Received: 12 February 2024 | Revised: 10 July 2024 | Accepted: 11 July 2024

DOI: 10.1002/pld3.625

RESEARCH ARTICLE

Association mapping and candidate gene identification for yield traits in European hazelnut (*Corylus avellana* L.)

Asena Akköse Baytar[1] | Ertuğrul Gazi Yanar[1] | Anne Frary[1] | Sami Doğanlar[1,2]

[1]Department of Molecular Biology and Genetics, Faculty of Science, Izmir Institute of Technology, Izmir, Türkiye

[2]Plant Science and Technology Application and Research Center, Izmir Institute of Technology, Izmir, Türkiye

Correspondence
Sami Doğanlar, Department of Molecular Biology and Genetics, Faculty of Science, Izmir Institute of Technology, Izmir, Türkiye.
Email: samidoganlar@iyte.edu.tr

Funding information
Izmir Institute of Technology, Grant/Award Number: 2021IYTE-1-00051

Abstract

European hazelnut (*Corylus avellana* L.) is an important nut crop due to its nutritional benefits, culinary uses, and economic value. Türkiye is the leading producer of hazelnut, followed by Italy and the United States. Quantitative trait locus studies offer promising opportunities for breeders and geneticists to identify genomic regions controlling desirable traits in hazelnut. A genome-wide association analysis was conducted with 5,567 single nucleotide polymorphisms on a Turkish core set of 86 hazelnut accessions, revealing 189 quantitative trait nucleotides (QTNs) associated with 22 of 31 traits ($p < 2.9E{-}07$). These QTNs were associated with plant and leaf, phenological, reproductive, nut, and kernel traits. Based on the close physical distance of QTNs associated with the same trait, we identified 23 quantitative trait loci. Furthermore, we identified 23 loci of multiple QTs comprising chromosome locations associated with more than one trait at the same position or in close proximity. A total of 159 candidate genes were identified for 189 QTNs, with 122 of them containing significant conserved protein domains. Some candidate matches to known proteins/domains were highly significant, suggesting that they have similar functions as their matches. This comprehensive study provides valuable insights for the development of breeding strategies and the improvement of hazelnut and enhances the understanding of the genetic architecture of complex traits by proposing candidate genes and potential functions.

KEYWORDS
genotyping by sequencing, GRAS-Di, GWAS, quantitative trait nucleotide (QTN)

1 | INTRODUCTION

European hazelnut (*Corylus avellana* L.) is an important nut crop due to its nutritional and health benefits, culinary applications, and economic value. The nutritional composition of hazelnut includes high levels of healthy fats, essential minerals, vitamins, amino acids, and fiber. Hazelnut consumption is reported to protect against cardiovascular diseases, hyperlipidemia, and diabetes by reducing cholesterol and blood sugar levels (Alasalvar et al., 2003; Brown et al., 1999; Özdemir et al., 1998). It is also beneficial for preventing some types of cancer due to its high dietary fiber content (Alasalvar et al., 2003; Brown et al., 1999; Kritchevsky, 1986; Tariq et al., 2000). Hazelnut has

diverse culinary uses, enhancing the flavor, texture, and nutritional value of a wide array of dishes. It is incorporated in both sweet and savory recipes, including desserts, baked goods, salads, spreads, and sauces. The versatility of hazelnut makes it highly coveted in the culinary world, elevating the taste and appeal of various food products. Additionally, hazelnut has applications in non-food industries like cosmetics (Monika & Anna, 2019), biofuel (Sayar et al., 2019), and woodworking, further enhancing its economic value.

Türkiye is the world leader in hazelnut production, with a 63% share (684,000 t) grown on 738,920 ha harvested in 2021. Türkiye is followed by Italy and the United States with 7.9% and 6.5% of total production (84,670 and 70,310 t), respectively (FAO, 2023). However, yield per hectare is lower in Türkiye (925.7 kg) than the United States (2,848.2 kg) and Italy (1,025.2 kg) (Figure S1a,b). The relatively low yield of hazelnut in Türkiye can be attributed to several factors, including the prevalence of old and unproductive orchards, poor soil, and poor cultivation practices. These factors have hindered growth and productivity, resulting in lower yields (Ekologos., 2020). One strategy for orchard renewal is pruning to remove old stems allowing new suckers to replace them (Silvestri et al., 2021). Because this younger growth is attached to the original root system, nut-bearing can be quicker than that achieved by establishment of a new orchard. However, the renewal of current cultivars may have limited future usefulness given the current status of climate change in hazelnut growing regions.

Climate change poses a serious threat to hazelnut production (An et al., 2020), as well as that of other tree nut crops such as pistachio (Benmoussa et al., 2018), walnut (Gauthier & Jacobs, 2011), and almond (Benmoussa et al., 2018; Parker & Abatzoglou, 2018). Climate change is expressed as changes in temperature, precipitation, relative humidity, and soil conditions, which, in turn, directly affect yield and quality. Hazelnut production has significant environmental impacts, particularly on soil, groundwater, and biodiversity. It is predicted that the Eastern Black Sea coastal strip (up to 250 m elevation), which has long been considered as the prime area for hazelnut production, will increasingly become less productive. Due to regional warming, cultivation is expected to shift to higher elevations and the Western Black Sea Region (Ekologos, 2020). As a result, there is a need for hazelnut cultivars that are adapted to non-traditional environments and that provide higher yield and better quality (Beltramo et al., 2016; Ferrucci et al., 2023; Valentini et al., 2021).

Many agronomic characters are influenced by a complex interplay of multiple genes and environmental factors, which complicates breeding for their improvement. Hazelnut breeding is also hampered by the longer juvenile phase and breeding cycle of trees as compared to annual crops. QTL (quantitative trait loci) studies are a promising area of research for plant geneticists. QTL studies can help identify the specific genomic regions that control desirable traits in hazelnut, such as nut size, shell thickness, and disease resistance. The identification of these regions can facilitate the development of new cultivars. QTL studies are conducted based on genotyping, phenotyping, and mapping and have led to the discovery of several useful marker-trait associated regions in hazelnut such as markers for self-incompatibility (Hill et al., 2021; Rowley et al., 2018) and for eastern filbert blight (EFB)-resistance/tolerance (Honig et al., 2019; Lombardoni et al., 2022; Talbot et al., 2024). Marker-assisted selection (MAS) is a promising approach for cultivar development that utilizes molecular markers linked to desirable traits (Beltramo et al., 2016).

To date, several QTL studies have been reported in *C. avellana*. The first QTL mapping was performed by Beltramo et al. (2016) in a population of 163 F1 progenies from a cross of Tonda Gentile delle Langhe x Merveille de Bollwiller (TGdL x MB). They constructed a linkage map covering 663.1 cM using 152 simple sequence repeat (SSR) markers and identified 15 QTLs for vegetative traits (vigor, sucker habit, and time of bud burst). Torello Marinoni et al. (2018a, 2018b) performed two QTL studies in the same F1 population but included more seedlings (213 individuals) (Torello Marinoni et al., 2018a, 2018b). These researchers constructed high density linkage maps for each parent using 1,236 and 1,211 markers (SSR and single nucleotide polymorphism, SNP) in TGdL and MB, respectively. They identified 29 QTLs for time of leaf bud burst and 19 and eight QTLs for morphological and nut traits, respectively (Torello Marinoni et al., 2018a, 2018b). More recently, Valentini et al. (2021) identified a total of 71 QTLs linked to phenology-related traits (25 for dichogamy, 20 for male flowering time, 15 for female flowering time, and 11 for nut maturity) using the two parental maps constructed by Torello Marinoni et al. (2018b). Using an association mapping approach, Frary et al. (2019) studied 390 *C. avellana* accessions (16 cultivars, 232 landraces, and 142 wild genotypes) and revealed 145 QTLs for 44 agro-morphological traits with 30 SSR markers. Another association mapping study (Öztürk et al., 2017) identified 49 marker-trait associations for nut and kernel traits using 49 SSR markers in germplasm consisting of 64 wild and cultivated *C. avellana* accessions. Söylemez (2020) associated five and three marker loci with oil and protein content, respectively, using 30 SSRs on a panel of 96 European hazelnut accessions. Despite this work, more research is needed to better understand other hazelnut traits and how to best utilize these QTLs in breeding programs. Therefore, in order to have a greater understanding of the genetic control of yield and quality, we conducted high-resolution genome-wide association analysis for 31 vegetative, nut, and kernel traits in a Turkish core set of 86 European hazelnut accessions and searched for candidate genes in those regions.

2 | MATERIALS AND METHODS

2.1 | Plant material

The diversity panel consisted of 86 European hazelnut individuals, including 14 cultivars, 48 landraces, and 24 wild accessions (Table S1). The panel was selected from a core set of Turkish national germplasm, which was chosen based on molecular genetic and morphological diversity (Öztürk et al., 2017). Plant material was collected from individual trees at the Hazelnut Research Institute located in Giresun, Türkiye. The Hazelnut Research Institute classified the individuals as cultivar, landrace, or wild.

2.2 | Phenotypic data

The phenotypic data of Çalışkan and Çetiner (1992), which were mainly based on UPOV descriptors (UPOV, 1979), were used in our analyses. Thirty-one characters (Table S2) were used to detect related QTLs in the European hazelnut genome. These characters included seven plant and leaf traits: plant height, plant suckering, plant area, shape of leaf blade, leaf blade size, stomata density of leaf blade, and petiole length. Eight phenological traits were assessed: time of leaf bud burst, time of male flowering, time of female flowering, male inflorescence pollen shedding period, time of female inflorescence receptivity, time of nut cluster filling (visible appearance of nut clusters relative to pollination time), time of kernel ripening, and time of leaf fall. Two reproductive traits were investigated: number of stigma in a bud and pollen germination ratio. Ten nut traits were examined: crop load (yield in kg per plant), tendency toward alternate bearing, percent cluster set (relative to number of pollinated female flowers), involucre length (compared to nut length), average number of nuts in a cluster, nut size, shell thickness of nut, percentage of kernel in the nut, number of double kernels in the nut, and frequency of empty nuts. Finally, four kernel traits were analyzed: kernelweight, kernel cavity diameter, kernel blanching, and fat content of kernel. Boxplots for traits were generated with R software (R Core Team, 2021) (Figure S2). Normality was assessed using the Shapiro–Wilk test with the Stats package in R software.

2.3 | Genotyping by sequencing and data filtration

DNA was extracted using the protocol of Fulton et al. (1995). Briefly, leaf tissue was ground using a mortar and pestle with a CTAB-based buffer and transferred to 2 ml tubes. Tubes were then incubated at 65°C for 30 min. Following lysis, the mixture was subjected to chloroform:isoamyl alcohol (24:1) extraction and centrifuged (10,000 rpm for 5 min) to separate phases. The aqueous phase was transferred to a new tube, and DNA was then precipitated with ice-cold isopropanol. The precipitate was pelleted by centrifugation, and the resulting pellet was washed with 70% ethanol and resuspended in TE buffer.

Genotyping by random amplicon sequencing-direct (GRAS-Di) was used for genotyping by sequencing (GBS) of the genetic materials. Library preparation and 100-bp paired-end sequencing were carried out by Eurofins Genomics (Ebersberg, Germany) using the Illumina HiSeq4000 system and v. 4 chemistry. SNP genotyping and data filtering were carried out based on the recommendations of O'Leary et al. (2018), as detailed in a previous study (Yanar et al., 2023). Briefly, SNP filtering utilized VCFtools and vcflib to eliminate low-quality data and ensure robust SNP calls. An R script, modified from TinHan et al. (2020) automated the filtering process. The first steps of filtering included removal of indels, accessions with high levels of missing data, and loci with low quality scores. Genotypic values with low genotype quality scores or low genotype depth were recorded as missing data. Additional filtering excluded loci with excessive missing data, insufficient mean depth, and excessive variation in read coverage. Using vcflib, SNPs were then filtered according to their INFO tag information by first removing loci outside specified thresholds for allele balance and quality-to-depth ratios. Further filtering targeted loci based on the specified thresholds for the ratio between mapping quality of alternate allele to reference allele and their properly paired status. A custom filter based on the quality to locus depth ratio, advised by Li (2014), further refined the dataset. Next, loci with no more than 10% missing genotypes were retained in the dataset. Final filtering steps involved removing loci suspected of being paralogous by using rad_haplotyper (Willis et al., 2017) and applying a minor allele count threshold. The resulting SNP dataset was mapped to the Tombul reference genome (Lucas et al., 2021), as described in Yanar et al. (2023), using BLASTN (Camacho et al., 2009) to determine their chromosomal locations. The resulting SNP dataset was then utilized in subsequent association analyses.

2.4 | Association analysis

Linkage disequilibrium (LD) was estimated with the software package PLINK 1.9 (Chang et al., 2015). A LD decay plot was generated with R software. Phenotypic data for 86 accessions were used in the association analysis. Genome-wide association mapping was conducted using the Heteroskedastic Linear Mixed Model (HLMM) and hetlm software with default settings (Young et al., 2018). The HLMM method was applied to calculate significant associations due to the non-normal distribution of the phenotypic data in the population (Table S2). Bonferroni correction (.05/[5,567 × 31]) was applied to reduce false positive rates. Quantile-quantile (Q-Q) plots for each trait were generated with R software (R Core Team, 2021). Manhattan plots were constructed with the CMplot package (Yin et al., 2021). Quantitative trait nucleotides (QTNs) associated with the same trait and in close proximity (20 kb) were merged and designated as a single QTL. The 20 kb threshold was chosen based on the linkage disequilibrium analysis results.

2.5 | Candidate gene identification and putative prediction of function

Candidate gene identification for the identified SNP dataset was performed with the SnpEff program (Cingolani et al., 2012) using the Tombul reference genome annotation (Lucas et al., 2021). Briefly, a SnpEff database was built with the reference genome files using the "build" command. SNPs were then imported into the SnpEff software as a VCF file to be annotated using the SnpEff database. Candidate gene protein sequences were searched against Pfam version 35.0 (Mistry et al., 2021) and NCBI's Conserved Domain Database (CDD) (Marchler-Bauer et al., 2015) to identify significant conserved domains and to predict their functions. Additionally, functional annotation of candidate genes was conducted using OmicsBox 3.2.2 software (Götz et al., 2008), incorporating BLASTp searches against the "nr

Viridiplantae" database, followed by GO mapping, InterProScan for domain identification, and enrichment analyses utilizing both GO and Gramene (Plant Reactome) databases.

3 | RESULTS

3.1 | Linkage disequilibrium and association analysis

The LD decay graph was plotted as r^2 against megabase (Mb) over a distance of 8 Mb and decayed rapidly at around 20 kb ($r^2 < .10$) for the hazelnut germplasm (Figure S3). The graph shows that there was significant LD between genetic markers over short physical distances, up to ~20 kb ($R^2 < .10$), but LD declined quickly as the physical distance increased.

Manhattan plots demonstrated significant SNP-trait linkages for all 31 characters when the less stringent significance threshold was used (uncorrected: $p = .0001$) in the hazelnut panel with 5,567 SNP markers (Figure S4). However, when the more stringent threshold of Bonferroni correction ($p = 2.9E{-}07$) was applied, 189 quantitative trait nucleotides (QTNs) were detected for 22 of the 31 characters. We found no QTN for nine traits after Bonferroni correction: plant height, shape of leaf blade, stomata density of leaf blade, involucre length (compared to nut length), crop load, tendency toward alternate bearing, percent cluster set (compared to female flowers), frequency of empty nuts, and kernel fat content. Q-Q plots compared observed and expected significance of markers for each trait (Figure S5). The significant QTNs were distributed across all 11 *C. avellana* chromosomes (Chr). Chr04 and Chr05 harbored the most QTNs, with 24 and 27 QTNs, respectively (Figure S6). QTNs for the same trait that mapped within 20 kb of each other and had the same direction of effect (positive or negative) were merged into a single QTL. The results are summarized according to the QTNs and QTLs detected for each trait with an emphasis on QTLs because they represent clusters of SNPs with significant associations with a single trait and, therefore, those loci are potentially more reliable than QTNs. As a result of this consolidation, 23 QTLs were identified, and the number of QTNs was reduced to 128 (Table S3).

A total of 21 QTN and five QTL were identified for four plant and leaf traits (Table S3 and Figure S7). Plant suckering was associated with two QTNs on Chr02 and Chr04 and one QTL that encompassed three SNPs on Chr01. Plant area was associated with only one QTN on Chr05. In contrast, leaf blade size was associated with 16 QTNs distributed on eight different chromosomes and three QTLs. The QTLs mapped to Chr07, Chr08, and Chr10 with the locus on Chr10 spanning five SNPs in a region of nearly 20 kb, the largest QTL identified in this study. Petiole length was associated with two QTN each on Chr05 and Chr09 and one QTL on Chr01, which was composed of two adjacent SNPs.

Eight phenological traits were associated with 66 QTNs and 13 QTLs. The most marker-trait associations were detected for time of male flowering with 22 QTNs and three QTLs. The QTNs were distributed on seven chromosomes with six and five QTNs on Chr05 and Chr08, respectively. The male flowering time QTLs were located on Chr02 and Chr05 (one and two QTLs, respectively), and each encompassed two or three SNPs. The Chr02 QTL was the second largest locus identified in this study, spanning 11.7 kb. Time of nut cluster filling was the phenological trait with the second most associations: 16 QTNs and two QTLs. The QTNs were located on eight chromosomes with multiple QTNs on Chr03, Chr04, Chr05, and Chr06. The filling QTLs mapped to Chr06 and Chr07 with two associated SNPs at each locus. Leaf fall was associated with 11 QTNs on six different chromosomes and one QTL on Chr02. The most QTLs for a phenological trait were identified for time of ripening with four QTLs distributed on Chr02, Chr03, Chr07, and Chr11. While most of these QTLs encompassed only two SNPs, six trait-associated SNPs were in the QTL region on Chr03. Five QTNs and three QTLs were detected for time of female inflorescence receptivity. The QTNs were located on Chr01, Chr03, and Chr05, while the QTLs were on Chr05, Chr07, and Chr09. Female flowering time was associated with five QTNs on Chr01, Chr02, Chr07, and Chr08. Single QTNs were identified for the remaining traits: time of leaf bud burst (Chr10) and pollen shedding period (Chr03).

Only QTNs were detected for the reproductive traits: two for number of stigma in a bud (Chr03 and Chr05) and three for pollen germination ratio (Chr07, Chr10, and Chr11). Marker-trait associations were identified for five nut traits: 21 QTNs and two QTLs. Shell thickness yielded the most loci with 12 QTNs and two QTLs. The QTNs were distributed across eight chromosomes with three QTNS each on Chr02 and Chr04. The shell thickness QTLs were on Chr04 and Chr05 with four and two associated SNPs, respectively, at these locations. Kernel percentage (percent kernel in the nut) was associated with four QTNs with two each on Chr01 and Chr04. Three QTNs were identified for average number of nuts per cluster with one QTN each on Chr05, Chr06, and Chr08. Nut size and number of double kernels were associated with one QTN each with these located on Chr06 and Chr04, respectively.

Fifteen QTNs and three QTLs were identified for the three kernel traits with kernel cavity diameter yielding the most marker-trait associations. Ten kernel cavity diameter QTNs were distributed on seven chromosomes with two on Chr04 and three on Chr09. Three QTLs were identified for this trait on Chr05, Chr06, and Chr08 with four SNPs associated with the Chr06 QTL. Two QTNs on Chr04 and Chr07 were associated with kernel weight, while three were associated with kernel blanching (Chr01 and two QTNs on Chr06).

We detected some chromosomal locations that were associated with more than one trait at close proximity or even at the same position (Figure 1 and Table 1), hereafter called Loci of Multiple-QTs (LMQ). Twenty-three LMQs accounting for 54 marker-trait associations were identified for the traits and were distributed across nine chromosomes. Only Chr09 and Chr11 did not harbor any LMQ. For the sake of brevity, only noteworthy LMQs are described here. For example, a LMQ on Chr01, LMQ1.3 was associated with the timing of both male and female flowering as well as kernel blanching and leaf blade size. Another flowering-related LMQ was located on Chr03 (LMQ3.1) and was associated with number of stigma in a bud, male

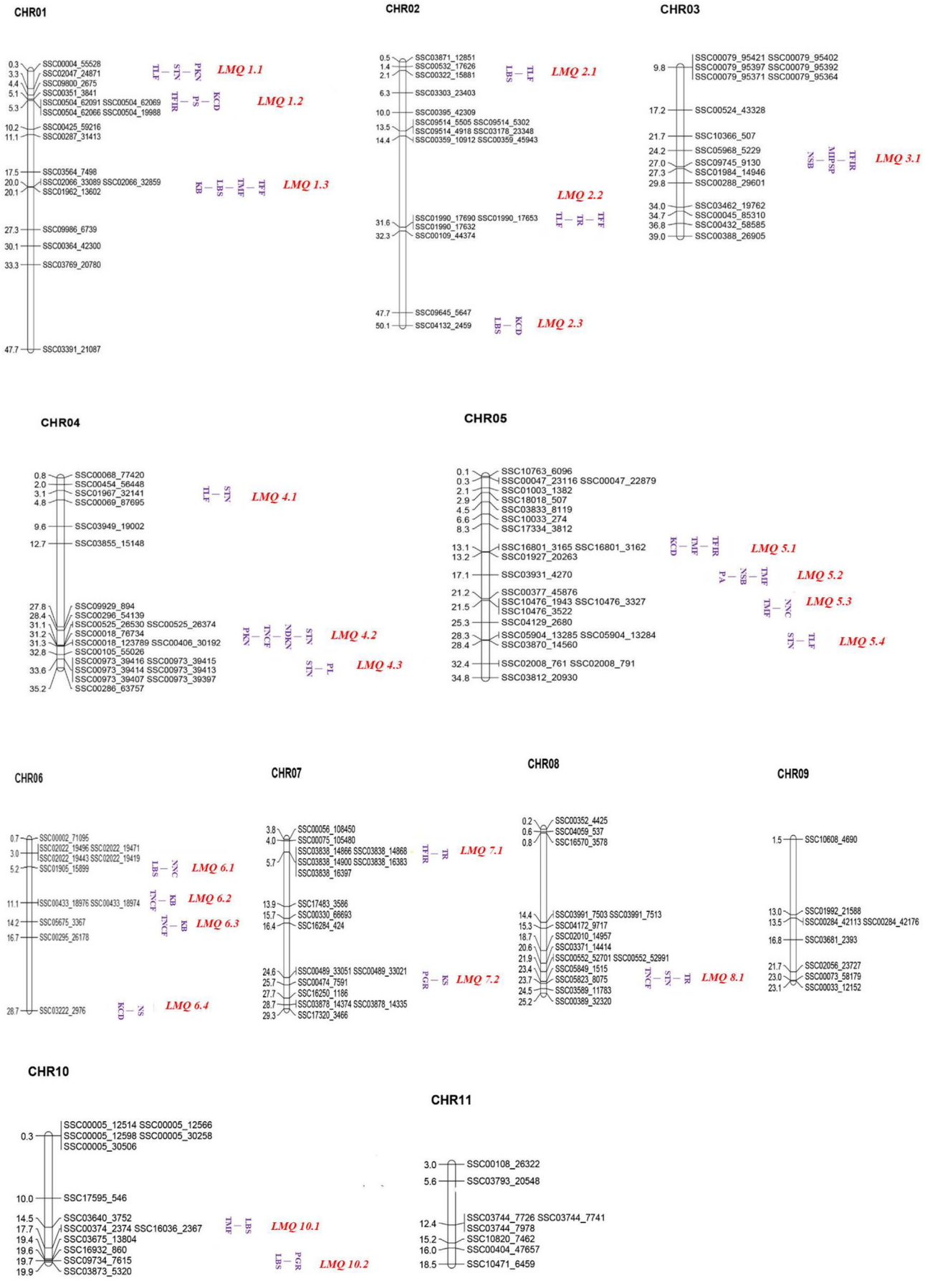
CHR01
0.3 SSC00004_55528
3.3 SSC02047_24871
4.4 SSC09800_2675
5.1 SSC00351_3841
5.3 SSC00504_62091 SSC00504_62069
SSC00504_62066 SSC00504_19988
10.2 SSC00425_59216
11.1 SSC00287_31413
17.5 SSC03564_7498
20.0 SSC02066_33089 SSC02066_32859
20.1 SSC01962_13602
27.3 SSC09986_6739
30.1 SSC00364_42300
33.3 SSC03769_20780
47.7 SSC03391_21087
TLF SSN PKN LMQ 1.1
TFTR PS KCD LMQ 1.2
KB LBS TMF TFF LMQ 1.3
CHR02
0.5 SSC03871_12851
1.4 SSC00532_17626
2.1 SSC00322_15881
6.3 SSC03303_23403
10.0 SSC00395_42309
13.5 SSC09514_5505 SSC09514_5302
14.4 SSC09514_4918 SSC03178_23348
SSC00359_10912 SSC00359_45943
31.6 SSC01990_17690 SSC01990_17653
32.3 SSC01990_17632
SSC00109_44374
47.7 SSC09645_5647
50.1 SSC04132_2459
LBS TLF LMQ 2.1
TLF TR TFF LMQ 2.2
LBS KCD LMQ 2.3
CHR03
9.8 SSC00079_95421 SSC00079_95402
SSC00079_95397 SSC00079_95392
SSC00079_95371 SSC00079_95364
17.2 SSC00524_43328
21.7 SSC10366_507
24.2 SSC05968_5229
27.0 SSC09745_9130
27.3 SSC01984_14946
29.8 SSC00288_29601
34.0 SSC03462_19762
34.7 SSC00045_85310
36.8 SSC00432_58585
39.0 SSC00388_26905
NSN MIPSP TFTR LMQ 3.1
CHR04
0.8 SSC00068_77420
2.0 SSC00454_56448
3.1 SSC01967_32141
4.8 SSC00069_87695
9.6 SSC03949_19002
12.7 SSC03855_15148
27.8 SSC09929_894
28.4 SSC00296_54139
31.1 SSC00525_26530 SSC00525_26374
31.2 SSC00018_76734
31.3 SSC00018_123789 SSC00406_30192
32.8 SSC00105_55026
33.6 SSC00973_39416 SSC00973_39415
SSC00973_39414 SSC00973_39413
SSC00973_39407 SSC00973_39397
35.2 SSC00286_63757
TLF SSN LMQ 4.1
PKN TNCF NDKN SSN LMQ 4.2
SSN TL LMQ 4.3
CHR05
0.1 SSC10763_6096
0.3 SSC00047_23116 SSC00047_22879
2.1 SSC01003_1382
2.9 SSC18018_507
4.5 SSC03833_8119
6.6 SSC10033_274
8.3 SSC17334_3812
13.1 SSC16801_3165 SSC16801_3162
13.2 SSC01927_20263
17.1 SSC03931_4270
21.2 SSC00377_45876
21.5 SSC10476_1943 SSC10476_3327
SSC10476_3522
25.3 SSC04129_2680
28.3 SSC05904_13285 SSC05904_13284
28.4 SSC03870_14560
32.4 SSC02008_761 SSC02008_791
34.8 SSC03812_20930
KCD TMF TFTR LMQ 5.1
PA NSN TMF LMQ 5.2
TMF NNC LMQ 5.3
SSN TL LMQ 5.4
CHR06
0.7 SSC00002_71095
3.0 SSC02022_19496 SSC02022_19471
SSC02022_19443 SSC02022_19419
5.2 SSC01905_15899
11.1 SSC00433_18976 SSC00433_18974
14.2 SSC05675_3367
16.7 SSC00295_26178
28.7 SSC03222_2976
LBS NNC LMQ 6.1
TNCF KB LMQ 6.2
TNCF KB LMQ 6.3
KCD SN LMQ 6.4
CHR07
3.8 SSC00056_108450
4.0 SSC00075_105480
5.7 SSC03838_14866 SSC03838_14868
SSC03838_14900 SSC03838_16383
SSC03838_16397
13.9 SSC17483_3586
15.7 SSC00330_66693
16.4 SSC16284_424
24.6 SSC00489_33051 SSC00489_33021
25.7 SSC00474_7591
27.7 SSC16250_1186
28.7 SSC03878_14374 SSC03878_14335
29.3 SSC17320_3466
TFTR TR LMQ 7.1
PGR KS LMQ 7.2
CHR08
0.2 SSC00352_4425
0.6 SSC04059_537
0.8 SSC16570_3578
14.4 SSC03991_7503 SSC03991_7513
15.3 SSC04172_9717
18.7 SSC02010_14957
20.6 SSC03371_14414
21.9 SSC00552_52701 SSC00552_52991
23.4 SSC05849_1515
23.7 SSC05823_8075
24.5 SSC03589_11783
25.2 SSC00389_32320
TNCF SSN TR LMQ 8.1
CHR09
1.5 SSC10608_4690
13.0 SSC01992_21588
13.5 SSC00284_42113 SSC00284_42176
16.8 SSC03681_2393
21.7 SSC02056_23727
23.0 SSC00073_58179
23.1 SSC00033_12152
CHR10
0.3 SSC00005_12514 SSC00005_12566
SSC00005_12598 SSC00005_30258
SSC00005_30506
10.0 SSC17595_546
14.5 SSC03640_3752
17.7 SSC00374_2374 SSC16036_2367
19.4 SSC03675_13804
19.6 SSC16932_860
19.7 SSC09734_7615
19.9 SSC03873_5320
TMF LBS LMQ 10.1
LBS PGR LMQ 10.2
CHR11
3.0 SSC00108_26322
5.6 SSC03793_20548
12.4 SSC03744_7726 SSC03744_7741
SSC03744_7978
15.2 SSC10820_7462
16.0 SSC00404_47657
18.5 SSC10471_6459

FIGURE 1 Locations of 189 significant QTNs on 11 chromosomes. A total of 23 LMQs are also indicated on the map with related traits. KB, kernel blanching; KCD, kernel cavity diameter; KS, kernel weight; LBS, leaf blade size; MIPSP, male inflorescence pollen shedding period; NDKN, number of double kernels in the nut; NNC, number of nuts in a cluster; NS, nut size; NSB, number of stigma in a bud; PA, plant area; PGR, pollen germination ratio; PKN, percentage of kernel in the nut; PL, petiole length; PS, plant suckering; STN, shell thickness of nut; TFF, time of female flowering; TFIR, time of female inflorescence receptivity; TLBB, time of leaf bud burst; TLF, time of leaf fall; TMF, time of male flowering; TNCF, time of nut cluster filling; TR, time of ripening

inflorescence pollen shedding period and female inflorescence receptivity time. LMQ4.2 was associated with multiple nut and kernel traits including nut filling time, shell thickness, kernel percentage, and percent/frequency of double kernels. Four LMQs were identified on Chr05 with LMQ5.1 associated with male flowering time, time of female receptivity, and kernel cavity size. Chr06 also contained four LMQs with two (LMQ6.2 and LMQ6.3) both associated with the same traits: kernel blanching and time of nut filling.

3.2 | Candidate gene identification and protein prediction

Candidate genes for the SNP dataset were identified based on the Tombul reference genome annotation. Conserved domains in the protein sequences of the candidate genes were identified and used to predict protein function. Our search gave hits within or near annotated genes for 188 of 189 QTNs (except SNP SSC00002_71095) (Table S4). Of these, 130 (69%) were in protein coding regions and 58 (31%) in intergenic regions. A total of 159 different candidate genes were identified for the 188 QTNs, with several QTNs detected in the same candidate genes. Among the QTNs found in protein coding regions, 49 were missense (a single amino acid change), 49 were located in introns, and 16 were synonymous variants. The other variants were in 5′ (7 QTNs) and 3′ (5 QTNs) UTRs or represented 5′ UTR premature start codon gain (1 QTN), splice region (1 QTN), and stop-gain (2 QTNs) variants (Table S4).

Of the 159 candidate genes, 122 genes (77%) had significant conserved domains as determined using the pfam and/or NCBI CCD databases (Table S5) with the most probable candidate proteins selected based on statistical significance and length of similarity (Table S6). Only the most likely candidates are described in Section 4.

4 | DISCUSSION

4.1 | LD decay in hazelnut

The rate of LD decay in a population is important for genetic association studies, including QTL mapping and genome-wide association analysis (GWAS). Rapid LD decay enables the identification of more precise and accurate QTLs and genetic variants associated with the traits of interest. Conversely, slow LD decay can lead to false positives or inflated effect sizes for QTLs, resulting in the identification of spurious associations that do not reflect true genetic effects (Gaut & Long, 2003; Soto-Cerda et al., 2013). Therefore, determination of the LD decay rate in a population is crucial for accurate genetic analyses and interpretation of results.

In our hazelnut panel with 5,567 SNP markers, LD decayed below an r^2 value of .10 at less than 20 kb. Beyond this distance, LD decayed rapidly; therefore, each marker provided independent information about the genetic variation in the population. This result is consistent with a previous study in European hazelnut with $r^2 = .33$ at 19.6 kb (Yang et al., 2023). Our LD rate is also comparable with other studies in trees such as walnut, *Juglans regia*, with $r^2 < .15$ at 41.2 kb (Bükücü et al., 2020) and $r^2 < .20$ at 10 kb (Arab et al., 2019). Our LD is longer than measured for *Ostrya rehderiana* ($r^2 < .50$ at ~400 kb, Yang et al., 2018) and shorter than *Eucalyptus globulus* ($r^2 = .2$ within ~4 kb, Butler et al., 2022) and *Pinus taeda* ($r^2 = .1$ within 192 bp, Lu et al., 2016). These findings of rapid LD decay are expected when a panel with high genetic diversity and extensive outcrossing is examined. Hazelnut is self-incompatible and requires cross-pollination by a different compatible type for nut set. In the wild and in collections of diverse genotypes, the female inflorescences may be pollinated by several different individuals. These reproductive strategies promote genetic diversity (Taghavi et al., 2018) and reduce the chance of inbreeding depression in individuals resulting in faster LD decay.

4.2 | Marker-trait associations

The present study is the largest high-resolution genome-wide association analysis conducted to date on European hazelnut accessions, with 31 agro-morphological traits evaluated across 86 individuals. Utilizing more than 5,000 SNP markers, we identified marker-trait associations for important hazelnut traits. Our study yielded many QTNs with p values between .0001 and $p < 2.9E{-}07$; however, we only present an association here if it was significant after Bonferroni correction. In this way, 189 QTNs were identified for 22 plant, leaf, phenological, reproductive, nut, and kernel characteristics. Among these QTNs, several were located at close physical distance and associated with the same trait. These loci were consolidated into 23 QTLs: 5 for plant and leaf traits, 13 for phenological traits, 2 for nut traits, and 3 for kernel traits. Compared to the QTNs, these QTLs are potentially more reliable because they represent statistically significant associations between at least two adjacent markers and the trait of interest. Based on this criterion, the most noteworthy QTL is kernel ripening QTL6.1, which spanned six associated SNPs. Similarly leaf blade size QTL10.1 encompassed five markers, while shell thickness QTL4.1 and kernel cavity diameter QTL6.1 each included four markers.

TABLE 1 Loci harboring multiple QTs identified in this study

LMQ	SNP	Chr	Physical position/range (bp)	Traits
LMQ 1.1	SSC02047_24871	1	3,271,224	TLF, STN, PKN
LMQ 1.2	SSC00351_3841	1	5,108,398–5,327,060	TFIR
	SSC00504_62091			PS
	SSC00504_62069			PS
	SSC00504_62066			PS
	SSC00504_19988			KCD
LMQ 1.3	SSC02066_33089	1	19,995,012–20,088,953	KB
	SSC02066_32859			LBS
	SSC01962_13602			TMF
	SSC02066_32859			TFF
LMQ 2.1	SSC00322_15881	2	2,080,809	LBS, TLF
LMQ 2.2	SSC01990_17632	2	31,610,804–31,610,862	TLF, TR
	SSC01990_17653			TFF
	SSC01990_17690			TLF, TR
LMQ 2.3	SSC04132_2459	2	50,079,549	LBS, KCD
LMQ 3.1	SSC09745_9130	3	26,981,854	MIPSP, TFIR, NSB
LMQ 4.1	SSC01967_32141	4	3,061,884	TLF, STN
LMQ 4.2	SSC00018_123789	4	31,091,524–31,255,449	PKN, STN, TNCF
	SSC00018_76734			NDKN
	SSC00525_26374			TNCF
	SSC00525_26530			PKN
LMQ 4.3	SSC00973_39397	4	33,640,893–33,640,912	STN
	SSC00973_39407			STN
	SSC00973_39413			PL
	SSC00973_39414			PL
	SSC00973_39415			STN
	SSC00973_39416			STN
LMQ 5.1	SSC16801_3165	5	13,055,552–13,235,456	KCD
	SSC16801_3162			KCD
	SSC01927_20263			TFIR, TMF
LMQ 5.2	SSC03931_4270	5	17,103,854	PA, NSB, TMF
LMQ 5.3	SSC10476_1943	5	21,539,749–21,540,633	TMF
	SSC10476_3327			NNC
	SSC10476_3522			TMF
LMQ 5.4	SSC05904_13285	5	28,285,331–28,366,561	STN
	SSC05904_13284			STN
	SSC03870_14560			STN, TLF
LMQ 6.1	SSC01905_15899	6	5,177,289	LBS, NNC
LMQ 6.2	SSC00433_18976	6	11,105,354–11,105,356	KB, TNCF
	SSC00433_18974			TNCF
LMQ 6.3	SSC05675_3367	6	14,158,104	TNCF, KB
LMQ 6.4	SSC03222_2976	6	28,735,711	NS, KCD
LMQ 7.1	SSC03838_14866	7	5,718,978–5,720,221	TFIR
	SSC03838_14868			TFIR
	SSC03838_14900			TFIR
	SSC03838_16383			TR

(Continues)

TABLE 1 (Continued)

LMQ	SNP	Chr	Physical position/range (bp)	Traits
	SSC03838_16397			TR
LMQ 7.2	SSC00474_7591	7	25,676,730	PGR, KS
LMQ 8.1	SSC05823_8075	8	23,742,601	TNCF, STN, TR
LMQ 10.1	SSC00374_2374	10	17,671,230-17,708,891	TMF
	SSC16036_2367			LBS
LMQ 10.2	SSC16932_860	10	19,645,470-19,864,509	LBS
	SSC09734_7615			LBS
	SSC03873_5320			PGR

Abbreviations: KB, kernel blanching;KCD, kernel cavity diameter; KS, kernel weight; LBS, leaf blade size; MIPSP, male inflorescence pollen shedding period; NDKN, number of double kernels in the nut; NNC, number of nuts in a cluster; NS, nut size; NSB, number of stigma in a bud; PA, plant area; PGR, pollen germination ratio; PKN, percentage of kernel in the nut; PL, petiole length; PS, plant suckering; STN, shell thickness of nut; TFF, time of female flowering; TFIR, time of female inflorescence receptivity; TLF, time of leaf fall; TMF, time of male flowering; TNCF, time of nut cluster filling; TR, time of ripening.

Comparing the positions of significant QTLs identified using the same and different types of molecular markers, such as SSRs and SNPs, can offer several advantages and valuable insights. For instance, it may aid in identifying robust and consistent marker-trait associations, which can inform the development of more effective molecular markers for use in MAS or genomic selection programs. As there have been limited agromorphological QTL studies in hazelnut utilizing GWAS, a comprehensive comparison of our findings with the existing literature is not feasible at this time. Nevertheless, we identified the genomic positions of the SSR markers used in our previous work (Frary et al., 2019) by performing a BLAST search of the primer sequences against the reference genome. This allowed us to determine the location of each marker in the hazelnut genome and compare it to the current results. Comparison of the coordinates of the previously identified SSR markers and the SNPs identified in this study and their consistency provides greater confidence in the identified loci. In this way, seven genomic regions in close proximity were identified in both studies highlighting the importance of these regions in contributing to the traits of interest (Table 2). The consistently identified regions are related to plant suckering, flowering time, nut cluster set, leaf fall time, and leaf blade size.

4.3 | Co-localization of QTNs and prediction of gene function

The identification and co-localization of loci containing multiple QTLs have significant implications for breeding programs. The co-localization of QTLs enables breeders to simultaneously select for multiple traits that are controlled by these loci, resulting in increased genetic gain for those traits in segregating populations. Therefore, understanding the genetic architecture of complex traits and identifying co-localized QTLs that are positively correlated can greatly enhance the efficiency and effectiveness of breeding programs (Li et al., 2016; Said et al., 2013). Here, we discuss some noteworthy loci containing multiple trait associations. Additionally, we considered the additive effect of each QTN. The additive effect values of SNPs associated with specific traits are valuable in identifying favorable genetic markers associated with traits, which can be used in breeding programs and to develop new cultivars with improved traits. Also, by selecting individuals with favorable alleles for specific SNPs, breeders can increase the frequency of those alleles in the population. Furthermore, when linked to candidate gene prediction, such knowledge gives insight on the genetic basis of complex traits leading to a better understanding of the pathways involved and potential targets for crop improvement (Chong et al., 2019).

LMQ2.1 (SNP SSC00322_15881) was associated with both leaf blade size and time of leaf fall in this study and in our previous work (Frary et al., 2019). This SNP was located within a gene predicted to code for a serine peptidase (g9250; Tables S4 and S5). Previous studies have suggested that serine proteases may be involved in regulating plant growth and development, as well as regulation of leaf aging and death (Diaz-Mendoza et al., 2016; Roberts et al., 2012). Roberts et al. (2003) purified and characterized the two main serine proteases found only in senescent wheat leaves. These enzymes accounted for over 70% of the total proteolytic activity in senescent leaves. Another study by Chauhan et al. (2009) found that serine proteases were the predominant proteases involved in monocarpic senescence in wheat, which was confirmed through the use of protease inhibitors and a protease inhibition assay. The process of leaf senescence is a natural and controlled aging process that leads to the breakdown of cellular components and, eventually, the death of the leaf. During this process, the nutrients stored in the leaf are broken down and recycled back into the plant (Gregersen et al., 2008; Krupinska et al., 2012). Thus, a later time of leaf fall indicates that the plant is taking more time to recycle the nutrients from its senescing leaves. This concept is supported by our finding that LMQ2.1 was associated with leaf fall.

A genomic region of 58 bp containing three SNPs was designated as LMQ2.2. This region was associated with times of leaf fall, nut ripening, and female flowering, favoring early senescence. These traits are all closely related to the timing and coordination of different stages of plant development. The genomic region containing LMQ2.2 encodes a protein with a conserved domain from a

TABLE 2 Comparison of our findings with previous studies

Associated trait in this study	SNP information				Marker name in Frary et al. (2019)	SSR marker information			Associated trait in Frary et al. (2019)	Associated trait in Beltramo et al. (2016)
	SNP in our study	*p* value	Chromosome	Position		Marker start[a]	Marker end[a]	Length (bp)		
PS	SSC00286_63757	2.91E−12	Chr04	35211750	B635	34420666	34421043	377	PS	-
LBS	SSC00056_108450	1.12E−07	Chr07	3800598	B631	3734330	3734524	194	LBS	-
LBS	SSC00075_105480	5.13E−09	Chr07	3967297						
TMF	SSC00388_26905	5.63E−08	Chr03	38957527	B641a	38761496	38761780	284	IAM	-
MIPSP	SSC09745_9130	1.54E−10	Chr03	26981854	B641b	38761530	38761781	251	IAF, TFIR	-
TFIR	SSC09745_9130	2.14E−10	Chr03	26981854						
TLF	SSC00322_15881	1.15E−08	Chr02	2080809	B660	1981869	1982100	231	TLF	-
NNC	SSC10476_3327	2.10E−07	Chr05	21540438	B625	21426325	21426586	261	TNCF	TC
NNC	SSC05849_1515	3.82E−10	Chr08	23432808	B648	21949832	21950058	226	RNCF, TNCF	-
TNCF	SSC05823_8075	5.71E−08	Chr08	23742601						

Abbreviations: IAM, inflorescences amount male; IAF, inflorescences amount female; LBS, leaf blade size; MIPSP, male inflorescence pollen shedding period; NNC, number of nuts in a cluster; PS, plant suckering; RNCF, Rate of nut cluster formation (compared to female flowers); TC, trunk circumference; TFIR, time of female inflorescence receptivity; TLF, time of leaf fall; TMF, time of male flowering; TNCF, time of nut cluster filling.

[a]The genomic positions of the SSR markers, used in the study by Frary et al. (2019), were identified in this study by performing a BLAST analysis of primer sequences against the reference genome (CavTom2PMs-1.0 GenBank assembly [GCA_901000735.2]) on NCBI.

retrotransposon gag protein (g13253), indicating a possible role in retrotransposon activity regulation (Tables S4 and S5). The retrotransposon-gag superfamily is a diverse group of proteins that are involved in various biological processes such as gene expression, DNA replication, and transposition. These proteins have been reported to play important roles in controlling genome evolution, stress response, and epigenetic regulation in plants (Orozco-Arias et al., 2019). Our results suggest that the LMQ2.2 region may play a significant role in the regulation of plant reproductive development, particularly during the transition from vegetative growth to reproductive growth and, later, senescence through the potential influence of a retrotransposon protein on the regulation of developmental pathways.

The SNP SSC09745_9130 in LMQ3.1 was negatively associated with three traits in this study: stigma number, male inflorescence pollen shedding period, and time of female inflorescence receptivity. The same region was identified as associated with female flower timing traits in previous work (Frary et al., 2019). In hazelnut, the female flower has a stigma, which receives pollen. In the vicinity of this SNP, a gene that potentially codes for a pectate lyase 3-like (PL3) enzyme was detected (g7175). Pectate lyase enzymes, including PL3, are involved in the degradation of pectin, a major component of the plant cell wall. During pollen development and shedding, the cell walls of the anther and pollen grains must undergo extensive remodeling and degradation to allow for proper pollen release. Studies have shown that pectate lyase activity was upregulated in the anthers of plants during pollen development and shedding (Jiang, Yao, Yu, Liang, et al., 2014; Jiang, Yao, Yu, Lv, et al., 2014; Wing et al., 1990; Zheng et al., 2018).

Actin-binding proteins are known to interact with the actin cytoskeleton, which is involved in various aspects of plant growth and development, including cell division, elongation, and differentiation. In our study, LMQ4.2 mapped to the same region as an ACR protein (g25372), a unique group of ACT domain-containing proteins. This group of proteins is characterized by having four copies of the ACT domain that span the entire polypeptide chain (Liu, 2006). Previously, ACR proteins were reported only in *Arabidopsis* and rice. In our study, we identified a single candidate ACR protein in hazelnut in which four ACT-domain repeats were tandemly aligned (Tables S4 and S5 and Figure S8). ACR genes were found to be differentially regulated by plant hormones, salt/cold stress, and light/dark treatment in *Arabidopsis* seedlings grown in tissue culture, which suggests that ACR proteins may function as novel regulatory or sensor proteins in plants (Hsieh & Goodman, 2002; Liu, 2006). These proteins allow plants to respond and adapt to their environment in a highly dynamic and complex manner. By sensing and responding to external cues, plants can adjust their growth and development to optimize their survival and reproduction in changing environmental conditions. In our study, two SNPs in LMQ4.2 were positively associated with percentage of kernel in the nut and time of nut cluster filling and were found in the predicted-ACR protein. Therefore, there may be a relationship between the time of nut cluster filling and sensor protein function, as the formation and development of nut clusters in plants can be regulated by various environmental and hormonal factors. Some of these factors may influence the expression or activity of sensor proteins, which could then affect downstream signaling pathways involved in nut development and maturation.

In our study, six different SNPs, located within a 19 bp region, were designated as LMQ4.3 and were negatively associated with nut shell thickness and positively to petiole length. We found that LMQ4.3 was located in close proximity to a gene that is predicted to code for a xyloglucan endotransglycosylase (XET) (g25732-g25733). This suggests that the identified SNPs may be associated with alterations in the function or expression of this XET gene, which could, in turn, affect both shell thickness and petiole length. XETs are important for plant growth and development and have been shown to play roles in cell wall degradation (Nishitani, 1997; van Sandt et al., 2007), cell expansion (van Sandt et al., 2007; Vissenberg et al., 2000), fruit ripening (Miedes et al., 2010; Nishikubo et al., 2007, 2011), xylem tissue development (Vissenberg et al., 2000), vascular tissue differentiation (Vissenberg et al., 2000), and response to environmental stresses (Xu et al., 1996). XET enzymes can break down cell walls, which facilitates the creation of the xylem and phloem conducting structures and intercellular spaces. In our study, LMQ4.3 was identified to negatively affect shell thickness and positively affect petiole length, which is consistent with its role in cell wall degradation.

The present study also revealed the presence of a SNP, SSC01905_15899, designated as LMQ6.1, in a transcript region that was predicted to be part of the *PIN4* gene (g41260). PIN4 protein is a member of the PIN family of auxin efflux transporters in plants. These proteins play a crucial role in regulating auxin transport, which is important for various developmental processes in plants such as root and shoot growth, organ formation, plant architecture, and tropisms. Previous studies have shown that *PIN* genes play a role in plant architecture (Chen et al., 2012; Xu et al., 2005) and canopy size affecting leaf blade size, leaf length, and branching (Borrell et al., 2022). Similarly, in our study, we found that the SNP was associated with reduced leaf blade size and increased number of nuts in a cluster. Thus, it is possible that energy and resources that would have been allocated to the development of larger leaves are instead directed toward the development of more nut.

Two SNPs in LMQ7.1 were associated with earlier nut maturity in our study. This region was designated as a candidate gene (g37162), which is predicted to be in the trehalose phosphatase family. Trehalose-6-phosphate phosphatase (TPP) is an important enzyme involved in the regulation of plant metabolism and development, as it has a critical role in the trehalose phosphate signaling pathway, which is involved in the regulation of various processes such as carbon partitioning, growth, and stress responses. Previous studies reported that changing TPP activity can affect biomass, inflorescence development, and senescence in plants (Meyer et al., 2007; Paul et al., 2008; Pellny et al., 2004; Ponnu et al., 2011; Zhang et al., 2009). These studies suggest that trehalose metabolism can play a role in fruit ripening and other plant processes. Therefore, the SNP loci identified herein may be important for breeding as ripening time is an important factor in hazelnut production.

5 | CONCLUSION

This genome-wide association analysis in hazelnut identified numerous QTNs associated with 22 traits. The consolidation of nearby QTNs into QTLs and comparison with previously mapped loci increases the reliability of the results and helps prioritize genomic regions for further study. Moreover, the significant association of single SNPs with different traits identified in this study can have implications for understanding the genetic control of pleiotropy and potentially enable breeders to simultaneously select for multiple traits, resulting in increased genetic gain. Similarly, the identification of candidate genes for selected loci sheds light on their hypothetical functions. However, further investigation is necessary to confirm the predicted gene/protein functions and elucidate their precise mechanisms to fully uncover their impact on plant growth and development. In addition, the LD decay analysis conducted in our study provides valuable insights into the extent of linkage disequilibrium in hazelnut germplasm, which can inform the design of future association mapping and genomic selection studies. Overall, our findings contribute to a better understanding of the genetic architecture of complex traits in hazelnut and might be useful for hazelnut geneticists and breeders.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

AUTHOR CONTRIBUTIONS

Asena Akköse Baytar: Data analysis; interpretation of data; manuscript draft and revision. **Ertuğrul Gazi Yanar**: Molecular data analysis; manuscript revision. **Anne Frary**: Conception and design; interpretation of data; manuscript revision. **Sami Doğanlar**: Conception and design; funding acquisition; manuscript revision. All authors gave final approval of the version to be published.

ACKNOWLEDGMENTS

The authors are grateful to Izmir Institute of Technology for funding and the Giresun Hazelnut Research Institute for providing plant material.

DATA AVAILABILITY STATEMENT

Data discussed in this study can be found in the supporting information.

ORCID

Asena Akköse Baytar https://orcid.org/0000-0002-1068-4544
Ertuğrul Gazi Yanar https://orcid.org/0000-0002-1325-0346
Anne Frary https://orcid.org/0000-0002-8973-0100
Sami Doğanlar https://orcid.org/0000-0002-5712-0826

REFERENCES

Alasalvar, C., Shahidi, F., Liyanapathirana, C. M., & Ohshima, T. (2003). Turkish Tombul hazelnut (*Corylus avellana* L.). 1. Compositional characteristics. *Journal of Agricultural and Food Chemistry*, *51*(13), 3790–3796. https://doi.org/10.1021/jf0212385

An, N., Turp, M. T., Türkeş, M., & Kurnaz, M. L. (2020). Mid-term impact of climate change on hazelnut yield. *Agriculture*, *10*(5), 159. https://doi.org/10.3390/agriculture10050159

Arab, M. M., Marrano, A., Abdollahi-Arpanahi, R., Leslie, C. A., Cheng, H., Neale, D. B., & Vahdati, K. (2019). Combining phenotype, genotype and environment to uncover genetic components underlying water use efficiency in Persian walnut. *Journal of Experimental Botany*, *71*, 1107–1127. https://doi.org/10.1093/jxb/erz467

Beltramo, C., Valentini, N., Portis, E., Torello Marinoni, D., Boccacci, P., Sandoval Prando, M. A., & Botta, R. (2016). Genetic mapping and QTL analysis in European hazelnut (*Corylus avellana* L.). *Molecular Breeding*, *36*(3), 27. https://doi.org/10.1007/s11032-016-0450-6

Benmoussa, H., ben Mimoun, M., Ghrab, M., & Luedeling, E. (2018). Climate change threatens central Tunisian nut orchards. *International Journal of Biometeorology*, *62*(12), 2245–2255. https://doi.org/10.1007/s00484-018-1628-x

Borrell, A. K., Wong, A. C. S., George-Jaeggli, B., van Oosterom, E. J., Mace, E. S., Godwin, I. D., Liu, G., Mullet, J. E., Klein, P. E., Hammer, G. L., McLean, G., Hunt, C., & Jordan, D. R. (2022). Genetic modification of PIN genes induces causal mechanisms of stay-green drought adaptation phenotype. *Journal of Experimental Botany*, *73*(19), 6711–6726. https://doi.org/10.1093/jxb/erac336

Brown, L., Rosner, B., Willett, W. W., & Sacks, F. M. (1999). Cholesterol-lowering effects of dietary fiber: A meta-analysis. *The American Journal of Clinical Nutrition*, *69*(1), 30–42. https://doi.org/10.1093/ajcn/69.1.30

Bükücü, Ş. B., Sütyemez, M., Kefayati, S., Paizila, A., Jighly, A., & Kafkas, S. (2020). Major QTL with pleiotropic effects controlling time of leaf budburst and flowering-related traits in walnut (*Juglans regia* L.). *Scientific Reports*, *10*(1), 15207. https://doi.org/10.1038/s41598-020-71809-x

Butler, J. B., Freeman, J. S., Potts, B. M., Vaillancourt, R. E., Kahrood, H. V., Ades, P. K., Rigault, P., & Tibbits, J. F. G. (2022). Patterns of genomic diversity and linkage disequilibrium across the disjunct range of the Australian forest tree *Eucalyptus globulus*. *Tree Genetics & Genomes*, *18*(3), 28. https://doi.org/10.1007/s11295-022-01558-7

Çalışkan, T., & Çetiner, E. (1992). *Bazi findik cesit ve tiplerinde karakterizasyon calismalari*. T.C. Tarim ve Koyisleri Bakanligi, Tarimsal Arastirmalar Genel Mudurlugu, Findik Arastirma Ensitusu Mudurlugu Yayinlar No: 25.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1), 7. https://doi.org/10.1186/s13742-015-0047-8

Chauhan, S., Srivalli, S., Nautiyal, A. R., & Khanna-Chopra, R. (2009). Wheat cultivars differing in heat tolerance show a differential response to monocarpic senescence under high-temperature stress and the involvement of serine proteases. *Photosynthetica*, *47*(4), 536–547. https://doi.org/10.1007/s11099-009-0079-3

Chen, Y., Fan, X., Song, W., Zhang, Y., & Xu, G. (2012). Over-expression of OsPIN2 leads to increased tiller numbers, angle and shorter plant height through suppression of OsLAZY1. *Plant Biotechnology Journal*, *10*(2), 139–149. https://doi.org/10.1111/j.1467-7652.2011.00637.x

Chong, X., Su, J., Wang, F., Wang, H., Song, A., Guan, Z., Fang, W., Jiang, J., Chen, S., Chen, F., & Zhang, F. (2019). Identification of favorable SNP alleles and candidate genes responsible for inflorescence-related traits via GWAS in chrysanthemum. *Plant Molecular Biology*, *99*(4–5), 407–420. https://doi.org/10.1007/s11103-019-00826-w

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating

and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly*, *6*(2), 80–92. https://doi.org/10.4161/fly.19695

Diaz-Mendoza, M., Velasco-Arroyo, B., Santamaria, M. E., González-Melendi, P., Martinez, M., & Diaz, I. (2016). Plant senescence and proteolysis: Two processes with one destiny. *Genetics and Molecular Biology*, *39*(3), 329–338. https://doi.org/10.1590/1678-4685-GMB-2016-0015

Ekologos. (2020). Fındık, İklim Değişikliği ve Çevresel Etkiler Türkiye Raporu. Report prepared by Ekologos Sustainability Management and Communications on behalf of Rainforest Alliance, UTZ. Downloadable at https://www.sivilsayfalar.org/raporlar/findik-iklim-degisikligi-ve-cevresel-etkileri-turkiye-raporu/

FAO. (2023). *FAOSTAT statistical database*. FAO. http://www.fao.org/faostat/en/#data/QCL

Ferrucci, A., Lupo, M., Turco, S., Pavese, V., Marinoni, D. T., Botta, R., Cristofori, V., Mazzaglia, A., & Silvestri, C. (2023). A roadmap of tissue culture and biotechnology in European hazelnut (*Corylus avellana* L.). *Plant Physiology and Biochemistry*, *205*, 108167. https://doi.org/10.1016/j.plaphy.2023.108167

Frary, A., Öztürk, S. C., Balık, H. I., Balık, S. K., Kızılcı, G., Doğanlar, S., & Frary, A. (2019). Association mapping of agro-morphological traits in European hazelnut (*Corylus avellana*). *Euphytica*, *215*(2), 21. https://doi.org/10.1007/s10681-019-2352-2

Fulton, T. M., Chunwongse, J., & Tanksley, S. D. (1995). Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Molecular Biology Reporter*, *13*(3), 207–209. https://doi.org/10.1007/BF02670897

Gaut, B. S., & Long, A. D. (2003). The lowdown on linkage disequilibrium. *The Plant Cell*, *15*(7), 1502–1506. https://doi.org/10.1105/tpc.150730

Gauthier, M.-M., & Jacobs, D. F. (2011). Walnut (*Juglans* spp.) ecophysiology in response to environmental stresses and potential acclimation to climate change. *Annals of Forest Science*, *68*(8), 1277–1290. https://doi.org/10.1007/s13595-011-0135-6

Götz, S., Garcia-Gomez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., Robles, M., Talon, M., Dopazo, J., & Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*, *36*(10), 3420–3435. https://doi.org/10.1093/nar/gkn176

Gregersen, P. L., Holm, P. B., & Krupinska, K. (2008). Leaf senescence and nutrient remobilisation in barley and wheat. *Plant Biology*, *10*, 37–49. https://doi.org/10.1111/j.1438-8677.2008.00114.x

Hill, R. J., Baldassi, C., Snelling, J. W., Vining, K. J., & Mehlenbacher, S. A. (2021). Fine mapping of the locus controlling self-incompatibility in European hazelnut. *Tree Genetics and Genomes*, *17*(1), 6. https://doi.org/10.1007/s11295-020-01485-5

Honig, J. A., Muehlbauer, M. F., Capik, J. M., Kubik, C., Vaiciunas, J. N., Mehlenbacher, S. A., & Molnar, T. J. (2019). Identification and mapping of Eastern Filbert Blight resistance quantitative trait loci in European hazelnut using double digestion restriction site associated DNA sequencing. *Journal of the American Society for Horticultural Science*, *144*(5), 295–304. https://doi.org/10.21273/JASHS04694-19

Hsieh, M.-H., & Goodman, H. M. (2002). Molecular characterization of a novel gene family encoding ACT domain repeat proteins in *Arabidopsis*. *Plant Physiology*, *130*(4), 1797–1806. https://doi.org/10.1104/pp.007484

Jiang, J., Yao, L., Yu, Y., Liang, Y., Jiang, J., Ye, N., Miao, Y., & Cao, J. (2014). PECTATE LYASE-LIKE 9 from *Brassica campestris* is associated with intine formation. *Plant Science*, *229*, 66–75. https://doi.org/10.1016/j.plantsci.2014.08.008

Jiang, J., Yao, L., Yu, Y., Lv, M., Miao, Y., & Cao, J. (2014). PECTATE LYASE-LIKE10 is associated with pollen wall development in *Brassica campestris*. *Journal of Integrative Plant Biology*, *56*(11), 1095–1105. https://doi.org/10.1111/jipb.12209

Kritchevsky, D. (1986). Dietary fiber and atherosclerosis. In G. V. Vahouny & D. Kritchevsky (Eds.), *Dietary fiber* (pp. 265–274). Springer. https://doi.org/10.1007/978-1-4613-2111-8_17

Krupinska, K., Mulisch, M., Hollmann, J., Tokarz, K., Zschiesche, W., Kage, H., Humbeck, K., & Bilger, W. (2012). An alternative strategy of dismantling of the chloroplasts during leaf senescence observed in a high-yield variety of barley. *Physiologia Plantarum*, *144*(2), 189–200. https://doi.org/10.1111/j.1399-3054.2011.01545.x

Li, H. (2014). Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics*, *30*(20), 2843–2851. https://doi.org/10.1093/bioinformatics/btu356

Li, C., Dong, Y., Zhao, T., Li, L., Li, C., Yu, E., Mei, L., Daud, M. K., He, Q., Chen, J., & Zhu, S. (2016). Genome-wide SNP linkage mapping and QTL analysis for fiber quality and yield traits in the upland cotton recombinant inbred lines population. *Frontiers in Plant Science*, *7*, 1356. https://doi.org/10.3389/fpls.2016.01356

Liu, Q. (2006). Computational identification and systematic analysis of the ACR gene family in *Oryza sativa*. *Journal of Plant Physiology*, *163*(4), 445–451. https://doi.org/10.1016/j.jplph.2005.05.011

Lombardoni, J. J., Honig, J. A., Revord, R. S., & Molnar, T. J. (2022). Segregation of eastern filbert blight disease response and single nucleotide polymorphism markers in three european–american interspecific hybrid hazelnut populations. *Journal of the American Society for Horticultural Science*, *147*(4), 196–207. https://doi.org/10.21273/JASHS05112-22

Lu, M., Krutovsky, K. V., Nelson, C. D., Koralewski, T. E., Byram, T. D., & Loopstra, C. A. (2016). Exome genotyping, linkage disequilibrium and population structure in loblolly pine (*Pinus taeda* L.). *BMC Genomics*, *17*(1), 730. https://doi.org/10.1186/s12864-016-3081-8

Lucas, S. J., Kahraman, K., Avşar, B., Buggs, R. J. A., & Bilge, I. (2021). A chromosome-scale genome assembly of European hazel (*Corylus avellana* L.) reveals targets for crop improvement. *The Plant Journal*, *105*(5), 1413–1430. https://doi.org/10.1111/tpj.15099

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M., Hurwitz, D. I., Lanczycki, C. J., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., & Bryant, S. H. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Research*, *43*(D1), D222–D226. https://doi.org/10.1093/nar/gku1221

Meyer, R. C., Steinfath, M., Lisec, J., Becher, M., Witucka-Wall, H., Törjék, O., Fiehn, O., Eckardt, Ä., Willmitzer, L., Selbig, J., & Altmann, T. (2007). The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences*, *104*(11), 4759–4764. https://doi.org/10.1073/pnas.0609709104

Miedes, E., Herbers, K., Sonnewald, U., & Lorences, E. P. (2010). Overexpression of a cell wall enzyme reduces xyloglucan depolymerization and softening of transgenic tomato fruits. *Journal of Agricultural and Food Chemistry*, *58*(9), 5708–5713. https://doi.org/10.1021/jf100242z

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D., & Bateman, A. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Research*, *49*(D1), D412–D419. https://doi.org/10.1093/nar/gkaa913

Monika, M., & Anna, K.-D. (2019). Nut oils and their dietetic and cosmetic significance: A review. *Journal of Oleo Science*, *68*(2), 111–120. https://doi.org/10.5650/jos.ess18216

Nishikubo, N., Awano, T., Banasiak, A., Bourquin, V., Ibatullin, F., Funada, R., Brumer, H., Teeri, T. T., Hayashi, T., Sundberg, B., & Mellerowicz, E. J. (2007). Xyloglucan endo-transglycosylase (xet) functions in gelatinous layers of tension wood fibers in poplar—A glimpse into the mechanism of the balancing act of trees. *Plant and*

Cell Physiology, *48*(6), 843–855. https://doi.org/10.1093/pcp/pcm055

Nishikubo, N., Takahashi, J., Roos, A. A., Derba-Maceluch, M., Piens, K., Brumer, H., Teeri, T. T., Stålbrand, H., & Mellerowicz, E. J. (2011). Xyloglucan endotransglycosylase-mediated xyloglucan rearrangements in developing wood of hybrid aspen. *Plant Physiology*, *155*(1), 399–413. https://doi.org/10.1104/pp.110.166934

Nishitani, K. (1997). The role of endoxyloglucan transferase in the organization of plant cell walls. *International Review of Cytology*, *173*, 157. https://doi.org/10.1016/S0074-7696(08)62477-8

O'Leary, S. J., Puritz, J. B., Willis, S. C., Hollenbeck, C. M., & Portnoy, D. S. (2018). These aren't the loci you're looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology*, *27*(16), 3193–3206. https://doi.org/10.1111/mec.14792

Orozco-Arias, S., Isaza, G., & Guyot, R. (2019). Retrotransposons in plant genomes: Structure, identification, and classification through bioinformatics and machine learning. *International Journal of Molecular Sciences*, *20*(15), 3837. https://doi.org/10.3390/ijms20153837

Özdemir, F., Topuz, A., Doğan, Ü., & Karkacier, M. (1998). Some physical and chemical properties of hazelnut cultivars. *GIDA*, *23*(1), 37–41.

Öztürk, S. C., Öztürk, S. E., Celik, I., Stampar, F., Veberic, R., Doğanlar, S., Solar, A., & Frary, A. (2017). Molecular genetic diversity and association mapping of nut and kernel traits in Slovenian hazelnut (*Corylus avellana*) germplasm. *Tree Genetics & Genomes*, *13*(1), 16. https://doi.org/10.1007/s11295-016-1098-4

Parker, L. E., & Abatzoglou, J. T. (2018). Shifts in the thermal niche of almond under climate change. *Climatic Change*, *147*(1–2), 211–224. https://doi.org/10.1007/s10584-017-2118-6

Paul, M. J., Primavesi, L. F., Jhurreea, D., & Zhang, Y. (2008). Trehalose metabolism and signaling. *Annual Review of Plant Biology*, *59*(1), 417–441. https://doi.org/10.1146/annurev.arplant.59.032607.092945

Pellny, T. K., Ghannoum, O., Conroy, J. P., Schluepmann, H., Smeekens, S., Andralojc, J., Krause, K. P., Goddijn, O., & Paul, M. J. (2004). Genetic modification of photosynthesis with *E. coli* genes for trehalose synthesis. *Plant Biotechnology Journal*, *2*(1), 71–82. https://doi.org/10.1111/j.1467-7652.2004.00053.x

Ponnu, J., Wahl, V., & Schmid, M. (2011). Trehalose-6-phosphate: Connecting plant metabolism and development. *Frontiers in Plant Science*, *2*, 70. https://doi.org/10.3389/fpls.2011.00070

R Core Team. (2021). R: A language and environment for statistical computing. https://www.r-project.org/

Roberts, I. N., Caputo, C., Criado, M. V., & Funk, C. (2012). Senescence-associated proteases in plants. *Physiologia Plantarum*, *145*(1), 130–139. https://doi.org/10.1111/j.1399-3054.2012.01574.x

Roberts, I. N., Murray, P. F., Caputo, C. P., Passeron, S., & Barneix, A. J. (2003). Purification and characterization of a subtilisin-like serine protease induced during the senescence of wheat leaves. *Physiologia Plantarum*, *118*(4), 483–490. https://doi.org/10.1034/j.1399-3054.2003.00142.x

Rowley, E. R., VanBuren, R., Bryant, D. W., Priest, H. D., Mehlenbacher, S. A., & Mockler, T. C. (2018). A draft genome and high-density genetic map of European hazelnut (*Corylus avellana* L.). *BioRxiv*, *1*, 1–25. https://doi.org/10.1101/469015

Said, J. I., Lin, Z., Zhang, X., Song, M., & Zhang, J. (2013). A comprehensive meta QTL analysis for fiber quality, yield, yield related and morphological traits, drought tolerance, and disease resistance in tetraploid cotton. *BMC Genomics*, *14*(1), 776. https://doi.org/10.1186/1471-2164-14-776

Sayar, N. A., Pınar, O., Kazan, D., & Sayar, A. A. (2019). Bioethanol production from Turkish hazelnut husk process design and economic evaluation. *Waste and Biomass Valorization*, *10*(4), 909–923. https://doi.org/10.1007/s12649-017-0103-y

Silvestri, C., Bacchetta, L., Bellincontro, A., & Cristofori, V. (2021). Advances in cultivar choice, hazelnut orchard management, and nut storage to enhance product quality and safety: An overview. *Journal of the Science of Food and Agriculture*, *101*(1), 27–43. https://doi.org/10.1002/jsfa.10557

Soto-Cerda, B. J., Diederichsen, A., Ragupathy, R., & Cloutier, S. (2013). Genetic characterization of a core collection of flax (*Linum usitatissimum* L.) suitable for association mapping studies and evidence of divergent selection between fiber and linseed types. *BMC Plant Biology*, *13*(1), 78. https://doi.org/10.1186/1471-2229-13-78

Söylemez, F. (2020). *Association mapping for oil and protein content in hazelnut (Corylus avellana L.)*. Izmir Institute of Technology.

Taghavi, T., Dale, A., Saxena, P., Galic, D., Rahemi, A., Kelly, J., & Suarez, E. (2018). Flowering of hazelnut cultivars and how it relates to temperature in southern Ontario. *Acta Horticulturae*, *1226*, 131–136. https://doi.org/10.17660/ActaHortic.2018.1226.18

Talbot, S. C., Vining, K. J., Snelling, J. W., Clevenger, J., & Mehlenbacher, S. A. (2024). A haplotype-resolved chromosome-level assembly and annotation of European hazelnut (*C. avellana* cv. Jefferson) provides insight into mechanisms of eastern filbert blight resistance. *G3 Genes|Genomes|Genetics*, *14*(6), jkae021. https://doi.org/10.1093/g3journal/jkae021

Tariq, N., Jenkins, D. J., Vidgen, E., Fleshner, N., Kendall, C. W., Story, J. A., Singer, W., D'Costa, M., & Struthers, N. (2000). Effect of soluble and insoluble fiber diets on serum prostate specific antigen in men. *The Journal of Urology*, *163*(1), 114–118. https://doi.org/10.1016/S0022-5347(05)67985-2

TinHan, T. C., O'Leary, S. J., Portnoy, D. S., Rooker, J. R., Gelpi, C. G., & Wells, R. J. D. (2020). Natural tags identify nursery origin of a coastal elasmobranch *Carcharhinus leucas*. *Journal of Applied Ecology*, *57*(7), 1222–1232. https://doi.org/10.1111/1365-2664.13627

Torello Marinoni, D., Valentini, N., Portis, E., Acquadro, A., Beltramo, C., & Botta, R. (2018a). Construction of a high-density genetic linkage map and QTL analysis for hazelnut breeding. *Acta Horticulturae*, *1226*, 25–30. https://doi.org/10.17660/ActaHortic.2018.1226.3

Torello Marinoni, D., Valentini, N., Portis, E., Acquadro, A., Beltramo, C., Mehlenbacher, S. A., Mockler, T. C., Rowley, E. R., & Botta, R. (2018b). High density SNP mapping and QTL analysis for time of leaf budburst in *Corylus avellana* L. *PLoS ONE*, *13*(4), e0195408. https://doi.org/10.1371/journal.pone.0195408

UPOV. (1979). Hazelnut (*Corylus avellana* L. and *Corylus maxima* mill.): Guidelines for the conduct of tests for distinctness, uniformity and stability. Hazelnut/Noisetier/Haselnuss, 79-03-28. Doc. No. TG/71/3. UPOV.

Valentini, N., Portis, E., Botta, R., Acquadro, A., Pavese, V., Cavalet Giorsa, E., & Torello Marinoni, D. (2021). Mapping the genetic regions responsible for key phenology-related traits in the European hazelnut. *Frontiers in Plant Science*, *12*, 749394. https://doi.org/10.3389/fpls.2021.749394

van Sandt, V. S. T., Suslov, D., Verbelen, J.-P., & Vissenberg, K. (2007). Xyloglucan endotransglucosylase activity loosens a plant cell wall. *Annals of Botany*, *100*(7), 1467–1473. https://doi.org/10.1093/aob/mcm248

Vissenberg, K., Martinez-Vilchez, I. M., Verbelen, J.-P., Miller, J. G., & Fry, S. C. (2000). In vivo colocalization of xyloglucan endotransglycosylase activity and its donor substrate in the elongation zone of *Arabidopsis* roots. *The Plant Cell*, *12*(7), 1229–1237. https://doi.org/10.1105/tpc.12.7.1229

Willis, S. C., Hollenbeck, C. M., Puritz, J. B., Gold, J. R., & Portnoy, D. S. (2017). Haplotyping RAD loci: An efficient method to filter paralogs and account for physical linkage. *Molecular Ecology Resources*, *17*(5), 955–965. https://doi.org/10.1111/1755-0998.12647

Wing, R. A., Yamaguchi, J., Larabell, S. K., Ursin, V. M., & McCormick, S. (1990). Molecular and genetic characterization of two pollen-expressed genes that have sequence similarity to pectate lyases of the plant pathogen *Erwinia*. *Plant Molecular Biology*, *14*(1), 17–28. https://doi.org/10.1007/BF00015651

Xu, W., Campbell, P., Vargheese, A. K., & Braam, J. (1996). The Arabidopsis XET-related gene family: Environmental and hormonal regulation of expression. *The Plant Journal*, *9*(6), 879–889. https://doi.org/10.1046/j.1365-313X.1996.9060879.x

Xu, M., Zhu, L., Shou, H., & Wu, P. (2005). A PIN1 family gene, OsPIN1, involved in auxin-dependent adventitious root emergence and tillering in rice. *Plant and Cell Physiology*, *46*(10), 1674–1681. https://doi.org/10.1093/pcp/pci183

Yanar, E. G., Frary, A., & Doğanlar, S. (2023). GRAS-Di SNP-based molecular characterization and fingerprinting of a Turkish *Corylus avellana* core set provide insights into the cultivation and breeding of hazelnut in Turkey. *Molecular Genetics and Genomics*, *298*(2), 413–426. https://doi.org/10.1007/s00438-022-01990-3

Yang, Y., Ma, T., Wang, Z., Lu, Z., Li, Y., Fu, C., Chen, X., Zhao, M., Olson, M. S., & Liu, J. (2018). Genomic effects of population collapse in a critically endangered ironwood tree *Ostrya rehderiana*. *Nature Communications*, *9*(1), 5449. https://doi.org/10.1038/s41467-018-07913-4

Yang, Z., Ma, W., Wang, L., Yang, X., Zhao, T., Liang, L., Wang, G., & Ma, Q. (2023). Population genomics reveals demographic history and selection signatures of hazelnut (Corylus). *Horticulture Research*, *10*(5), uhad065. https://doi.org/10.1093/hr/uhad065

Yin, L., Zhang, H., Tang, Z., Xu, J., Yin, D., Zhang, Z., Yuan, X., Zhu, M., Zhao, S., Li, X., & Liu, X. (2021). rMVP: A memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. *Genomics, Proteomics & Bioinformatics*, *19*(4), 619–628. https://doi.org/10.1016/j.gpb.2020.10.007

Young, A. I., Wauthier, F. L., & Donnelly, P. (2018). Identifying loci affecting trait variability and detecting interactions in genome-wide association studies. *Nature Genetics*, *50*(11), 1608–1614. https://doi.org/10.1038/s41588-018-0225-6

Zhang, Y., Primavesi, L. F., Jhurreea, D., Andralojc, P. J., Mitchell, R. A. C., Powers, S. J., Schluepmann, H., Delatte, T., Wingler, A., & Paul, M. J. (2009). Inhibition of SNF1-related protein kinase1 activity and regulation of metabolic pathways by trehalose-6-phosphate. *Plant Physiology*, *149*(4), 1860–1871. https://doi.org/10.1104/pp.108.133934

Zheng, Y., Yan, J., Wang, S., Xu, M., Huang, K., Chen, G., & Ding, Y. (2018). Genome-wide identification of the pectate lyase-like (PLL) gene family and functional analysis of two PLL genes in rice. *Molecular Genetics and Genomics*, *293*(6), 1317–1331. https://doi.org/10.1007/s00438-018-1466-x

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Baytar, A. A., Yanar, E. G., Frary, A., & Doğanlar, S. (2024). Association mapping and candidate gene identification for yield traits in European hazelnut (*Corylus avellana* L.). *Plant Direct*, *8*(8), e625. https://doi.org/10.1002/pld3.625