Tree Genetics & Genomes (2017) 13: 16
DOI 10.1007/s11295-016-1098-4

ORIGINAL ARTICLE

Molecular genetic diversity and association mapping of nut and kernel traits in Slovenian hazelnut (*Corylus avellana*) germplasm

Süleyman Can Ozturk[1] · Saniye Elvan Ozturk[1] · Ibrahim Celik[1] · Franci Stampar[2] · Robert Veberic[2] · Sami Doganlar[1] · Anita Solar[2] · Anne Frary[1]

Received: 16 November 2016 / Revised: 20 December 2016 / Accepted: 25 December 2016 / Published online: 12 January 2017

Abstract European hazelnut (*Corylus avellana* L.), cultivated in several areas of the world including Europe, Anatolia, and the USA, is an economically important nut crop due to its high mineral, oleic acid, amino acid, and phenolic compound content and pleasant flavor. This study examined molecular genetic diversity and population structure of 54 wild accessions and 48 cultivars from the Slovenian national hazelnut collection using amplified fragment length polymorphism (AFLP) and simple sequence repeat (SSR) markers. Eleven AFLP primer combinations and 49 SSR markers yielded 532 and 504 polymorphic fragments, respectively. As expected for a wind-pollinated, self-incompatible species, levels of genetic diversity were high with cultivars and wild accessions having mean dissimilarity values of 0.50 and 0.60, respectively. In general, cultivars and wild accessions clustered separately in dendrogram, principal coordinate, and population structure analyses with regional clustering of the wild material. The accessions were also characterized for ten nut and seven kernel traits and some wild accessions were shown to have breeding potential. Morphological principal component analysis showed distinct clustering of cultivars and wild accessions. An association mapping panel composed of 64 hazelnut cultivars and wild accessions had considerable variation for the nut and kernel quality traits. Morphological and molecular data were associated to identify markers controlling the traits. In all, 49 SSR markers were significantly associated with nut and kernel traits [$P < 0.0001$ and LD value (r^2) = 0.15–0.50]. This work is the first use of association mapping in hazelnut and has identified molecular markers associated with important quality parameters in this important nut crop.

Keywords AFLP · *Corylus avellana* · Filbert · Morphological characterization · QTL · SSR

Introduction

Corylus avellana L. (European hazelnut) is a diploid (2n = 22), monoecious, dichogamous, and wind-pollinated species belonging to the Betulaceae family. This species is the source of the commercially important hazelnut cultivars grown in Europe, Anatolia, and the USA (Boccacci et al. 2006). Turkey and Italy are major hazelnut producers with 77% of world production with the remainder grown by countries such as the USA, Georgia, Azerbaijan, Spain, and Slovenia (FAO 2013). Although Slovenia is a minor hazelnut producing country with less than 1% of the world total, the country has extensive hazelnut genetic resources including wild accessions and cultivars such as 'Istrska dolgoplodna leska' which originated in Croatia but was domesticated in Slovenia. Orchards with mean surface of 2.2 ha are mainly located in the Stajerska, Dolenjska, and Celjska kotlinam regions, producing approximately 200 tons of in-shell nuts per year (Slatnar et al. 2014). On the basis, of a long-term investigation, some international cultivars, such as 'Tonda di Giffoni' and 'Daria' from Italy, 'Ennis' from the USA and 'Pauetet' from Spain are

SC. Ozturk and SE Ozturk contributed equally to this paper

Communicated by M. Troggio

Electronic supplementary material The online version of this article (doi:10.1007/s11295-016-1098-4) contains supplementary material, which is available to authorized users.

✉ Anne Frary
annefrary@iyte.edu.tr

[1] Department of Molecular Biology and Genetics, Izmir Institute of Technology, Urla, 35430 Izmir, Turkey

[2] Biotechnical Faculty, Department of Agronomy, University of Ljubjana, Ljubjana, Slovenia

recommended for commercial growth in contemporary orchards in Slovenia (Solar and Stampar, 2009, 2011). In addition, many local wild populations are distributed throughout the country, representing diverse hazelnut genetic resources. These populations are interesting for further selection and breeding but must first be genetically characterized.

The genetic diversity of hazelnut cultivars was first assessed using randomly amplified polymorphic DNA (RAPD) (Galderisi et al. 1999; Miaja et al. 2001; Kafkas et al. 2009) and amplified fragment length polymorphism (AFLP) (Ferrari et al. 2004; Kafkas et al. 2009) markers. AFLP was used to fingerprint 57 clones (Ferrari et al. 2004) and, in combination with other markers, to assess diversity in 18 Turkish hazelnut cultivars (Kafkas et al. 2009). In other work, Martins et al. (2014) used AFLP and inter simple sequence repeat (ISSR) markers to measure diversity in wild and cultivated hazelnuts from Portugal.

Simple sequence repeat (SSR) markers have also been used to assess genetic diversity in hazelnut and are more popular than other marker types because they are simpler to assay, highly polymorphic, and reproducible. To date, more than 300 genomic SSR markers have been developed and mapped in the hazelnut genome (Boccacci et al. 2005; Bassil et al. 2005, 2013; Mehlenbacher et al. 2006; Gurcan and Mehlenbacher 2010; Gurcan et al. 2010a). Most recently, Betulaceae expressed sequence tag (EST) were used to identify more than 1000 genic SSRs, 36 of which were found to be useful in hazelnut (Boccacci et al. 2015). Much of the research assessing genetic diversity in hazelnut has been done under the auspices of the SAFENUT European Commission Action which focused on characterization, conservation and use of European hazelnut germplasm (reviewed in Bacchetta et al. 2014). As part of this project, analyses of SSR loci in hazelnut revealed high levels of genetic diversity in accessions from Spain (Boccacci et al. 2005, 2006, 2008) and Southern Europe (Boccacci et al. 2013). In other work, Black Sea region hazelnuts from Turkey, Georgia, and Azerbaijan were also found to be highly diverse (Gurcan et al. 2010b).

Most genetic diversity studies in hazelnut have focused on cultivars (Bacchetta et al. 2014) with only recent interest in wild individuals and landraces (Ferreira et al. 2010; Campa et al. 2011; Leinemann et al. 2013; Martins et al. 2013, 2014, 2015). In addition to its contribution to biodiversity, wild germplasm is widely recognized as a potential resource of interesting traits for improved cultivars (Tanskley and McCouch, 1997). Thus, the molecular and morphological variation of Slovenian hazelnut genetic resources should be examined for valuable features. In addition, these resources can be used to reveal the molecular bases of agronomic traits by quantitative trait locus (QTL) mapping approaches such as association mapping.

Association mapping (AM), also called linkage disequilibrium (LD) mapping, was first developed for QTL identification in medical genomics studies and is now frequently implemented in plant genomics studies. Association mapping is more practical than QTL mapping performed in bi-parental mapping population because it does not require the development of experimental populations such as F2 and BC (backcross). The development of such populations is time-consuming especially in tree species like hazelnut (Gómez et al. 2011). Instead, AM uses an association panel consisting of naturally occurring plant germplasm/populations. AM also has higher resolution than bi-parental QTL mapping because AM uses LD generated by historical recombination and can detect more alleles than are found in bi-parental populations (Zhu et al. 2008; Gómez et al. 2011; Khan and Korban 2012). In a recent study, 275 'Tonda Gentile delle Langhe' X 'Merveille de Bollwiller' hazelnut F1 hybrids were used for QTL analysis, which was performed for vigor, sucker habit, and time of bud burst characters (Beltramo et al. 2016). However, to date no association mapping has been performed with hazelnut.

Nut and kernel traits are important yield and quality parameters for hazelnut. Although these traits have been characterized for a limited number of reference and local cultivars (Ozdemir and Akinci 2004; Balta et al. 2006; Cristofori et al. 2008; Ferreira et al. 2010; Xu and Hanna 2010; Solar and Stampar 2011), to our knowledge, wild hazelnut accessions have not been examined in this way. Morphological and molecular characterization of wild accessions for nut and kernel traits is important to assess their breeding potential. In addition, identification of molecular markers linked to QTLs for quality traits is essential for the implementation of marker-assisted selection in hazelnut for targeted breeding of nut and kernel traits.

The purpose of this study was to analyze the genetic diversity and population structure of 102 wild and cultivated hazelnut accessions grown in Slovenia. The clonal accessions included 54 wild accessions collected in five regions in Slovenia and 48 cultivars originating from Europe and the USA. These accessions were evaluated with molecular marker data from 11 AFLP primer combinations and 49 SSR markers. The germplasm was also evaluated for nut and kernel traits and these data were used to identify QTLs for these parameters via association mapping. Thus, this study is the first AM QTL report for nut and kernel quality traits in hazelnut. The results of this work will be useful for breeding of hazelnut using molecular approaches.

Materials and methods

Plant sampling and DNA isolation

For genetic diversity analysis, 48 individuals of *C. avellana* were sampled from the national hazelnut collection in

Maribor, NE, Slovenia. These accessions represent cultivars that have been introduced into Slovenia. Leaves and catkins were taken from one single, true-to-type plant of the three replicates that were planted per cultivar. An additional 54 samples were obtained by in situ collection of wild accessions from five hazelnut growing regions in Slovenia (Suppl. Fig. 1, Suppl. Table 1). The Koroska region is characterized by a humid continental climate (Dwb) and is one of the coldest areas in Slovenia beside the Alps (Suppl. Table 2). Maribor and Dolenjska, two regions with extensive vineyard production, have temperate climate with dry winters (Cwb). The Vipava-Razdrto and Bovec regions have a similar climate but without a dry season (Cfwb) and are areas where a Mediterranean influence can be felt. Total genomic DNA was isolated from leaf or catkin tissue according to a microprep method (Fulton et al. 1995).

A panel composed of 24 cultivars and 40 wild accessions was randomly chosen from the germplasm described above for morphological characterization and association mapping of nut and kernel traits.

AFLP amplification

AFLP Core Reagent and AFLP Starter Primer Kits from Invitrogen (Carlsbad, CA, USA) were used according to the manufacturer's protocol (Vos et al. 1995). Sixty-four selective *EcoRI/MseI* primer combinations were tested on 'Willamette' (accession S58) and the wild accession B9. Based on these results, 11 combinations (*M-CAC* + *E-AGC, M-CAA* + *E-ACG, M-CAA* + *E-ACC, M-CAG* + *E-ACT, M-CTC* + *E-AGG, M-CTC* + *E-ACA, M-CAT* + *E-ACA, M-CAT* + *E-ACT, M-CTA* + *E-ACA, M-CTG* + *E-AGC* and *M-CTT* + *E-AGG*) were chosen as the most polymorphic and subsequently applied to the 102 hazelnut accessions (Suppl. Table 3). After selective PCR, fragments with labeled EcoRI primer signals were detected using a Genetic Analysis System CEQ 8800 machine (Beckman-Coulter, Fullerton, CA, USA). Amplification products were diluted 1:10 in sample loading solution (SLS) with 0.5 μl size standard 600. The mixture for each accession was then run on a Beckman CEQ8800 capillary electrophoresis device using the frag2 method (capillary temperature 35 °C, denaturation 90 °C for 120 s, injection voltage 2.0 kV for 30 s, separation voltage 6.0 kV for 60 min). PCR fragments were scored binomially (presence 1, absence 0).

SSR amplification

A total of 49 SSR marker pairs was used to accession the 102 hazelnut accessions. SSR markers were selected based on their polymorphic allele content as reported by Bassil et al. (2005), Boccacci et al. (2005), and Gurcan et al. (2010a) (Suppl. Table 4). PCR amplification was performed with 20 ng DNA in a 20-μl reaction containing 10 pmol of each primer pair, 200 μm dNTPs, 2 μl 10× Taq polymerase buffer and 0.6 Unit Taq polymerase. The same reaction conditions were used for all primers: 30 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s and extension at 72 °C for 30 s). These cycles were preceded by a denaturing step at 94 °C for 3 min and ended with an extension step at 72 °C for 5 min. The PCR amplifications were performed in a GeneAmp PCR system 9700 (Perkin Elmer Applied Biosystems). After amplification, samples were separated by capillary electrophoresis using a Fragment Analyzer™ (Applied Biosystems) with the DNF-900 dsDNA Reagent Kit (Advanced Analytical) according to the manufacturer's instructions. PCR fragments were scored binomially (presence 1, absence 0) because many of the SSR markers yielded more than two fragments and allelism could not be determined.

Data analysis

Average gene diversity (Nei 1973) was calculated for each AFLP primer combination and SSR marker with the formula: average gene diversity$_i$ $=\left(\sum_{i=1}^{n} 2fi(1-fi)\right)/n$ (Roldan-Ruiz et al. 2000), where *fi* is the frequency of band presence for the *i*th allele and *n* is the number of alleles. Calculated in this way, the diversity value of a locus ranges from 0 (monomorphic) to 0.5 (highly informative). Cluster analysis was performed using the Dice coefficient (Dice, 1945) and unweighted neighbor joining (NJ) algorithm in DARwin 5 software (Perrier and Jacquemoud-Collet 2006). DARwin 5 was also used for principal coordinate analysis (PCoA). Population structure was determined using the computer program Structure (Pritchard et al. 2000). Ad hoc statistics were used to find the best reflected subpopulation number for the hazelnut accessions (Evanno et al. 2005). For this analysis, the data were evaluated for 2 to 20 subpopulations (K = 2 to 20) with a burn-in time of 10,000 cycles. Each model was tested 10 times with 300,000 iterations per K. The probability change of each group (ΔK) was calculated using the program Structure Harvester (Earl and Von Holdt 2012). The highest ΔK was determined to be the best fit. Clusters were determined according to a threshold of ≥0.70 inferred ancestry. Accessions that did not meet this threshold were considered as admixed. A second population structure computer program, InStruct (Gao et al. 2007), was used to confirm the results of Structure and to test K = 1. As the conclusion did not differ, these results are not shown in the paper.

Morphological evaluation

The hazelnut association panel was characterized over two consecutive years for 17 nut and kernel traits using 30 samples per accession. The 10 nut traits including length, width,

thickness, shape index, caliber, shell thickness, weight, shape uniformity, and proportions of healthy and empty nuts. Nut length, width and thickness of randomly selected in-shell nuts from each hazelnut accession (cultivars and wild accessions) were measured using calipers in millimeters (mm). The largest value among the three dimensions (nut length, width, and thickness) was recorded as caliber. Shape index was calculated according to the following formula: width + thickness/(2× length). Shell thickness was determined on hand-cracked nuts using calipers to measure the convex side of each half of the shell. Nut weight was recorded in grams (g). Nut shape uniformity was visually determined for each hazelnut accession using a scale from 1 to 9 (1 = least uniform, 9 = most uniform). Proportion of healthy and empty nuts were calculated by cracking 30 nuts for each hazelnut accession.

In addition to nut traits, the hazelnut association panel was characterized for seven kernel traits including weight, kernel percentage, shape uniformity, and proportions of kernels with brown spots, mold, deformation, and twin kernels. Kernel weight was recorded in grams (g). Kernel percentage was calculated as: (kernel weight/nut weight) × 100. Shape uniformity was visually determined for each hazelnut accession using a scale from 1 to 9 (1 = least uniform, 9 = most uniform). Proportion of kernels with brown spots, mold, deformation, and twin kernels were recorded using 1 kg nut samples harvested for each hazelnut accession

Means and coefficients of variation for hazelnut cultivars and wild accessions from each region of Slovenia were calculated separately for comparison. Principal component analysis (PCA) was performed with DARwin (Perrier and Jacquemoud-Collet 2006) and PASW software (Norusis 2010). Basic statistics such as correlation analysis between traits, paired sample Student's *t* tests, ANOVA, and discriminant analyses were performed using PASW software. Stepwise discriminant analysis of the nut and kernel traits was done using subpopulation, dendrogram cluster, and region as grouping variables.

Association mapping

The binary data generated for the SSR markers assayed on the association panel were associated to the nut and kernel trait data using the GLM and MLM models of TASSEL v2.1 (Trait Analysis by aSSociation, Evolution and Linkage software) (Bradbury et al. 2007). Linkage disequilibrium (LD) values (r^2 and *P* values) between SSR markers were calculated using the same software. Several association mapping (AM) models were tested to identify the one with the best fit for AM of nut and kernel traits. Tested models were GLM model without correction; GLM model corrected with the Q-matrix of population structure (subgroup number = 2) [GLM (Q)], principal components (PC) [GLM (PC)] and both Q and PC [GLM (Q + PC)]; MLM model corrected with kindship matrix (*K*) [MLM (*K*)], Q-matrix [GLM (Q)], principal components (PC) [GLM (PC)], and both Q and PC [GLM (Q + PC)] (Suppl. Table 6). Principal components (PC) were calculated in TASSEL software. The *P* values of the eight models were analyzed with QVALUE (Storey 2002) software using a false discovery rate (FDR) of 0.05 (Storey and Tibshirani 2003). The model with the highest probability of significant results ($\pi 1$) was accepted as the one with the best fit and only those results are reported here. The percent value of $\pi 1$ was calculated based on the probability that a given hypothesis is null, $\pi 0$, such that $\pi 1$ (%) = [100 − $\pi 0$ (%)]. Marker-trait associations with *P* values lower than 0.0001 [−Log (*P* value) = 4] were selected as significant associations.

Results

The 102 hazelnut accessions studied in this work included 48 cultivars from the national collection in Maribor, Slovenia (Suppl. Table 1). These cultivars were introduced into Slovenia from other countries including Italy (12 accessions), the USA (11 accessions), France (5 accessions), Spain, the UK, and Germany (4 accessions each) with one or two cultivars of Croatian, Hungarian, Romanian, and unknown origin. Some of the cultivars have been adopted by Slovenian hazelnut growers while others are being studied for their adaptation to growth and climate conditions in the country (Solar and Stampar 2009, 2011). The remaining 54 accessions represented wild hazelnut resources and were collected in situ from five hazelnut growing regions: Vipava-Razdrto (12 accessions), Dolenjska, Koroska and Bovec (11 accessions each), and Maribor (9 accessions) (Suppl. Fig. 1). Thus, the material represented both naturally occurring and introduced genetic diversity.

AFLP and SSR marker polymorphism

A total of 532 polymorphic fragments was scored from the 11 selective AFLP primer combinations, with 27 to 69 polymorphic alleles per combination (Suppl. Table 3). Thus, AFLP provided an average of 48.4 alleles per primer combination. Average gene diversity values, which indicate the informativeness of each combination were calculated and ranged from 0.20 (for *M-CAC* + *E-AGC*) to 0.30 (for *M-CTA* + *E-ACA*) with an average of 0.26.

The 49 SSR primer pairs yielded a total of 504 polymorphic fragments in the 102 accessions with an average of 10.3 alleles per SSR marker. Number of polymorphic fragments ranged from four to 28 with SSRs B625 and B777 each yielding more than 25 fragments (Suppl. Table 4). Average gene diversity values for the SSRs ranged from 0.20 to 0.45 with B790 and A602 as the most polymorphic markers. Average gene diversity for all 49 markers was 0.30.

Genetic diversity

The AFLP and SSR data were used to construct separate distance matrices and dendrograms using the Dice coefficient and unweighted neighbor-joining algorithm. Mantel tests showed very high correlations between the dendrograms and distance matrices ($r = 0.96$ for both data sets). The distance matrices for the AFLP and SSR data were also tested for correlation using a Mantel test which indicated a very low correlation ($r = 0.33$). For that reason, the two data sets were not combined.

With the AFLP data, the hazelnut accessions grouped into two main clusters: cluster A with 46 accessions and cluster B with 54 (data not shown). A third cluster contained only two accessions. The minimum and maximum genetic dissimilarities between hazelnut accessions were 0.06 and 0.52, respectively, with a mean value of 0.32. All but seven of the cultivars (85%) fell in cluster A while all but five of the wild accessions (91%) fell in cluster B which also contained five cultivars. The remaining two cultivars, 'Romische Zellernuss' and 'Valcea', clustered separately (C). The cultivars in cluster A did not show any grouping based on geographical origin. In contrast, some clustering by origin was observed for the wild material. For example, a distinct subcluster of cluster B contained 18 of the 23 wild accessions from Vipava-Razdrto and Bovec (78%), 'Willamette' and a single wild accession from Maribor. In addition, seven of the eight remaining Maribor accessions were closely grouped in the AFLP dendrogram, the rest of the wild accessions in cluster B were intermixed.

The dendrogram constructed with the SSR data consisted of four clusters with 31, 46, 21, and 4 accessions in clusters A to D, respectively (Fig. 1). The minimum genetic dissimilarity between hazelnut accessions was 0.22 and the maximum dissimilarity was 0.85 with a mean of 0.58. All but three (94%) of the cultivars ('Tonda di Giffoni', 'Pauetet' and 'Valcea') were found in cluster B which only contained one wild accession (accession 10 from Dolenjska). As with the AFLP dendrogram, the cultivars did not show any clustering by geographical origin. In addition, most of the wild accessions from Vipava-Razdrto and Bovec were intermixed and separate from the other accessions in cluster A. Similar intermixing was seen for wild accessions from Maribor, Koroska, and Dolenjska in clusters A and C.

Principal coordinate analysis of both molecular marker datasets showed clear separation of the wild accessions from the cultivars (Suppl. Fig. 2). As with the dendrogram analysis, the Vipava-Razdrto and Bovec accessions clustered together and were distinct from the other wild accessions which were intermixed in the lower half of the two-dimensional PCoA plot. Nearly all of the cultivars fell in the upper right quadrant of the PCoA plot and were more tightly clustered than the wild material.

Average Dice coefficient dissimilarity values were calculated for the SSR dataset (Suppl. Table 5) to compare the diversity present in wild vs. cultivated accessions and in accessions from different regions. As expected, the wild material was more diverse than the cultivars with mean dissimilarity values of 0.60 and 0.50, respectively. Among the different regions where wild accessions were collected, Vipava-Razdrto (0.61), Bovec (0.57), and Maribor (0.55) had the most diverse material. Dissimilarity values calculated using the AFLP data showed similar trends (data not shown).

Population structure

Population structure analysis was performed with both the AFLP and SSR datasets and similar results were obtained. For that reason, only the SSR results are given here. According to the analysis, the data were best described by a $K = 2$ model, indicating that the material fell into two subpopulations. Based on a subpopulation identity threshold of $P \geq 0.7$, 62 individuals were assigned to subpopulation 1, 21 individuals were assigned to subpopulation 2, and 19 individuals were admixed. All but five of the hazelnut cultivars belonged to subpopulation 1 with the remaining accessions (101, 'F-104', 'Bandnuss', 'Pauetet', and 'Valcea') showing an admixed ancestry (Suppl. Fig. 3, Suppl. Table 1). The wild accessions were nearly equally divided between subpopulations 1 and 2 with 19 and 21 individuals in each subpopulation, respectively. The remaining 14 (26%) wild accessions were admixed. When the wild material was examined by region, all of the wild accessions from Bovec and most from Vipava-Razdrto (8 of 12 accessions) belonged to subpopulation 2 while the Dolenjska accessions (8 of 11) primarily fell into subpopulation 1 (Suppl. Fig. 3). Both Koroska and Maribor had higher incidence of admixed accessions with 36 and 56%, respectively (Suppl. Fig. 3).

Cultivar origin

The cultivars were subjected to PCoA and plotted in two dimensions (Fig. 2) to see if clustering was explained by the genetic background of the material. Seven of the cultivars were related to 'Tonda Gentile delle Langhe' while the 'Cosford' background was found in five cultivars (Suppl. Table 1), 'Fertile de Coutard' and 'Nocchione' backgrounds were found in four cultivars each while 'Compton' background was found in three cultivars. 'Tonda Gentile delle Langhe' and 'Cosford'-related cultivars showed no grouping in the PCoA. In contrast, 'Fertile de Coutard', 'Nocchione', and 'Compton'-related material all showed similar clustering on the left side of the graph.

Nut and kernel traits

A total of 64 accessions was randomly selected for the association panel including 24 cultivars and 40 wild accessions.

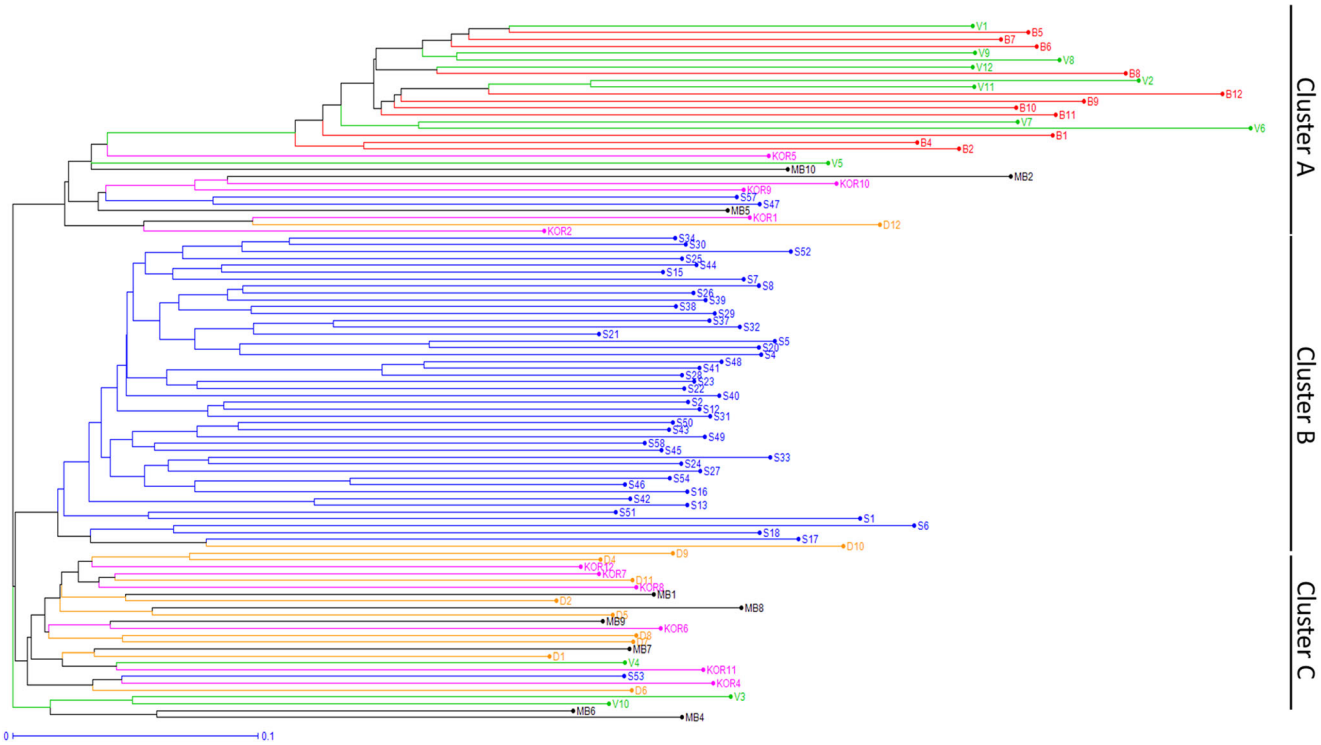

Fig. 1 Unweighted neighbor-joining dendrogram of the 102 hazelnut accessions based on SSR data. Accessions are color coded by origin: *blue* cultivar, *red* Bovec, *orange* Dolenjska, *purple* Koroska, *black* Maribor, *green* Vipava-Razdrto

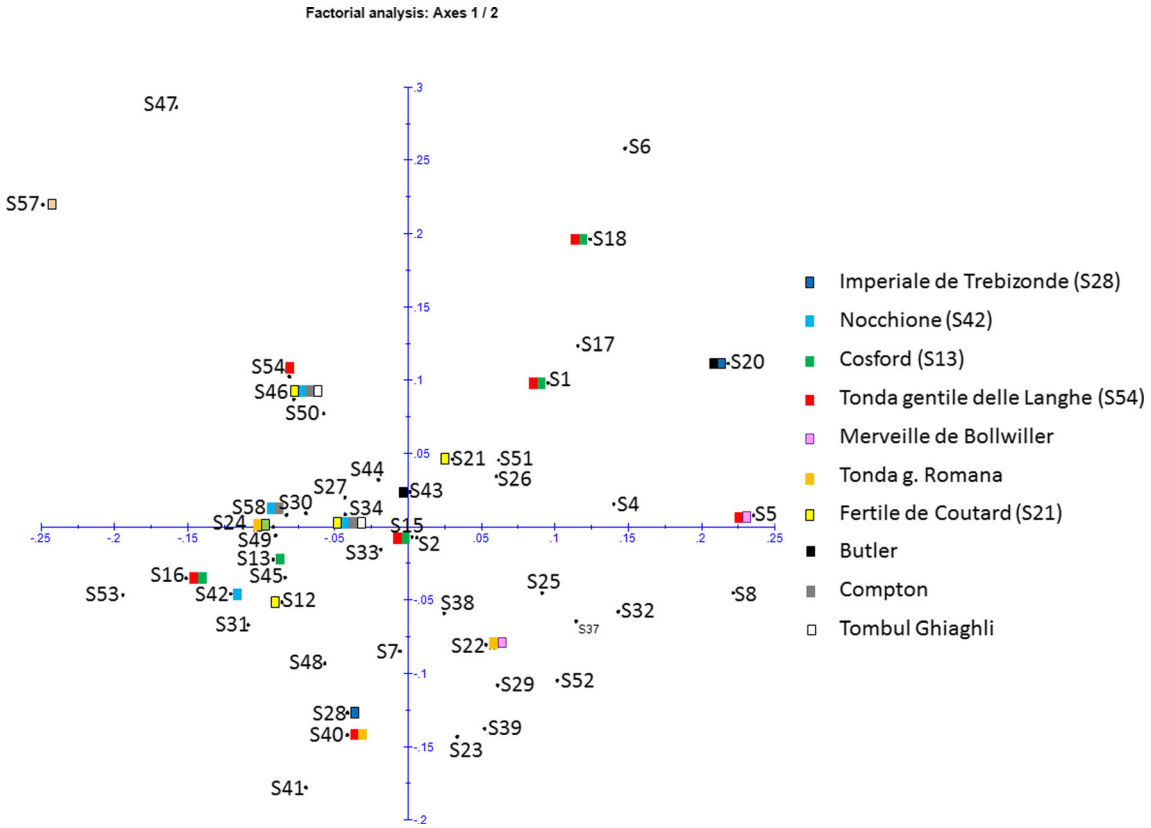

Fig. 2 PCoA of hazelnut cultivars based on SSR results. Genetic background of cultivars is indicated by *colored boxes*

The panel was characterized for 10 nut and seven kernel traits. Cultivars and wild accessions were analyzed separately. The wild accessions were also analyzed according to region.

Nut traits

Mean nut length, width, and thickness were significantly higher for the cultivars (means of 21, 20, 17.5 mm, respectively) than for wild accessions (means of 17.4, 14.5, 12.3 mm, respectively) (Table 1). Cultivars also had slightly higher coefficient of variation (CV) than wild accessions for these traits. ANOVA and LSD (least significant difference) tests showed that there was no significant difference between the means for these three parameters for wild accessions from different regions of Slovenia ($P \leq 0.05$). Nut width and thickness had the least variation in the wild accessions from Maribor (Suppl. Table 7).

Mean caliber of wild accessions (17.4 mm) was significantly higher than that of cultivars (15.8 mm) (Table 1); however, both sets of material had similar amounts of variation for this trait (CV 10 to 12%). There were no significant differences between calibers for each region of Slovenia ($P \leq 0.05$). Wild accessions from all regions with the exception of Bovec (5%) preserved the variation of the wild accessions for caliber (Suppl. Table 7). Although hazelnut cultivars had slightly higher mean shape index (0.9) than wild accessions (0.8), cultivars and wild accessions had similar amounts of variation for shape index. Wild accessions from different regions had similar shape index with the least variation for this trait in Bovec and Koroska (Suppl. Table 7).

Although mean nut shell thickness of the cultivars and wild accessions were the same (1.1 mm), wild accessions had more variation (CV = 20%) than cultivars (CV = 15%) (Table 1). There were no significant differences between wild accessions from each region for shell thickness (Suppl. Table 7).

Hazelnut cultivars had twofold higher mean nut weight (2.8 g) than wild accessions (1.3 g) (Table 1). Wild accessions had higher CV (32%) than cultivars (24%) for this trait with nuts as small as 0.6 g (Table 1). Hazelnuts from Maribor were the lightest (0.9 g); however, this difference was not significant. Wild accessions from Koroska and Vipava-Razdrto had higher nut weight variation than the other regions with one accession from Koroska having nearly the same mean weight as the cultivars (2.6 vs. 2.8 g) (Suppl. Table 7). Mean nut shape uniformity of wild accessions was higher (8) than that of the cultivars (7.5) (Table 1) with very little variation for this trait in both sets of material. There was no significant difference for nut shape uniformity in different regions of Slovenia (Suppl. Table 7).

The proportion of healthy nuts for cultivars (90.8%) was much higher than for wild accessions (57.3%) with much more variation for this parameter in the wild accessions

Table 1 Nut and kernel traits for hazelnut accessions

	Cultivars (S)			Wild material (all regions)		
Trait	Mean ± SE	Range	CV (%)	Mean ± SE	Range	CV (%)
Nut						
Length, mm	21.0 ± 0.7 a	16.5–29.2	15	17.4 ± 0.3 b	12.9–21.7	11
Width, mm	20.0 ± 0.6 a	17.0–27.6	14	14.5 ± 0.3 b	12.1–17.5	9
Thickness, mm	17.5 ± 0.5 a	14.2–24.4	14	12.3 ± 0.3 b	10.3–14.9	11
Caliber, mm	15.8 ± 0.4 a	12.2–20.3	12	17.4 ± 0.3 b	13.0–21.7	10
Shape index	0.9 ± 0.1 a	0.7–1.2	12	0.8 ± 0.1 b	0.5–0.9	12
Shell thickness, mm	1.1 ± 0.1 a	0.9–1.6	15	1.1 ± 0.1 a	0.7–1.7	20
Weight, g	2.8 ± 0.1 a	1.9–4.3	24	1.3 ± 0.1 b	0.6–2.6	32
Shape uniformity	7.5 ± 0.1 a	6.0–8.0	9	8.0 ± 0.1 b	7.0–8.0	2
Healthy, %	90.8 ± 1.2 a	82.0–100.0	7	57.3 ± 5.8 b	0–100.0	65
Empty, %	4.2 ± 0.8 a	0–13.3	96	3.7 ± 2.4 a	0–90.0	404
Kernel						
Weight, g	1.2 ± 0.1 a	1.0–1.9	20	0.4 ± 0.1 b	0.1–1.1	50
Kernel percentage	46.3 ± 0.1 a	33.8–52.8	10	32.1 ± 1.8 b	11.1–73.3	34
Shape uniformity	6.5 ± 0.3 a	4.0–8.0	17	6.5 ± 0.3 a	2.0–8.0	22
Brown spots, %	0.6 ± 0.2 a	0–2.5	156	0	–	–
Moldy, %	0.4 ± 0 a	0–2.5	165	0	–	–
Twins, %	2.0 ± 0.7 a	0–15.0	179	0	–	–
Deformed, %	2.7 ± 0.5 a	0–9.0	83	38.7 ± 5.4 b	0–100.0	89

Means with different letters within a row are significantly different according to analysis of variance (ANOVA) and least significant difference (LSD) test ($P < 0.05$)

(CV = 65% vs. 7%) (Table 1). The Dolenjska and Vipava-Razdrto accessions had significantly healthier nuts (80 to 88%) than the Koroska accessions (12.5%). Cultivars and wild material had similarly low proportions of empty nuts. However, the wild accessions had much more variation for this trait with one accession producing 90% empty nuts. All wild accessions except those from Koroska and Maribor had no empty nuts (Suppl. Table 7).

Kernel traits

Hazelnut cultivars had threefold higher mean kernel weight (1.2 g) than wild accessions (0.4 g) with more variation for this trait in the wild material (Table 1). The wild accessions from all regions except Dolenjska (0.6 g) had similar kernel weight (0.3 to 0.4 g) with the most variation in Koroska (CV = 63%) (Suppl. Table 7). Mean kernel percentage for cultivars (46.3%) was significantly higher than for wild accessions (32.1%) which had more variation for this trait (Table 1). The wild accessions from Dolenjska had significantly higher mean kernel percentage than the other regions (Suppl. Table 7).

There was no difference between cultivars and wild accessions for kernel shape uniformity as both sets of material had moderate uniformity values of 6.5 (out of 9) (Table 1). Wild accessions from different regions had similar uniformity; however, those from Maribor had much higher variation (CV = 46%) than the other regions (Suppl. Table 7).

Proportion of kernels with brown spots and mold were low for both cultivars and wild accessions (0 to 0.6%) with absolutely no variation for this trait in the wild material (Table 1, Suppl. Table 7). Similarly, none of the wild accessions had twin kernels while a low percentage of the cultivars (2.0%) had this trait. Wild accessions had a 14-fold higher proportion of deformed kernels (38.7%) than cultivars (2.7%) with some wild accessions having 100% deformed kernels. Wild accessions from Dolenjska (19.2%) and Vipava-Razdrto (11.6%) had significantly fewer deformed kernels with most of the variation for this trait in Dolenjska (CV = 96%) and Maribor (CV = 130%) (Table 1, Suppl. Table 7).

Trait correlations and principal component analysis

Nut length had a high positive correlation to caliber (r^2 = 0.99); moderate positive correlations to nut (r^2 = 0.52) and kernel weights (r^2 = 0.45); and a moderate negative correlation with nut shape index (r^2 = −0.64) (data not shown). There was high correlation between nut width and thickness (r^2 = 0.90) and nut weight (r^2 = 0.72). In addition, there were moderate positive correlations between width and nut shape index (r^2 = 0.5), shell thickness (r^2 = 0.35), and kernel weight (r^2 = 0.56). There was also moderate positive correlation between shell thickness and shape index (r^2 = 0.54) and nut weight (r^2 = 0.60). There was negative moderate correlation between shape index and caliber (r^2 = −0.62). Positive moderate correlations were observed between caliber and nut (r^2 = 0.53) and kernel (r^2 = 0.50) weights. As expected, there was positive correlation between nut and kernel weight; however, this correlation was only moderate (r^2 = 0.65). Proportion of healthy nuts was correlated to four traits: kernel weight (r^2 = 0.45), kernel percentage (r^2 = 0.50) and kernel shape uniformity (r^2 = 0.70); and proportion of empty nuts (r^2 = −0.40). There was negative moderate correlation between proportion of empty nuts and kernel shape uniformity (r^2 = −0.60). Kernel weight was highly correlated to kernel percentage (r^2 = 0.85).

Principal component analysis (PCA) of the nut and kernel traits was performed. The first three Eigen vectors explained 61.6% of the morphological variation. A total of 39% of the morphological variation was explained by PC1 with high positive correlations (r^2 > 0.7) to nut and kernel weight; nut length, width, and thickness; and kernel percentage (Suppl. Table 8). PC2 explained 12.2% of the morphological variation with moderate positive correlations to deformed kernels, caliber, and nut length and negative moderate correlation to proportion of healthy nuts. PC3 explained 10.4% of the morphological variation with moderate positive correlations to caliber and length and high negative correlation to nut shape index (Suppl. Table 8). The two-dimensional PCA plot of the morphological data showed that the hazelnut cultivars formed a tight cluster compared to the wild materials which were widely distributed (Suppl. Fig 4). Thus, the PCA plot did not show region-specific clustering of wild accessions based on morphological traits.

Stepwise discriminant analysis was performed using subpopulation (as determined from population structure) as a grouping variable and all nut and kernel traits as independents. Kernel weight and percentage made significant contributions to multivariate discrimination of the three subpopulations with moderate correlations of r^2 = 0.71 and r^2 = 0.57, respectively. Kernel weight also made significant contributions to discrimination when dendrogram cluster was used as the grouping variable. When discriminant analysis was performed using regions as the grouping variable, only the proportion of healthy nuts made a significant but moderate (r^2 = 0.47) contribution to discrimination of the five regions.

Association mapping

A total of 504 SSR fragments generated from 49 SSR markers was associated with nut and kernel traits. Significant (−Log (*P* value) >3) linkage disequilibrium (LD) was detected for 855 (0.65%) SSR marker pairs. LD values (r^2) of these SSR marker pairs ranged from 0.17 to 1 with a mean of 0.32.

Different AM models [GLM, GLM (Q), GLM (PC), GLM (Q + PC), MLM (K), MLM (Q + K), MLM (PC + K), MLM

(Q + PC + K)] were compared and used to calculate the proportion of significant results (Suppl. Table 6). The GLM models had higher proportions of significant results than the MLM models. The GML model corrected with the population structure Q-matrix had the highest proportion of significant results among the GLM models [$\pi 1$ (%) = 9.9] and was used for AM of the nut and kernel traits (Suppl. Table 6).

Nut traits

Nine SSR markers were associated with nut length (Table 2). LD values (r^2) of these SSR markers ranged from 0.22 (B791-270) to 0.46 (A622-125). A622-125 also had a significance level [−Log (*P* value) = 6.55]. Ten SSR markers were associated with nut thickness. The LD values (r^2) of markers associated with the trait ranged from 0.22 to 0.40. Six of these markers (A622-125, A622-130, B791-270, A604-161, B777-48 and B716-216) were also associated with length. There were no SSR markers associated with width and shape index. Ten markers were identified for nut caliber with LD values ranging from 0.23 (CAC-B005-440) to 0.41 (A613-150). Nine of the ten caliber markers were also identified for nut length and thickness.

No SSR markers were significantly associated with shape index; however, two markers were identified for shell thickness. The LD values of these markers were 0.16 and 0.15 for B602-341 and B648 -261, respectively (Table 2). No markers were detected for nut weight, shape uniformity, and proportion of healthy nuts. Proportion of empty nuts was associated with the most SSRs, 22 markers. The LD values of these markers ranged from 0.22 to 0.49 with the greatest effects seen for A604-161 and A613-150.

Kernel traits

Only one SSR marker (B726-281) was associated with kernel weight with an LD value of 0.30. No SSR markers were associated with kernel percentage. Seven SSR markers were associated with kernel shape uniformity with LD values ranging from 0.23 (B640-81) to 0.42 (A604-161). Association mapping was also performed for proportion of kernel with brown spots, moldy kernels, and twin kernels. Among these, a significant association was only identified for proportion of moldy kernels. This SSR marker, B640-81, had an LD value of 0.30.

Discussion

Marker polymorphism

Both AFLP and SSR markers had good polymorphism in the 102 hazelnut accessions. The AFLP technique gave an average of 48.4 alleles per primer combination, a value that

Table 2 Hazelnut SSR markers associated with nut and kernel traits

Trait	SSR locus	−Log (*P* value)*	LD value (r^2)
Nut			
Length	A622-125	6.55	0.46
Length	B777-48	5.96	0.32
Length	A622-130	5.53	0.39
Length	A604-161	4.86	0.30
Length	A613-150	4.61	0.36
Length	B741-201	4.42	0.26
Length	B709-226	4.27	0.26
Length	B791-270	4.11	0.22
Length	B716-216	4.08	0.23
Thickness	B791-270	7.01	0.37
Thickness	B777-48	5.53	0.30
Thickness	B777-60	5.48	0.30
Thickness	A622-125	5.46	0.40
Thickness	B628-304	5.45	0.31
Thickness	A604-161	5.29	0.32
Thickness	B758-165	5.11	0.30
Thickness	A622-130	4.21	0.30
Thickness	B716-216	4.13	0.23
Thickness	CAC-B005-440	4.05	0.22
Caliber	B791-270	7.62	0.40
Caliber	A604-161	6.11	0.37
Caliber	B777-48	5.68	0.31
Caliber	B628-304	5.53	0.31
Caliber	A613-150	5.47	0.41
Caliber	B758-165	5.01	0.29
Caliber	A622-125	4.48	0.33
Caliber	B777-60	4.47	0.25
Caliber	B777-42	4.23	0.23
Caliber	CAC-B005-440	4.20	0.23
Shell thickness	B602-341	4.72	0.16
Shell thickness	B648-261	4.62	0.15
Empty	B628-304	9.10	0.48
Empty	A604-161	8.74	0.49
Empty	B791-270	8.41	0.43
Empty	A606-154	7.15	0.45
Empty	B758-165	6.80	0.39
Empty	A613-150	6.78	0.49
Empty	B640-81	6.31	0.35
Empty	B628-315	5.84	0.33
Empty	A606-146	5.60	0.37
Empty	B603-287	5.55	0.32
Empty	B777-48	5.35	0.29
Empty	B625-253	4.95	0.27
Empty	CAC-B005-440	4.87	0.26
Empty	CAT-C504-235	4.48	0.28
Empty	B791-343	4.37	0.24
Empty	CAC-B005-476	4.32	0.24
Empty	A622-172	4.26	0.31

Table 2 (continued)

Trait	SSR locus	−Log (*P* value)*	LD value (r^2)
Empty	B790-211	4.09	0.22
Empty	A606-184	4.06	0.27
Empty	A606-159	4.06	0.27
Empty	B648-55	4.05	0.23
Empty	A606-150	4.04	0.27
Kernel			
Weight	B726-281	4.76	0.30
Shape uniformity	A604-161	7.18	0.42
Shape uniformity	B791-270	5.75	0.31
Shape uniformity	B716-216	5.47	0.30
Shape uniformity	B777-48	5.21	0.29
Shape uniformity	A622-130	4.92	0.35
Shape uniformity	A622-125	4.70	0.35
Shape uniformity	B640-81	4.01	0.23
Moldy	B640-81	5.32	0.30

*Negative log10-transformed *P* values

was much higher than reported in previous work which averaged 29.9 and 24.4 polymorphic fragments, respectively. This difference may be attributed to the fact that our work examined more accessions and many more wild accessions than Kafkas et al. (2009) who examined only 18 Turkish cultivars while Martins et al. (2014) studied 58 accessions, only 13 of which were wild accessions. As expected, SSR markers gave fewer alleles than AFLP but proved to be more informative with an average gene diversity of 0.30 compared to 0.26 for AFLP. The 49 SSR markers had an average of 10.3 alleles per SSR which was consistent with other work with the same or similar genomic SSR markers which resulted in 3 to 10.6 alleles per marker (Bassil et al. 2005; Boccacci et al. 2005, 2008; Gokirmak et al. 2009; Boccacci and Botta 2010; Gurcan and Mehlenbacher 2010; Gurcan et al. 2010a, 2010b; Campa et al. 2011; Bassil et al. 2013). The distance matrices generated from the AFLP and SSR marker data had only low correlation to each other perhaps because they sampled different parts of the hazelnut genome. While both AFLP and SSR markers are expected to occur in coding and noncoding regions throughout the genome, the majority (96%) of SSR markers used in this study were previously selected to provide only one or two amplification products (Gurcan et al. 2010a). The existence of fewer alleles for these markers may indicate that they are located in coding regions under selective pressure.

Diversity and population structure of wild and cultivated hazelnuts

Gene diversity values for the AFLP and SSR markers and the genetic dissimilarity matrix calculated with the Dice coefficient indicated high genetic diversity of hazelnut in agreement with many other studies (Gokirmak et al. 2009; Gurcan et al. 2010a, 2010b; Campa et al. 2011; Boccacci and Botta 2010; Boccacci et al. 2006, 2008, 2013). Levels of diversity are expected to be high in hazelnut because it has a self-incompatible mating system which prevents self-pollination. In addition, hazelnut is a wind-pollinated species. Overall diversity of the cultivars was lower than the wild accessions with average dissimilarity coefficients of 0.50 and 0.60, respectively. Higher levels of diversity in wild accessions as compared to cultivars were also reported by Campa et al. (2011) and Martins et al. (2015) who studied material from northern Spain and northern Portugal, respectively. Such results are expected as hazelnut cultivars have been selected for certain similar features, are clonally propagated, and therefore, genetically fixed. In contrast, wild individuals are the result of sexual reproduction with cross-pollination which allows greater gene flow and an increased probability of recombination events leading to new alleles and greater genetic variability.

Genetic diversity and relationships among the hazelnut accessions were determined using dendrogram and PCoA analyses of the AFLP and SSR data. Based on these results, it was clear that most cultivars (91–94%, depending on dataset) grouped separately from the wild material. This was also observed in comparisons between cultivated and wild material sampled in Spain, Italy, and Portugal (Campa et al. 2011; Boccacci et al. 2013; Martins et al. 2014, 2015). Similar separation was also seen when reference cultivars were analyzed with local cultivated germplasm from northern Spain (Ferreira et al. 2010). Unlike other studies (Gokirmak et al. 2009; Boccacci and Botta 2009, 2010; Gurcan et al. 2010a; Boccacci et al. 2013), cultivars did not cluster according to geographic origin. This was not surprising given the relatively similar genetic origins of hazelnut breeding material which has been spread to several different countries. However, clustering by origin was observed for the wild accessions. Thus, most of the accessions from western Slovenia (Bovec and Vipava-Razdrto) formed a group which was distinct from the cluster of accessions from central and eastern Slovenia (Dolenjska, Koroska and Maribor). Vipava-Razdrto and Bovec are located at high elevation, 533 m and 452 m above sea level, respectively (Suppl. Table 2). Both regions have temperate climate with dry winters, no dry season, at least 60 mm of precipitation per month and a Mediterranean influence. In such an environment, the genetically close relationship between accessions could be preserved, explaining the occurrence of all accessions from Bovec and 75% of Vipava-Razdrto accessions in the same cluster (A).

The plant material from Koroska grows in a small, isolated location in north Slovenia, characterized by a humid continental climate that is colder than the other regions examined in this work (Suppl. Table 2). In this area, the local forestry

service protects hazelnut bushes that are at least 20 years old in order to keep the species in the region. According to foresters, hazelnut is very rarely spread in the region and of unknown origin. The intermixed clustering of some Koroska accessions with those from Bovec and Vipava-Razdrto (Suppl. Fig. 1), suggests that some Koroska hazelnuts originated from the western part of Slovenia.

The high genetic dissimilarity of the accessions from Maribor region with 33, 45, and 22% of accessions belonging to clusters A, C, and D, respectively, could be explained by the possible influence of commercial cultivars. Maribor belongs to the wider Stajerska region, one of the most important areas of commercial hazelnut growth. Therefore, wind pollination among trees could contribute to genetic variability in this region.

Population structure analysis was consistent with the dendrogram and PCoA results. For example, 95% of the subpopulation 1 accessions were located in dendrogram clusters B and C while 100% of subpopulation 2 accessions were in dendrogram cluster A. Regardless of marker type, the hazelnut accessions fell into two subpopulations which appeared to correspond to the cultivated (subpopulation 1) and wild (subpopulation 2) gene pools. Thus, with only a few cases of admixing, all but 10.4% of the cultivars fell into subpopulation 1. In contrast, the wild material fell mainly into subpopulation 2 (39%) with significant proportions belonging to subpopulation 1 (35%) and the admixed group (26%). These results may indicate that the wild accessions in subpopulation 1 originated from cultivars which escaped cultivation. Similarly, the admixed accessions may be the result of ancestral, natural crosses between wild and cultivated trees. Interestingly, the vast majority (83%) of wild accessions from the western part of Slovenia belonged to subpopulation 2. It would be useful to compare these accessions with wild material from adjacent northern Italy to see if they share the same gene pool. If so, the results may suggest the movement of wild material from Italy, a proposed origin of domestication and diversification for hazelnut (Boccacci and Botta 2009, 2010; Boccacci et al. 2013), to parts of Slovenia. Such west to east spread of hazelnut was first proposed by Boccaci and Botta (2009). In contrast to the western accessions, most (56%) of the accessions from the northeast (Maribor) had admixed ancestry. As previously stated, Maribor is in the Stajerska region which is one of the primary areas of hazelnut cultivation in Slovenia. Thus, it is likely that the collected accessions are the result of cross-pollination between truly wild individuals and cultivars. In general, our results suggest that there is gene flow within Slovenian western and eastern hazelnut populations but very little intermixing between them. Moreover, the differentiation of wild and cultivated hazelnuts into two distinct subpopulations suggests that the wild germplasm could be a useful source of genetic diversity and new traits for hazelnut improvement (Martins et al. 2015). Of course, it is also necessary to combine these molecular genetic results with morphological data in order to make selections that will ensure the future diversity and improvement of the crop in Slovenia and other hazelnut growing regions.

Morphological evaluation

Improvement of hazelnut cultivars for nut and kernel traits is essential to increase market value. Only about 20 cultivars are grown worldwide for the confectionery industry or in-shell marketing (Cristofori et al. 2008). Wild accessions could be used in hazelnut breeding to introduce new variation and improve cultivars for nut and kernel traits.

Volume of the nut and kernel is important because larger nuts are preferred for in-shell marketing, while medium and small nuts are preferred for confectionery (Solar and Stampar 2011). Nut and kernel volume and uniformity is also important for correct operation of processing machines used in the food industry (Ozdemir and Akinci 2004). Cultivars had greater nut length, width, and thickness than wild accessions. This result was expected because nut volume is determined by these dimensional traits which have positive effects on increasing yield. The cultivars analyzed in the present study were, on average, longer, and wider than the 20 hybrids (from Arbor Day Farm, Nebraska City, NE) analyzed by Xu and Hanna (2010). They were also longer, wider, and thicker than four Turkish cultivars analyzed by Ozdemir and Akinci (2004). Wild Slovenian accessions had similar length and width as the aforementioned 20 hybrids but were, on average, shorter, and narrower than two Turkish cultivars (‘Palaz’ and ‘Tombul’) (Xu and Hanna 2010; Ozdemir and Akinci 2004). Mean caliber of wild accessions was higher than for hazelnut cultivars because these nuts were much longer than the cultivars. The wild accessions also had higher mean caliber (17.4) than the 16 cultivars (15.6) reported by Solar and Stampar (2011). Thus, this present study demonstrated that hazelnut wild accessions may be a good source of alleles for increased nut caliber.

Shape index is an important trait in hazelnut breeding. Globular cultivars (shape index = 1) can be more efficiently processed (Ozdemir and Akinci 2004). Cultivars and wild accessions in this study were nearly globular. Although the mean shape index of cultivars (0.9) was similar to the 16 hazelnut cultivars (0.89) studied by Solar and Stampar (2011), it was lower than that for 24 Italian and foreign cultivars (1.05) (Cristofori et al. 2008) and four Turkish hazelnut cultivars (0.96) (Ozdemir and Akinci 2004). Some of the wild accessions had shape indices similar to the cultivars.

Thin shell is preferred for efficient processing of hazelnut. Interestingly, there was no difference between the shell thicknesses of cultivars and wild accessions. Both cultivars and wild accessions had similar mean shell thickness as 16 previously examined cultivars (Solar and Stampar 2011). Two wild

accessions from Dolenjska and Maribor had the thinnest shell (0.7 mm) and could be used to develop cultivars with thinner shell.

Nut weight has a direct effect on yield and cultivars are expected to have big nuts. Wild accessions did not have superior alleles for nut weight. This demonstrated that cultivars were primarily selected and adapted from wild accessions to have greater nut weight. The cultivars examined in this study had higher mean nut weight than 20 hybrids (0.6 g) (Xu and Hanna 2010), 24 Italian and foreign cultivars (2.4 g) (Cristofori et al. 2008), and four Turkish cultivars (1.8 g) (Ozdemir and Akinci 2004). The cultivars had similar mean weight as the 16 cultivars (3.0 g) analyzed by Solar and Stampar (2011).

In addition to a globular shape, uniform nuts are best for processing. Thus, cultivars and wild accessions were evaluated for nut shape uniformity. Interestingly, wild accessions had more uniform nuts than the cultivars analyzed in present study and the 16 cultivars (7.5) examined by Solar and Stampar 2011. This showed that wild accessions contain good genetic potential for nut uniformity which can be introduced to cultivars.

The proportions of healthy and empty nuts are yield-related traits and are also affected by environmental conditions and can vary greatly by year. During the tested years, wild accessions did not contain high genetic potential for proportion of healthy nuts; however, some wild accessions from Vipava-Razdrto, Dolenjska, and Maribor had 100% healthy nuts and can be used in breeding programs for improvement of nut health. Both cultivars and wild accessions had very few empty nuts. Wild accessions from Bovec, Dolenjska, and Vipava-Razdrto did not contain any empty nuts and they can be used to decrease the proportion of empty nuts in cultivars.

A main objective of hazelnut breeding programs is improvement of kernel weight because of its direct effect on yield. Unfortunately, wild accessions did not have good alleles for kernel weight. The cultivars had higher mean kernel weight than the 20 hazelnut hybrids (0.6 g) studied by Xu and Hanna (2010), the 24 Italian and foreign cultivars (1.1 g) examined by Cristofori et al. 2008 and the four Turkish cultivars (1.0 g) (Ozdemir and Akinci 2004). These results demonstrated that the cultivars of the present study were, on average, superior to the previously studied material and/or grown under more favorable conditions.

Kernel percentage represents the proportion of nut weight that is accounted for by the kernel and is important for the food industry. Wild accessions had low genetic potential for kernel percentage. Mean kernel percentage of cultivars was higher than the 20 hazelnut hybrids (38.8) studied by Xu and Hanna (2010) and slightly higher than 24 Italian and foreign cultivars (44.3%) (Cristofori et al. 2008) and 16 cultivars (44.6%) examined by Solar and Stampar (2011).

Although kernel shape uniformity is an important trait for hazelnut processing and could be considered as a breeding goal, cultivars, and wild accessions had similar, moderate levels of uniformity. Indeed the material had similar kernel shape uniformity as the previously examined 16 hazelnut cultivars (6.2) (Solar and Stampar 2011). These results indicate that new source of alleles for improved kernel shape uniformity should be introduced to breeding programs.

Hazelnut kernels with brown spots are an important problem in the processing industry because they cannot be processed because they split easily. Brown spots in kernels can decrease hazelnut production up to 30% (Romero et al. 2003). Cultivars had a very low proportion of brown spots. More interestingly, wild hazelnut accessions did not have any kernels with brown spot. Similar results were obtained for the proportion of moldy kernels with no mold found in the wild accessions. Kernel mold is caused by fungal species, decreases quality for both the in-shell and confectionery industry markets, and is, of course, influenced by climatic conditions (Teviotdale et al. 2002). These results show that wild accessions can be used as allele sources to decrease the proportions of brown-spotted and moldy kernels in hazelnut. Twin kernels result from the development of two kernels in one nut. A higher proportion of twin kernels decreases kernel quality. The cultivars analyzed in the present study had extra quality according to the quality classification of the United Nations (UNECE 2010) because cultivars had just 2% twin kernels. According to the classification system, nuts with more than 2% twins are placed in the lower quality classes. Interestingly, wild accessions did not have any twin kernels. In contrast to the proportions of kernel with brown spots, moldy kernels, and twin kernels, wild accessions did not have superior alleles for deformed kernels. Thus, such material is not useful for improvement of this trait.

Clustering of the hazelnut cultivars in the PCoA was expected because cultivars are more improved than wild accessions for most of the nut and kernel traits. Despite the seemingly limited genetic potential of wild accessions when examined using mean values, sizable variation was observed for all traits expect nut shape uniformity. Similarly, interesting genetic variability was seen when local cultivated germplasm from northern Spain was examined for morphological traits (Ferreira et al. 2010). Thus, wild and local accessions may be useful to increase genetic diversity of hazelnut breeding material as well as improve some breeding targets.

Association mapping

The proportion of SSR markers (0.65%) showing significant LD in the present study was slightly lower than the proportion of SSR markers (11 and 6.5%) with significant LD in AM panels of *Gossypium hirsutum* germplasm and opium poppy (Abdurakhmonov et al. 2008; Celik et al. 2016, respectively).

This can be due to the higher genetic diversity of trees as compared to annual plants. LD identified in this study might be due to linkage of the markers but this could not be confirmed because the SSR markers used in this study are not mapped in the hazelnut genome. LD can also be due to selection and relatedness of hazelnut accessions which can lead to false positive associations between markers and traits (Pritchard et al. 2000; Stich et al. 2005). To avoid false positive associations, association mapping was corrected by the population structure Q-matrix.

In the present study, QTLs controlling nut and kernel traits were identified for the first time. A total 49 SSR markers associated with nine of 17 traits was identified using an association mapping approach. Some of these markers could be useful for marker-assisted selection of hazelnut accessions for morphological traits. No QTLs were detected for eight traits. This lack of QTLs for some parameters may be due to reduced genetic diversity of these traits in the AM panel.

In conclusion, hazelnut cultivars and wild accessions from Slovenia were examined for genetic diversity and morphological variation. The wild accessions were more diverse than cultivars at the molecular level with clustering of the wild material by region. Characterization of nut and kernel traits was done to assess the breeding potential of the wild germplasm. Wild accessions were shown to have breeding potential for most of the traits except nut and kernel weight and to have sizable variation for most traits. In addition, the first association mapping of hazelnut was performed with the identification of SSR markers associated with traits including the length, thickness, and caliber and nuts, as well as, kernel weight and shape uniformity. These SSR markers provide initial molecular information for marker-assisted selection in hazelnut.

Acknowledgments This study was supported by The Scientific and Technological Research Council of Turkey (TUBITAK, project no: 212T201). It was carried out in the frame of The Bilateral Scientific and Technological Cooperation between the Republic of Slovenia and Turkey, BI-TR/13-15-003. We thank Amy Frary for review of the manuscript.

Author contribution statement SCO and SEO performed laboratory experiments and drafted manuscript; SCO analyzed the genetic diversity data; IC analyzed the association mapping data; FS, RV, and AS performed all morphological characterization; AF helped design the study and write the manuscript; SD designed the study and received funding.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Data archiving Data will be available at http://plantmolgen.iyte.edu.tr/data/ upon publication.

References

Abdurakhmonov IY, Kohel RJ, Yu JZ, Pepper AE, Abdullaev AA, Kushanov FN, Abdukarimov A (2008) Molecular diversity and association mapping of fiber quality traits in exotic *G. hirsutum* L. germplasm. Genomics 92:478–487

Bacchetta L, Avanzato D, Di Giovanni B, Botta R, Boccacci P, Drogoudi P, Metzidakis I, Rovira M, Sarraquigne JP, Silva AP, Solar A, Spera D (2014) The reorganisation of European hazelnut genetic resources in the SAFENUT (AGRI GEN RES) project. Acta Hort 1052:67–74

Balta MF, Yarılgaç T, Aşkın MA, Kuçuk M, Balta F, Özrenk K (2006) Determination of fatty acid compositions, oil contents and some quality traits of hazelnut genetic resources grown in eastern Anatolia of Turkey. J Food Compos Anal 19:681–686

Bassil NV, Botta R, Mehlenbacher SA (2005) Additional microsatellite markers of the European hazelnut. Acta Hortic 686:105

Bassil N, Boccacci P, Botta R, Postman J, Mehlenbacher SA (2013) Nuclear and chloroplast microsatellite markers to assess genetic diversity and evolution in hazelnut species, hybrids and cultivars. Gen Res Crop Evol 60:543–568

Beltramo C, Valentini N, Portis E, Torello D, Boccacci P, Angelica M, Botta R (2016) Genetic mapping and QTL analysis in European hazelnut (*Corylus avellana* L.). Mol Breed 36:27

Boccacci P, Botta R (2009) Investigating the origin of hazelnut (*Corylus avellana* L.) cultivars using chloroplast microsatellites. Gen Res Crop Evol 56:851–859

Boccacci P, Botta R (2010) Microsatellite variability and genetic structure in hazelnut (*Corylus avellana* L.) cultivars from different growing regions. Sci Hortic 124:128–133

Boccacci P, Akkak A, Bassil NV, Mehlenbacher SA, Botta R (2005) Characterization and evaluation of microsatellite loci in European hazelnut (*Corylus avellana* L.) and their transferability to other Corylus species. Mol Ecol Notes 5:934–937

Boccacci P, Akkak A, Botta R (2006) DNA typing and genetic relationship among European hazelnut (*Corylus avellana* L.) cultivars using microsatellite markers. Genome 49:598–611

Boccacci P, Botta R, Rovira M (2008) Genetic diversity of hazelnut (*Corylus avellana* L.) germplasm in northeastern Spain. Hort Sci 43:667–672

Boccacci P, Aramini M, Valentini N, Bacchetta L, Rovira M, Drogoudi P, Silva AP, Solar A, Calizzano F, Erdogan V, Cristoferi V, Ciarmiello LF, Contessa C, Ferreira JJ, Marra FP, Botta R (2013) Molecular and morphological diversity of on-farm hazelnut (*Corylus avellana* L.) landraces from southern Europe and their role in the origin and diffusion of cultivated germplasm. Tree Genet Genomes 9:1465–1480

Boccacci P, Beltramo C, Prando MS, Lembo A, Sartor C, Mehlenbacher SA, Botta R, Marinoni DT (2015) In silico mining, characterization and cross-species transferability of EST-SSR markers for European hazelnut (*Corylus avellana* L.). Mol Breed 35:1–14

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635

Campa A, Trabanco N, Pérez-Vega E, Rovira M, Ferreira JJ (2011) Genetic relationship between cultivated and wild hazelnuts (*Corylus avellana* L.) collected in northern Spain. Plant Breed 130:360–366

Celik I, Camci H, Kose A, Kosar FC, Doganlar S, Frary A (2016) Molecular genetic diversity and association mapping of morphine content and agronomic traits in Turkish opium poppy (Papaver somniferum) germplasm. Mol Breed 36(4):1–13

Cristofori V, Ferramondo S, Bertazza G, Bignami C (2008) Nut and kernel traits and chemical composition of hazelnut (*Corylus avellana* L.) cultivars. J Sci Food Agric 88:1091–1098

Dice LR (1945) Measures of the amount of ecologic association between species. Ecology 26:297–302

Earl DA, von Holdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Ferrari M, Gori M, Monnanni R, Buiatti M, Scarascia Mugnozza GT, De Pace C (2004) DNA fingerprinting of *Corylus avellana* L. accessions revealed by AFLP molecular markers. Acta Hortic 686:125–134

Ferreira JJ, Garcia-Gonzalez C, Tous J, Rovira M (2010) Genetic diversity revealed by morphological traits and ISSR markers in hazelnut germplasm from northern Spain. Plant Breed 129:435–441

Food and Agriculture Organization of the United Nations, FAOSTAT (2013) http://www.fao.org/corp/statistics/en. Accessed 07.07.2014

Fulton TM, Chunwongse J, Tanksley SD (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Mol Biol Rep 13:207–207

Galderisi U, Cipollaro M, Di Bernanrdo G, De Masi L, Galano G, Cascino A (1999) Identification of hazelnut (*Corylus avellana* L.) cultivars by RAPD analysis. Plant Cell Rep 18:652–655

Gao H, Williamson S, Bustamante CD (2007) A Markov chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. Genetics 176:1635–1651

Gökirmak T, Mehlenbacher SA, Bassil NV (2009) Characterization of European hazelnut (*Corylus avellana* L.) cultivars using SSR markers. Genet Resour Crop Evol 56:147–172

Gómez G, Álvarez MF, Mosquera T (2011) Association mapping, a method to detect quantitative trait loci: statistical bases. Agronomía Colombiana 29:367–376

Gurcan K, Mehlenbacher SA (2010) Development of microsatellite marker loci for European hazelnut (*Corylus avellana* L.) from ISSR fragments. Mol Breed 26:551–559

Gurcan K, Mehlenbacher SA, Botta R, Boccacci P (2010a) Development, characterization, segregation, and mapping of microsatellite markers for European hazelnut (*Corylus avellana* L.) from enriched genomic libraries and usefulness in genetic diversity studies. Tree Genet Genomes 6:513–531

Gurcan K, Mehlenbacher SA, Erdogan V (2010b) Genetic diversity in hazelnut (*Corylus avellana* L.) cultivars from Black Sea countries assessed using SSR markers. Plant Breed 129:422–434

Kafkas S, Doğan Y, Sabır A, Turan A, Seker H (2009) Genetic characterization of hazelnut (*Corylus avellana* L.) cultivars from Turkey using molecular markers. Hortscience 44:1557–1561

Khan MA, Korban SS (2012) Association mapping in forest trees and fruit crops. J Exp Bot 63:4045–4060

Leinemann L, Steiner W, Hosius B, Kuchma O, Arenhövel W, Fussi B, Finkeldey R (2013) Genetic variation of chloroplast and nuclear markers in natural populations of hazelnut (*Corylus avellana* L.) in Germany. Plant Syst Evol 299:369–378

Martins S, Simões F, Mendonça D, Matos J, Silva AP, Carnide V (2013) Chloroplast SSR genetic diversity indicates a refuge for *Corylus avellana* in northern Portugal. Gen Res Crop Evol 60:1289–1295

Martins S, Simões F, Matos J, Silva AP, Carnide V (2014) Genetic relationship among wild, landraces and cultivars of hazelnut (*Corylus avellana* L.) from Portugal revealed through ISSR and AFLP markers. Plant Syst Evol 300:1035–1046

Martins S, Simões F, Mendonça D, Matos J, Silva AP, Carnide V (2015) Western European wild and landraces hazelnuts evaluated by SSR markers. Plant Mol Biol Rep 1–9

Mehlenbacher SA, Brown RN, Nouhra ER, Gokirmak T, Bassil NV, Kubisiak TL (2006) A genetic linkage map for hazelnut (*Corylus avellana* L.) based on RAPD and SSR markers. Genome 49:122–133

Miaja ML, Vallania R, Me C, Akkak O, Nassi O, Lepori G (2001) Varietal characterization in hazelnut by RAPD markers. Acta Hortic 556: 247–250

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci U S A 70:3321–3323

Norusis MJ (2010) PASW statistics 18 advanced statistical procedures. Prentice Hall Press, Upper Saddle River, NJ

Ozdemir F, Akinci I (2004) Physical and nutritional properties of four major commercial Turkish hazelnut varieties. J Food Eng 63:341–347

Perrier X, Jacquemoud-Collet JP (2006) DARwin software. Available: http://darwin.cirad.fr/darwin

Pritchard JK, Stephens M, Rosenberg NA, Donnelly P (2000) Association mapping in structured populations. Am J Hum Genet 67:170–181

Roldàn-Ruiz I, Dendauw J, Van Bockstaele E, Depicker A, De Loose M (2000) AFLP markers reveal high polymorphic rates in ryegrasses (Lolium spp.). Mol Breed 6:125–134

Romero A, Tous J, Durfort M, Rius M (2003) Histology of hazelnut (*Corylus avellana* L.) kernel affected by brown spots in kernel cavity physiopathy. Span J Agr Res 1:47–53

Slatnar A, Mikulic-Petkovsek M, Stampar F, VebericR SA (2014) HPLC-MS identification and quantification of phenolic compounds in hazelnut kernels, oil and bagasse pellets. Food Res Int 64:783–789

Solar A, Stampar F (2009) Performance of hazelnut cultivars from Oregon in northeastern Slovenia. HortTechnology 19:653–659

Solar A, Stampar F (2011) Characterisation of selected hazelnut cultivars: phenology, growing and yielding capacity, market quality and nutraceutical value. J Sci Food Agr 91:1205–1212

Stich B, Melchinger AE, Frisch M, Maurer HP, Heckenberger M, Reif JC (2005) Linkage disequilibrium in European elite maize germplasm investigated with SSRs. Theor Appl Genet 111:723–730

Storey JD (2002) A direct approach to false discovery rates. J R Stat Soc Ser B 64:479–498

Storey JD, Tibshirani R (2003) R: statistical significance for genome-wide experiments. Proc Natl Acad Sci U S A 100:9440–9445

Tanksley SD, McCouch SR (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. Science 277:1063–1066

Teviotdale BL, Michailides TJ, Pscheidt JW (2002) Compendium of nut crop diseases in temperate zones. APS Press, St. Paul, MN

United Nations (2010) Unece Standard DDP-04. https://www.nutfruit.org/wp-continguts/exportacions_mycored/standard-11_4597.pdf. Accessed 07.07.2014

Vos F, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new concept for DNA fingerprinting. Nucl Acids Res 23:4407–4414

Xu YX, Hanna MA (2010) Evaluation of Nebraska hybrid hazelnuts: nut/kernel characteristics, kernel proximate composition, and oil and protein properties. Ind Crop Prod 31:84–91

Zhu C, Gore M, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. The Plant Genome 1:5–20