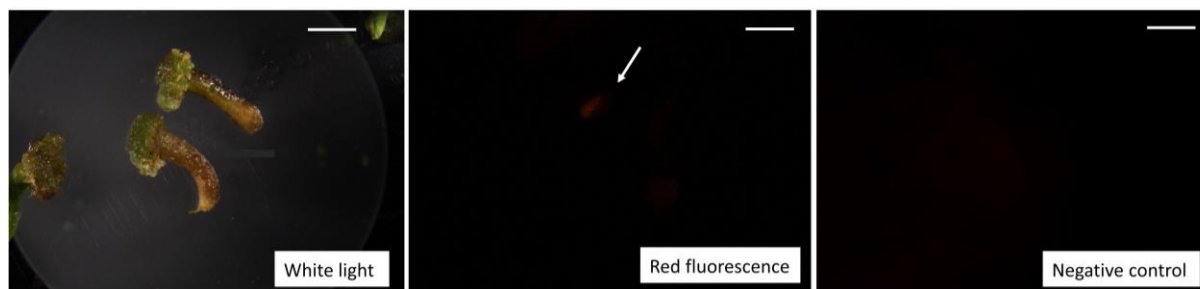
Supplementary Figure 23. Vectors used in this study. (A) pCambia1305.1: kanamycin selection against *Agrobacteria*, GUS reporter gene with catalase intron and hygromycin resistance gene. (B) pCas9_TPC_Red vector. Spectinomycin selection against *Agrobacteria* with a red fluorescence marker gene inserted in the DsRed cassette.



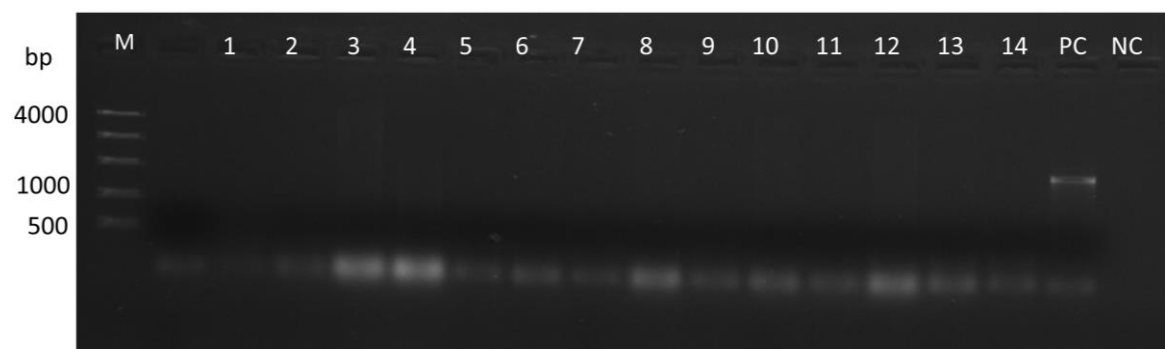
Supplementary Figure 24. Illustrations of the quinoa regeneration protocol for the accession BO-63's petiole explants. (A) Sterile seedlings are grown for two weeks; later, (B) petiole explants are cultured under 16 h light for one week. After six weeks, (C) shoots are obtained. Next, two weeks-old shoots are transferred to rooting media, where the (D) first roots are visible after six weeks. (E) Regenerated plants two weeks after transference to soil. (F) Regenerated plant seven weeks after transference to soil. (G) Regenerated plant reached maturity. Scale bars= 1cm.



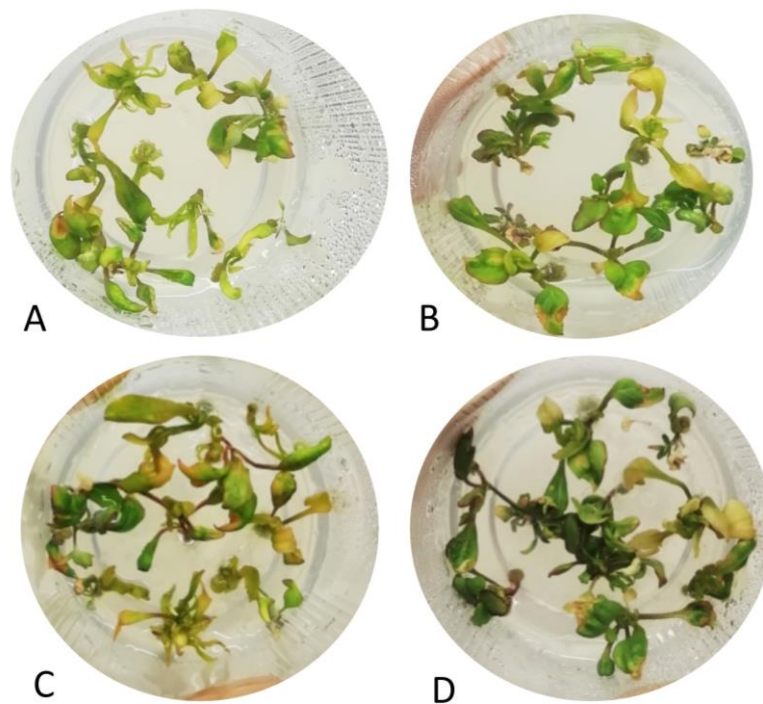
Supplementary Figure 25. Calli from BO-63 after GUS staining. (A) Non-infected callus (control). (B) Callus showing GUS expression. *Agrobacterium*-mediated transformation was performed with the vector pCambia1305.1. Calli were grown from petiole explants under 16 h light; mMS media, 0.25 mg/l BAP + 0.50 mg/l 2,4-D. Liquid and solid co-cultivation were performed for 25 min and 72 hours, respectively. Scale bar= 5 mm.



Supplementary Figure 26. Red fluorescence selection of putatively transformed calli. From left to right: plain view of exemplary calli under white light; view under red fluorescence exposure; and a negative control constituted by non-transformed calli. The white arrow shows red fluorescence, which indicates a putatively transformed callus; exposure time: 30 s. Scale bar= 5000 μm .



Supplementary Figure 27. Agarose gel electrophoresis of PCR products with primers Cas9F+Cas9R for verification of Cas9 insertion. DNA was isolated from 14 putative transformed calli, which show blue staining or red fluorescence. Lane M=high range DNA ladder; lanes 1-14=calli DNA; PC=bacterial culture used for inoculation; NC=water. Expected band size of the PC: 1,735 bp. Agarose gel conditions: 1.0%, 100 V, 30 min.



Supplementary Figure 28. Non-*Agrobacterium* infected regenerated shoots from the accession BO-63 grown in mMS media + 2000 μ l BAP + 200 μ l IBA, with different concentrations of phosphinothricin: (A) 0.5 (B) 1.0 (C) 2.0 and (D) 2.5 mg/l. Photos were taken six weeks after transference to shooting media.

9. References

- Adamczyk, B. J., Lehti-Shiu, M. D., & Fernandez, D. E. (2007). The MADS domain factors *AGL15* and *AGL18* act redundantly as repressors of the floral transition in *Arabidopsis*. *The Plant Journal*, 50(6), 1007-1019.
- Ahmadzai, H. (2020). Trends in Quinoa Adoption in Marginal Areas. *Journal of Agribusiness Rural Development*, 57(3), 235–247-235–247.
- Alandia, G., Odone, A., Rodriguez, J. P., Bazile, D., & Condori, B. (2021). Quinoa—Evolution and future perspectives. In *The Quinoa Genome* (pp. 179-195): Springer.
- Alandia, G., Rodriguez, J., Jacobsen, S.-E., Bazile, D., & Condori, B. (2020). Global expansion of quinoa and challenges for the Andean region. *Global Food Security*, 26, 100429.
- Ali, O. I. E.-D. (2019). Nutritional value of germinated quinoa seeds and their protective effects on rats' health injected by nicotine. *Egyptian Journal of Food Science*, 47(2), 227-241.
- Andrés, F., & Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. *Nature Reviews Genetics*, 13(9), 627-639.
- Angeli, V., Miguel Silva, P., Crispim Massuela, D., Khan, M. W., Hamar, A., Khajehei, F., . . . Piatti, C. (2020). Quinoa (*Chenopodium quinoa* Willd.): An overview of the potentials of the “golden grain” and socio-economic and environmental aspects of its cultivation and marketization. *Foods*, 9(2), 216.
- Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., . . . Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS one*, 3(10), e3376.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., . . . Prjibelski, A. D. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, 19(5), 455-477.
- Bao, S., Hua, C., Shen, L., & Yu, H. (2020). New insights into gibberellin signaling in regulating flowering in *Arabidopsis*. *Journal of Integrative Plant Biology*, 62(1), 118-131.
- Bazile, D., Fuentes, F., & Mujica, Á. (2013). Historical perspectives and domestication: CABI.
- Bazile, D., Jacobsen, S.-E., & Verniau, A. (2016). The global expansion of quinoa: trends and limits. *Frontiers in plant science*, 7, 622.
- Bhargava, A., Shukla, S., & Ohri, D. (2006). *Chenopodium quinoa*—an Indian perspective. *Industrial crops products*, 23(1), 73-87.
- Bhargava, A., & Srivastava, S. (2013). Quinoa: Botany, production and uses: CABI.
- Blümel, M., Dally, N., & Jung, C. (2015). Flowering time regulation in crops—what did we learn from *Arabidopsis*? *Current Opinion in Biotechnology*, 32, 121-129.
- Boopathi, N. M. (2020). QTL analysis. In *Genetic Mapping and Marker Assisted Selection* (pp. 253-326): Springer.
- Byrne, A., Cole, C., Volden, R., & Vollmers, C. (2019). Realizing the potential of full-length transcriptome sequencing. *Philosophical Transactions of the Royal Society* 374(1786), 20190097.
- Cervantes, D. P., & van Loo, E. (2017). QTL mapping for agromorphological traits in Quinoa (*Chenopodium quinoa* Willd.). Wageningen University, Wageningen, The Netherlands.
- Cháb, D., Kolář, J., Olson, M. S., & Štorchová, H. (2008). Two *FLOWERING LOCUS T (FT)* homologs in *Chenopodium rubrum* differ in expression patterns. *Planta*, 228(6), 929.

References

- Che, P., Anand, A., Wu, E., Sander, J. D., Simon, M. K., Zhu, W., . . . Liu, D. (2018). Developing a flexible, high-efficiency *Agrobacterium*-mediated sorghum transformation system with broad application. *Plant Biotechnology Journal*, 16(7), 1388-1395.
- Cheng, S., Chen, P., Su, Z., Ma, L., Hao, P., Zhang, J., . . . Wang, H. (2021). High-resolution temporal dynamic transcriptome landscape reveals a GhCAL-mediated flowering regulatory pathway in cotton (*Gossypium hirsutum* L.). *Plant Biotechnology Journal*, 19(1), 153-166.
- Cheng, Y., Zhou, Y., Yang, Y., Chi, Y.-J., Zhou, J., Chen, J.-Y., . . . Zhou, Y.-H. (2012). Structural and functional analysis of VQ motif-containing proteins in *Arabidopsis* as interacting proteins of WRKY transcription factors. *Plant Physiology*, 159(2), 810-825.
- Chiurugwi, T., Holmes, H. F., Qi, A., Chia, T. Y., Hedden, P., & Mutasa-Göttgens, E. (2013). Development of new quantitative physiological and molecular breeding parameters based on the sugar-beet vernalization intensity model. *J. Agric. Sci.*, 151, 492-505.
- Clough, S. J., & Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal*, 16(6), 735-743.
- Collard, B. C., Jahufer, M., Brouwer, J., & Pang, E. (2005). An introduction to markers, quantitative trait loci (QTL) mapping, and marker-assisted selection for crop improvement: the basic concepts. *Euphytica*, 142(1), 169-196.
- Dally, N., Eckel, M., Batschauer, A., Höft, N., & Jung, C. (2018). Two *CONSTANS-LIKE* genes jointly control flowering time in beet. *Scientific reports*, 8(1), 1-10.
- Danielsen, S., & Munk, L. (2004). Evaluation of disease assessment methods in quinoa for their ability to predict yield loss caused by downy mildew. *Crop Protection*, 23(3), 219-228.
- Darbani, B., Farajnia, S., Toorchi, M., Zakerbostanabad, S., & Noeparvar, S. (2008). DNA-delivery methods to produce transgenic plants. *Biotechnology research international*, 7(3), 385-402.
- Datta, S., Jankowicz-Cieslak, J., Nielen, S., Ingelbrecht, I., & Till, B. J. (2018). Induction and recovery of copy number variation in banana through gamma irradiation and low-coverage whole-genome sequencing. *Plant Biotechnology Journal*, 16(9), 1644-1653.
- Drabešová, J., Černá, L., Mašterová, H., Koloušková, P., Potocký, M., & Štorchová, H. (2016). The evolution of the *FT/TFL1* genes in Amaranthaceae and their expression patterns in the course of vegetative growth and flowering in *Chenopodium rubrum*. *G3: Genes, Genomes, Genetics*, 6(10), 3065-3076.
- Drabešová, J., Cháb, D., Kolář, J., Haškovcová, K., & Štorchová, H. (2014). A dark–light transition triggers expression of the floral promoter *CrFTL1* and downregulates *CONSTANS*-like genes in a short-day plant *Chenopodium rubrum*. *Journal of Experimental Botany*, 65(8), 2137-2146.
- Dun, E. A., Ferguson, B. J., & Beveridge, C. A. (2006). Apical dominance and shoot branching. Divergent opinions or divergent mechanisms? *Plant Physiology*, 142(3), 812-819.
- Fang, C., Chen, L., Nan, H., Kong, L., Li, Y., Zhang, H., . . . Hou, Z. (2019). Rapid identification of consistent novel QTLs underlying long-juvenile trait in soybean by multiple genetic populations and genotyping-by-sequencing. *Molecular breeding*, 39(6), 1-11.
- FAO. (2010). The second report on the state of the world's plant genetic resources for food and agriculture: Food & Agriculture Org. UN Food and Agriculture Organization (FAO).
- FAO. (2021). FAOSTAT. FAO. Food and Agriculture Organization of the United Nations, Rome, Italy.

References

- Feldmann, K. A., & David Marks, M. (1987). *Agrobacterium*-mediated transformation of germinating seeds of *Arabidopsis thaliana*: a non-tissue culture approach. *Molecular General Genetics*, 208(1), 1-9.
- Galvão, V. C., Horrer, D., Küttner, F., & Schmid, M. (2012). Spatial control of flowering by DELLA proteins in *Arabidopsis thaliana*. *Development*, 139(21), 4072-4082.
- Garcia, M., Condori, B., & Del Castillo, C. (2015). Agroecological and agronomic cultural practices of quinoa in South America. *Quinoa: Improvement and sustainable production*, 25-46.
- Gleason, K. J., Yang, F., Pierce, B. L., He, X., & Chen, L. S. (2020). Primo: integration of multiple GWAS and omics QTL summary statistics for elucidation of molecular mechanisms of trait-associated SNPs and detection of pleiotropy in complex traits. *Genome Biology*, 21(1), 1-24.
- Golicz, A. A., Steinfort, U., Arya, H., Singh, M. B., & Bhalla, P. L. (2020). Analysis of the quinoa genome reveals conservation and divergence of the flowering pathways. *Functional Integrative Genomics*, 20(2), 245-258.
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., . . . Zeng, Q. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29(7), 644-652.
- Gurel, S. J. S. T. (2021). Sand-Wounding of Shoot and Petiole Explants Enhances Transformation Efficiency in Sugar Beet (*Beta vulgaris* L.) by *Agrobacterium*-mediated Transformation. 23(2), 415-427.
- Gurusaravanan, P., Vinoth, S., & Jayabalan, N. (2020). An improved *Agrobacterium*-mediated transformation method for cotton (*Gossypium hirsutum* L. 'KC3') assisted by microinjection and sonication. *In Vitro Cellular Developmental Biology-Plant*, 56(1), 111-121.
- Halevy, A. H. (2018). *Handbook of Flowering: Volume I*: CRC press.
- Hatakeyama, M., Aluri, S., Balachadran, M. T., Sivarajan, S. R., Patrignani, A., Grüter, S., . . . Francoijs, K.-J. (2018). Multiple hybrid de novo genome assembly of finger millet, an orphan allotetraploid crop. *DNA Research*, 25(1), 39-47.
- Hesami, M., & Daneshvar, M. H. (2016). Development of a regeneration protocol through indirect organogenesis in *Chenopodium quinoa* Willd. *Indo Am J Agric Vet Sci*, 4(2), 25-32.
- Hesami, M., Naderi, R., & Yoosefzadeh-Najafabadi, M. (2018). Optimizing sterilization conditions and growth regulator effects on *in vitro* shoot regeneration through direct organogenesis in *Chenopodium quinoa*. *BioTechnologia. Journal of Biotechnology Computational Biology and Bionanotechnology*, 99(1).
- Huang, X., Feng, Q., Qian, Q., Zhao, Q., Wang, L., Wang, A., . . . Huang, T. (2009). High-throughput genotyping by whole-genome resequencing. *Genome Research*, 19(6), 1068-1076.
- Jacobsen, S.-E., Monteros, C., Christiansen, J., Bravo, L., Corcuera, L., & Mujica, A. (2005). Plant responses of quinoa (*Chenopodium quinoa* Willd.) to frost at various phenological stages. *European Journal of Agronomy*, 22(2), 131-139.
- Jacobsen, S.-E., Monteros, C., Corcuera, L. J., Bravo, L. A., Christiansen, J. L., & Mujica, A. (2007). Frost resistance mechanisms in quinoa (*Chenopodium quinoa* Willd.). *European Journal of Agronomy*, 26(4), 471-475.
- Stanschewski, C. S., Rey, E., Fiene, G., Craine, E., . . . Johansen, K. (2021). Quinoa Phenotyping Methodologies: An International Consensus. *Plants*, 10(9).

References

- Jarvis, D., Kopp, O., Jellen, E., Mallory, M., Pattee, J., Bonifacio, A., . . . Maughan, P. (2008). Simple sequence repeat marker development and genetic mapping in quinoa (*Chenopodium quinoa* Willd.). *Journal of Genetics*, 87(1), 39-51.
- Jarvis, D. E., Ho, Y. S., Lightfoot, D. J., Schmöckel, S. M., Li, B., Borm, T. J., . . . Saber, N. (2017). The genome of *Chenopodium quinoa*. *Nature*, 542(7641), 307.
- Kale, S. M., Jaganathan, D., Ruperao, P., Chen, C., Punna, R., Kudapa, H., . . . Doddamani, D. (2015). Prioritization of candidate genes in “QTL-hotspot” region for drought tolerance in chickpea (*Cicer arietinum* L.). *Scientific reports*, 5(1), 1-14.
- Kanitz, A., Gypas, F., Gruber, A. J., Gruber, A. R., Martin, G., & Zavolan, M. (2015). Comparative assessment of methods for the computational inference of transcript isoform abundance from RNA-seq data. *Genome Biology*, 16(1), 1-26.
- Kavipriya, C., Yuvaraja, A., & Senthil, C. M. (2019). Genetic Transformation Methods for Crop Improvement: A Brief Review. *Agricultural Reviews*, 40(4).
- Keshavareddy, G., Kumar, A., Ramu, V. S., & Sciences, A. (2018). Methods of plant transformation-a review. *International Journal of Current Microbiology*, 7(7), 2656-2668.
- Koboldt, D. C., Steinberg, K. M., Larson, D. E., Wilson, R. K., & Mardis, E. R. (2013). The next-generation sequencing revolution and its impact on genomics. *Cell*, 155(1), 27-38.
- Komari, T. (1990). Transformation of cultured cells of *Chenopodium quinoa* by binary vectors that carry a fragment of DNA from the virulence region of pTiBo542. *Plant Cell Reports*, 9(6), 303-306.
- Kong, F., Liu, B., Xia, Z., Sato, S., Kim, B. M., Watanabe, S., . . . Harada, K. (2010). Two coordinately regulated homologs of *FLOWERING LOCUS T* are involved in the control of photoperiodic flowering in soybean. *Plant Physiology*, 154(3), 1220-1231.
- Korani, W., O'Connor, D., Chu, Y., Chavarro, C., Ballen, C., Guo, B., . . . Clevenger, J. (2021). De novo QTL-seq Identifies Loci Linked to Blanchability in Peanut (*Arachis hypogaea*) and Refines Previously Identified QTL with Low Coverage Sequence. *Agronomy*, 11(11), 2201.
- Kumar, P., Choudhary, M., Jat, B., Kumar, B., Singh, V., Kumar, V., . . . Rakshit, S. (2021). Skim sequencing: an advanced NGS technology for crop improvement. *Journal of Genetics*, 100(2), 1-10.
- Kyriakidou, M., Tai, H. H., Anglin, N. L., Ellis, D., & Strömvik, M. V. (2018). Current strategies of polyploid plant genome sequence assembly. *Frontiers in plant science*, 9, 1660.
- Lander, E. S., & Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, 121(1), 185-199.
- Leijten, W., Koes, R., Roobeek, I., & Frugis, G. (2018). Translating flowering time from *Arabidopsis thaliana* to *Brassicaceae* and *Asteraceae* crop species. *Plants*, 7(4), 111.
- Li, J., Manghwar, H., Sun, L., Wang, P., Wang, G., Sheng, H., . . . Rui, H. (2019). Whole genome sequencing reveals rare off-target mutations and considerable inherent genetic or/and somaclonal variations in CRISPR/Cas9-edited cotton plants. *Plant Biotechnology Journal*, 17(5), 858-868.
- Liu, C., Qu, X., Zhou, Y., Song, G., Abiri, N., Xiao, Y., . . . Yang, D. (2018). OsPRR37 confers an expanded regulation of the diurnal rhythms of the transcriptome and photoperiodic flowering pathways in rice. *Plant, Cell & Environment*, 41(3), 630-645.
- Luo, Y., Guo, Z., & Li, L. (2013). Evolutionary conservation of microRNA regulatory programs in plant flower development. *Developmental Biology*, 380(2), 133-144.

References

- Lutz, U., Pose, D., Pfeifer, M., Gundlach, H., Hagmann, J., Wang, C., . . . Schwechheimer, C. (2015). Modulation of ambient temperature-dependent flowering in *Arabidopsis thaliana* by natural variation of *FLOWERING LOCUS M*. *PLoS Genetics*, 11(10), e1005588.
- Ma, X., Fu, Y., Zhao, X., Jiang, L., Zhu, Z., Gu, P., . . . Tan, L. (2016). Genomic structure analysis of a set of *Oryza nivara* introgression lines and identification of yield-associated QTLs using whole-genome resequencing. *Scientific reports*, 6(1), 1-12.
- Matar, S., Kumar, A., Holtgräwe, D., Weisshaar, B., & Melzer, S. (2021). The transition to flowering in winter rapeseed during vernalization. *Plant, Cell & Environment*, 44(2), 506-518.
- Maughan, P., Smith, S., Rojas-Beltran, J., Elzinga, D., Raney, J., Jellen, E., . . . Fairbanks, D. (2012). Single nucleotide polymorphism identification, characterization, and linkage mapping in quinoa. *The Plant Genome*, 5(3).
- Melini, V., & Melini, F. (2021). Functional components and anti-nutritional factors in gluten-free grains: A focus on quinoa seeds. *Foods*, 10(2), 351.
- Michaels, S. D., & Amasino, R. M. (1999). *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. *The Plant Cell*, 11(5), 949-956.
- Mouradov, A., Cremer, F., & Coupland, G. (2002). Control of flowering time: interacting pathways as a basis for diversity. *The Plant Cell*, 14, S111-S130.
- Munz, M., Wohlers, I., Simon, E., Reinberger, T., Busch, H., Schaefer, A. S., & Erdmann, J. (2020). Qtlizer: comprehensive QTL annotation of GWAS results. *Scientific reports*, 10(1), 1-8.
- Murase, K., Hirano, Y., Sun, T.-p., & Hakoshima, T. (2008). Gibberellin-induced DELLA recognition by the gibberellin receptor GID1. *Nature*, 456(7221), 459-463.
- Mutasa-Göttgens, E. S., Qi, A., Zhang, W., Schulze-Buxloh, G., Jennings, A., Hohmann, U., . . . Hedden, P. (2010). Bolting and flowering control in sugar beet: relationships and effects of gibberellin, the bolting gene *B* and vernalization. *AoB Plants*, 2010.
- Myles, C., & Wayne, M. (2008). Quantitative trait locus (QTL) analysis. *Nature Education* 1, 208.
- Noh, Y. S., Bizzell, C. M., Noh, B., Schomburg, F. M., & Amasino, R. M. (2004). *EARLY FLOWERING 5* acts as a floral repressor in *Arabidopsis*. *The Plant Journal*, 38(4), 664-672.
- Ogata, T., Toyoshima, M., Yamamizo-Oda, C., Kobayashi, Y., Fujii, K., Tanaka, K., . . . Nagatoshi, Y. (2021). Virus-Mediated Transient Expression Techniques Enable Functional Genomics Studies and Modulations of Betalain Biosynthesis and Plant Height in Quinoa. *Frontiers in plant science*, 12.
- Opabode, J. T. J. B., & Reviews, M. B. (2006). *Agrobacterium*-mediated transformation of plants: emerging factors that influence efficiency. 1(1), 12-20.
- Park, Y.-J., Lee, J.-H., Kim, J. Y., & Park, C.-M. (2019). Alternative RNA splicing expands the developmental plasticity of flowering transition. *Frontiers in plant science*, 10, 606.
- Patiranage, D. S., Asare, E., Maldonado-Taipe, N., Rey, E., Emrani, N., Tester, M., & Jung, C. (2021). Haplotype variations of major flowering time genes in quinoa unveil their role in the adaptation to different environmental conditions. *Plant, Cell & Environment*.
- Patirange, D. S., Rey, E., Emrani, N., Wellman, G., Schmid, K., Schmöckel, S. M., . . . Jung, C. (2020). Genome-wide association study in the pseudocereal quinoa reveals selection pattern typical for crops with a short breeding history. *bioRxiv*.
- Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, 33(3), 290-295.

References

- Peterson, A. J., & Murphy, K. M. (2015). Quinoa cultivation for temperate North America: Considerations and areas for investigation. *Quinoa: Improvement and sustainable production*, 173-192.
- Pin, P. A., Benlloch, R., Bonnet, D., Wremmerth-Weich, E., Kraft, T., Gielen, J. J., & Nilsson, O. (2010). An antagonistic pair of FT homologs mediates the control of flowering time in sugar beet. *Science*, 330(6009), 1397-1400.
- Pin, P. A., Zhang, W., Vogt, S. H., Dally, N., Büttner, B., Schulze-Buxloh, G., . . . Dohm, J. C. (2012). The role of a pseudo-response regulator gene in life cycle adaptation and domestication of beet. *Current Biology*, 22(12), 1095-1101.
- Pose, D., Verhage, L., Ott, F., Yant, L., Mathieu, J., Angenent, G. C., . . . Schmid, M. (2013). Temperature-dependent regulation of flowering by antagonistic *FLM* variants. *Nature*, 503(7476), 414-417.
- Risi, J., Galwey, N., Wickens, G., Haq, N., & Day, P. (1989). New crops for food and industry. In.
- Roe, J. L., Rivin, C. J., Sessions, R. A., Feldmann, K. A., & Zambryski, P. C. (1993). The *Tousled* gene in *A. thaliana* encodes a protein kinase homolog that is required for leaf and flower development. *Cell*, 75(5), 939-950.
- Rojas, W., Pinto, M., Alanoca, C., Gomez Pando, L., Leon-Lobos, P., Alercia, A., . . . Bazile, D. (2015). Quinoa genetic resources and *ex situ* conservation. State of the art report on quinoa around the world in 2013.
- Sedaghati, B., Haddad, R., & Bandehpour, M. (2021). Development of an efficient in-planta *Agrobacterium*-mediated transformation method for Iranian purslane (*Portulaca oleracea* L.) using sonication and vacuum infiltration. *Acta Physiologiae Plantarum*, 43(2), 1-9.
- Shah, S., Weinholdt, C., Jedrusik, N., Molina, C., Zou, J., Große, I., . . . Emrani, N. (2018). Whole-transcriptome analysis reveals genetic factors underlying flowering time regulation in rapeseed (*Brassica napus* L.). *Plant, Cell & Environment*, 41(8), 1935-1947.
- Shi, P., & Gu, M. (2020). Transcriptome analysis and differential gene expression profiling of two contrasting quinoa genotypes in response to salt stress. *BMC Plant Biology*, 20(1), 1-15.
- Shindo, C., Aranzana, M. J., Lister, C., Baxter, C., Nicholls, C., Nordborg, M., & Dean, C. (2005). Role of *FRIGIDA* and *FLOWERING LOCUS C* in determining variation in flowering time of *Arabidopsis*. *Plant Physiology*, 138(2), 1163-1173.
- Singh, B., & Singh, A. K. (2015). Marker-assisted plant breeding: principles and practices: Springer.
- Štorchová, H., Hubáčková, H., Abeyawardana, O. A., Walterová, J., Vondráková, Z., Eliášová, K., & Mandák, B. (2019). *Chenopodium ficifolium* flowers under long days without upregulation of *FLOWERING LOCUS T (FT)* homologs. *Planta*, 250(6), 2111-2125.
- Subedi, M., Neff, E., & Davis, T. M. (2021). Developing *Chenopodium ficifolium* as a potential B genome diploid model system for genetic characterization and improvement of allotetraploid quinoa (*Chenopodium quinoa*). *BMC Plant Biology*, 21(1), 1-18.
- Telahigue, D., & Toumi, L. (2017). Influence of medium and growth regulators on callogenesis of quinoa (*Chenopodium quinoa* Willd.) and effect of hydrous stress induced by PEG 6000 on the callus. *Horticultural Biotechnology Research*, 01-09.
- Tian, H., Li, Y., Wang, C., Xu, X., Zhang, Y., Zeb, Q., . . . Li, L. (2021). Photoperiod-responsive changes in chromatin accessibility in phloem companion and epidermis cells of *Arabidopsis* leaves. *The Plant Cell*, 33(3), 475-491.

References

- Tong, Z., Jiang, S., He, W., Chen, X., Yin, L., Fang, D., . . . Zeng, J. (2021). Construction of high-density genetic map and QTL mapping in *Nicotiana tabacum* backcrossing BC4F3 population using whole-genome sequencing. *Czech Journal of Genetics and Plant Breeding*, 57(3), 102-112.
- Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M. J., . . . Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, 28(5), 511-515.
- Tsuji, H., Taoka, K.-i., & Shimamoto, K. (2011). Regulation of flowering in rice: two florigen genes, a complex gene network, and natural variation. *Current Opinion in Plant Biology*, 14(1), 45-52.
- Valverde, F., Mouradov, A., Soppe, W., Ravenscroft, D., Samach, A., & Coupland, G. (2004). Photoreceptor regulation of *CONSTANS* protein in photoperiodic flowering. *Science*, 303(5660), 1003-1006.
- Van Dam, S., Vosa, U., van der Graaf, A., Franke, L., & de Magalhaes, J. P. (2018). Gene co-expression analysis for functional classification and gene-disease predictions. *Briefings in Bioinformatics*, 19(4), 575-592.
- Vasudevan, V., Siva, R., Krishnan, V., & Manickavasagam, M. (2020). Polyamines, sonication and vacuum infiltration enhances the *Agrobacterium*-mediated transformation in watermelon (*Citrullus lanatus* Thunb.). *South African Journal of Botany*, 128, 333-338.
- Vita, F., Ghignone, S., Bazihizina, N., Rasouli, F., Sabbatini, L., Kiani-Pouya, A., . . . Mancuso, S. (2021). Early responses to salt stress in quinoa genotypes with opposite behavior. *Physiologia Plantarum*, 173(4), 1392-1420.
- Voichkek, Y., & Weigel, D. (2020). Identifying genetic variants underlying phenotypic variation in plants without complete genomes. *Nature genetics*, 52(5), 534-540.
- Wang, B., Zhu, Y., Zhu, J., Liu, Z., Liu, H., Dong, X., . . . Gao, C. (2018). Identification and fine-mapping of a major maize leaf width QTL in a re-sequenced large recombinant inbred lines population. *Frontiers in plant science*, 9, 101.
- Wang, H., Xu, X., Vieira, F. G., Xiao, Y., Li, Z., Wang, J., . . . Chu, C. (2016). The power of inbreeding: NGS-based GWAS of rice reveals convergent evolution during rice domestication. *Molecular Plant*, 9(7), 975-985.
- Wang, Q., Zuo, Z., Wang, X., Gu, L., Yoshizumi, T., Yang, Z., . . . Han, Y.-J. (2016). Photoactivation and inactivation of *Arabidopsis* cryptochrome 2. *Science*, 354(6310), 343-347.
- Wang, Y.-Y., Xiong, F., Ren, Q.-P., & Wang, X.-L. (2019). Regulation of flowering transition by alternative splicing: the role of the U2 auxiliary factor. *Journal of Experimental Botany*, 71(3), 751-758. doi:10.1093/jxb/erz416 %J Journal of Experimental Botany
- Ward, S. M. (1998). A new source of restorable cytoplasmic male sterility in quinoa. *Euphytica*, 101(2), 157-163.
- Wei, J., Li, J., Yu, J., Cheng, Y., Ruan, M., Ye, Q., . . . Equipment, B. (2020). Construction of high-density bin map and QTL mapping of horticultural traits from an interspecific cross between *Capsicum annuum* and Chinese wild *Capsicum frutescens*. 34(1), 549-561.
- Wickland, D. P., Battu, G., Hudson, K. A., Diers, B. W., & Hudson, M. (2017). A comparison of genotyping-by-sequencing analysis methods on low-coverage crop datasets shows advantages of a new workflow, GB-eaSy. *BMC Bioinformatics*, 18(1), 1-12.

References

- Wu, Q., Bai, X., Wu, X., Xiang, D., Wan, Y., Luo, Y., . . . Qin, P. (2020). Transcriptome profiling identifies transcription factors and key homologs involved in seed dormancy and germination regulation of *Chenopodium quinoa*. *Plant Physiology Biochemistry*, 151, 443-456.
- Wu, Q., Bai, X., Zhao, W., Shi, X., Xiang, D., Wan, Y., . . . Peng, L. (2019). Investigation into the underlying regulatory mechanisms shaping inflorescence architecture in *Chenopodium quinoa*. *BMC genomics*, 20(1), 1-20.
- Wu, Q., Luo, Y., Wu, X., Bai, X., Ye, X., Liu, C., . . . Zou, L. (2021). Identification of the specific long-noncoding RNAs involved in night-break mediated flowering retardation in *Chenopodium quinoa*. *BMC genomics*, 22(1), 1-18.
- Wu, Z., Fang, X., Zhu, D., & Dean, C. (2020). Autonomous pathway: *FLOWERING LOCUS C* repression through an antisense-mediated chromatin-silencing mechanism. *Plant Physiology*, 182(1), 27-37.
- Yamaguchi, S. (2008). Gibberellin metabolism and its regulation. *Annu. Rev. Plant Biol.*, 59, 225-251.
- Zeng, Z.-B. (1994). Precision mapping of quantitative trait loci. *Genetics*, 136(4), 1457-1468.
- Zhang, Y.-M., & Xu, S. (2004). Mapping quantitative trait loci in F2 incorporating phenotypes of F3 progeny. *Genetics*, 166(4), 1981-1993.
- Zhou, C.-M., Zhang, T.-Q., Wang, X., Yu, S., Lian, H., Tang, H., . . . Wang, J.-W. (2013). Molecular basis of age-dependent vernalization in *Cardamine flexuosa*. *Science*, 340(6136), 1097-1100.
- Zhou, Y., Zhao, X., Li, Y., Xu, J., Bi, A., Kang, L., . . . Wang, Y.-g. (2020). Triticum population sequencing provides insights into wheat adaptation. *Nature genetics*, 52(12), 1412-1422.
- Zhu, C., Huang, J., & Zhang, Y.-M. (2007). Mapping binary trait loci in the F2: 3 design. *Journal of Heredity*, 98(4), 337-344.

10. Supplementary data on CD/DVD

The following supplemental data are available on a DVD and can be distributed upon request (contact: Prof. Dr. Christian Jung, c.jung@plantbreeding.uni-kiel.de).

File name	Content	Format
Chapter 2 Supplementary tables	Supplementary Table 7 and Supplementary Table 9	.xlsx
Chapter 2 Supplementary figures	Supplementary Figure 7 and Supplementary Figure 10	.pdf
Chapter 3 Supplementary tables	Supplementary Tables 12 to 21	.xlsx
Phenotype and genotype data	Genotypes of the F ₂ population and phenotypes of the F ₂ and F ₃ populations used for QTL mapping (5,218 binned markers)	.csv
Data for genetic map construction	Raw data for genetic map construction with MSTmap (133,913 markers)	.map
QTL mapping README	Captions of phenotype and genotype data, and data for genetic map construction	.txt
Seed codes	Seed codes of all plant material used in this study	.xlsx

11. Curriculum Vitae and Publications

11.1. Curriculum vitae

Personal Data

Name:	Nathaly Fernanda Maldonado Taipe
Date of birth:	07 May 1991
Place of birth:	Quito, Ecuador
Nationality:	Ecuador

Education

03/2019 - present	Ph.D. student at the Plant Breeding Institute, University of Kiel. Research title: "Identification of agronomically important QTL in quinoa and their application in practical quinoa breeding" Supervisors: Dr. Nazgol Emrani, Prof. Dr. Christian Jung
03/2018 – 10/2018	Master's thesis at the Plant Breeding Institute, University of Kiel Research title: "Expression analysis of flowering time genes in quinoa (<i>Chenopodium quinoa</i>)" Supervisors: Dr. Nazgol Emrani, Prof. Dr. Christian Jung
10/2016 – 02/2018	Master of Science in AgriGenomics, University of Kiel
09/2009 – 03/2016	Bachelor of Engineering in Agroindustry, Faculty of Chemistry and Agroindustry, Escuela Politécnica Nacional, Quito - Ecuador Research title: "Development of a carrageenan-based edible coating for minimally processed pineapple" Supervisors: Dr. Silvia Valencia

11.2. Publications

Articles

Maldonado-Taipe, N., Patiranage, D. S., Schmöckel, S. M., Jung, C., & Emrani, N. (2021). Validation of suitable genes for normalization of diurnal gene expression studies in *Chenopodium quinoa*. PloS one, 16(3), e0233821.

Patiranage, D. S., Asare, E., **Maldonado-Taipe, N.**, Rey, E., Emrani, N., Tester, M., & Jung, C. (2021). Haplotype variations of major flowering time genes in quinoa unveil their role in the adaptation to different environmental conditions. Plant, Cell & Environment.

Emrani, N., Hasler, M., Patiranage, D. S., **Nathaly, M. T.**, Rey, E., & Jung, C. (2020). An efficient method to produce segregating populations in quinoa (*Chenopodium quinoa*). Plant Breeding, 139(6), 1190-1200.

Maldonado-Taipe, N., & Valencia-Chamorro, S. (2016). Development of a carrageenan-based edible coating for minimally processed pineapple. Revista Agronomía Colombiana.

Oral and poster presentations

Maldonado-Taipe, N., Barbier, F., Schmid, K., Jung, C., Emrani, N. (2022). High-density mapping of quantitative trait loci for agronomically important traits in quinoa. KPC Mini-Symposium, 9 March, Kiel, Germany. Oral presentation.

Maldonado-Taibe, N., Barbier, F., Schmid, K., Jung, C., Emrani, N. (2021). High-density mapping of quantitative trait loci for agronomically important traits in quinoa. XLV Jornadas Nacionales de Biología, 26 November, Quito, Ecuador. Oral presentation.

Maldonado-Taibe, N., Barbier, F., Schmid, K., Jung, C., Emrani, N. (2020). High-density mapping of quantitative trait loci for agronomically important traits in quinoa. KPC Summer School, 2 September, Kiel, Germany. Oral presentation.

Maldonado-Taibe, N., Barbier, F., Schmid, K., Jung, C., Emrani, N. (2020). High-density mapping of quantitative trait loci for agronomically important traits in quinoa. International Quinoa Research Symposium. 17-19 August, Washington, United States. Virtual event. Poster presentation.

Patirange D. S. R, Asare E., **Maldonado-Taibe, N.**, Emrani N., Jung C. (2020). Unraveling genetic mechanisms of flowering time control in quinoa. International Quinoa Research Symposium. 17-19 August, Washington, United States. Virtual event. Poster presentation.

Maldonado, N., Patirange D. S. R, Jung C., Emrani N. (2020). Towards understanding the phenological development of quinoa by expression analysis of putative flowering time genes. International Symposium of the Society for Plant Breeding (GPZ). Digital breeding. 11- 13 February, Tulln, Austria. Poster presentation.

12. Publications and declaration of own contribution

Chapter two has been published. Chapter three was prepared as a manuscript. The respective declarations of co-authorship are presented on the following page. Meanwhile, chapter four comprises the work of the doctoral candidate.

Declaration of co-authorship**1. Doctoral Candidate**

Name: Nathaly Fernanda Maldonado Taipe

2. This co-author declaration applies to the following article:High-density mapping of QTL controlling agronomically important traits in quinoa (*Chenopodium quinoa* Willd.)

The extent of the doctoral candidate's contribution to the article is assessed on the following scale:

- A. Has contributed to the work (0-33%)
 B. Has made a substantial contribution (34-66%)
 C. Did the majority of the work independently (67-100%)

3. Declaration on the individual phases of the scientific work (A,B,C)

Extent

Concept: Formulation of the basic scientific problem based on theoretical questions which require clarification, including a summary of the general questions which, it is assumed, will be answerable via analyses or concrete experiments/investigations

B

Planning: Planning of experiments/analyses and formulation of investigative methodology, including choice of method and independent methodological development, in such a way that the scientific questions asked can be expected to be answered

C





Execution: Involvement in the analysis or the concrete experiments/investigation

C

Manuscript preparation: Presentation, interpretation and discussion of the results obtained in article form

C

4. Signature of all co-authors

Date	Name	Signature
22.04.2022	Nazgol Emrani	
16.5.2022	Christian Jung	
17.05.2022	Federico Barbier	
22.04.2022	Karl Schmid	

5. Signature of doctoral candidate

Date	Name	Signature
22.04.2022	Nathaly Fernanda Maldonado Taipe	

Declaration of co-authorship**1. Doctoral Candidate**

Name: Nathaly Fernanda Maldonado Taipe

2. This co-author declaration applies to the following article:Leaf and shoot apical meristem transcriptomes of quinoa (*Chenopodium quinoa*) in response to photoperiod and plant development

The extent of the doctoral candidate's contribution to the article is assessed on the following scale:

- A. Has contributed to the work (0-33%)
 B. Has made a substantial contribution (34-66%)
 C. Did the majority of the work independently (67-100%)

3. Declaration on the individual phases of the scientific work (A,B,C)

Extent

Concept: Formulation of the basic scientific problem based on theoretical questions which require clarification, including a summary of the general questions which, it is assumed, will be answerable via analyses or concrete experiments/investigations

B

Planning: Planning of experiments/analyses and formulation of investigative methodology, including choice of method and independent methodological development, in such a way that the scientific questions asked can be expected to be answered

C



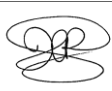
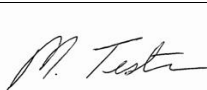
Execution: Involvement in the analysis or the concrete experiments/investigation

C

Manuscript preparation: Presentation, interpretation and discussion of the results obtained in article form

C

4. Signature of all co-authors

Date	Name	Signature
22.04.2022	Nazgol Emrani	
16.5.2022	Christian Jung	
18.5.2022	Elodie Rey	
18.5.2022	Mark Tester	

5. Signature of doctoral candidate

Date	Name	Signature
22.04.2022	Nathaly Fernanda Maldonado Taipe	

13. Acknowledgements

First, I would like to express my gratitude to Prof. Christian Jung for the opportunity to pursue my doctoral research at the Institute of Crop Science and Plant Breeding. I appreciate the resources provided by him for the successful execution of this study on quinoa. I also want to thank Dr. Nazgol Emrani, for sharing her knowledge and guidance, but overall, her time. Thanks for your advice. To both Prof. Jung and Dr. Emrani, I will always treasure the memories in South America.

Then, I thank Dilan, Moni, Fede, Bettina, Brigitte, Birgit, Siegbert, Sarah, Srijan, Verena, Ehsan, Annika and other colleagues at the Institute of Crop Science and Plant Breeding. They were always available when I needed help.

My genuine thanks go to Prof. Dr. Mark Tester, Prof. Dr. Karl Schmid and Dr. Elodie Rey for the excellent collaboration. It was a great pleasure to conduct research with all of you.

I want to thank my parents, the pillars of my life. I thank them for their unconditional love and support. I thank them for motivating me to pursue my dreams. I thank them for being my parents. Love you, Fernando and Ofelia.

I appreciate the financial support by Schleswig-Holstein Stiftung, Proj. 2019/59: Züchtung von Quinoa für den Anbau in Schleswig-Holstein.