

RESEARCH ARTICLE

Open Access



Association mapping for total polyphenol content, total flavonoid content and antioxidant activity in barley

Zhigang Han¹, Jingjie Zhang¹, Shengguan Cai¹, Xiaohui Chen¹, Xiaoyan Quan² and Guoping Zhang^{1*}

Abstract

Background: The interest has been increasing on the phenolic compounds in plants because of their nutritive function as food and the roles regulating plant growth. However, their underlying genetic mechanism in barley is still not clear.

Results: A genome-wide association study (GWAS) was conducted for total phenolic content (TPC), total flavonoid content (FLC) and antioxidant activity (AOA) in 67 cultivated and 156 Tibetan wild barley genotypes. Most markers associated with phenolic content were different in cultivated and wild barleys. The markers bPb-0572 and bPb-4531 were identified as the major QTLs controlling phenolic compounds in Tibetan wild barley. Moreover, the marker bPb-4531 was co-located with the UDP- glycosyltransferase gene (*HvUGT*), which is a homolog to Arabidopsis *UGTs* and involved in biosynthesis of flavonoid glycosides .

Conclusions: GWAS is an efficient tool for exploring the genetic architecture of phenolic compounds in the cultivated and Tibetan wild barleys. The DArT markers applied in this study can be used in barley breeding for developing new barley cultivars with higher phenolics content. The candidate gene (*HvUGT*) provides a potential route for deep understanding of the molecular mechanism of flavonoid synthesis.

Keywords: Antioxidant activity, Genome-wide association study (GWAS), UDP- glycosyltransferases, Phenolic compounds, Tibetan wild barley (*Hordeum spontaneum* L.)

Background

Barley (*Hordeum vulgare* L.) is mainly used as animal feed and raw material for malt and beer production. In recent decades, its use in the production of functional or healthy food has been increasingly focused because of its rich phytochemicals [1, 2]. Some phytochemicals in cereal grains, including phenolic acids, flavonoids and anthocyanins, can reduce risk of human chronic inflammation, cardiovascular diseases, certain cancers, and diabetes, and they are also involved in cell wall formation by forming bridges between polysaccharide and lignin [3, 4]. These phytochemicals in barley grains could be extracted as natural antioxidants. In fact, antioxidants have distinct effects on malting and brewing processes, including foam stability

and beer bitterness, and flavonoids may affect beer taste, color and haze formation [5–8].

Identification of genes or even QTLs responsible for phenolic metabolism is necessary for the genetic improvement of the trait. Although multiple studies have identified QTLs associated with phenolic compounds in rice and sorghum, there were few studies on total phenolic content (TPC), total flavonoid content (FLC) and antioxidant activity (AOA) in barley [9, 10]. Glycosylation is one of the key steps in flavonoid biosynthesis, as it promotes the solubility, stability and bioactivity of flavonoids [11]. UDP-glycosyltransferases (UGTs) are often characterized by a conserved plant secondary product glycosyltransferase (PSPG) box of 44 amino acids binding UDP-conjugates as their activated sugar donor substrates. In plants, UGTs are involved in the transferring of glycosyl moieties into a wide range of acceptors including flavonoids in the process of glycosylation [12]. A few UGT protein coding genes are associated with biosynthesis of

* Correspondence: zhanggp@zju.edu.cn

¹Department of Agronomy, Zhejiang Key Laboratory of Crop Germplasm, Zhejiang University, Hangzhou 310058, China

Full list of author information is available at the end of the article



flavonoid glycosides in *Arabidopsis thaliana* and soybean (*Glycine max*) [13, 14]. However, little is known about the relevant genes or QTLs controlling UGTs in barley. Moreover, the genetic diversity of cultivated barley became bottlenecked due to its domestication, posing a limitation for success of barley breeding. Tibetan wild barley, as one of the progenitors of cultivated barley, shows a wider genetic diversity in agronomic traits and abiotic stress tolerance [15–17]. Phenolics might have important functions in developing tolerance to salinity, high radiation and low temperature [18–20]. Wild barley may accumulate high polyphenol content in order to adapt to harsh environments with strong ultraviolet (UV) radiation and other abiotic stresses, including extremely variable temperature [20–23]. Therefore it is meaningful to identify the genetic factors associated with polyphenol synthesis and accumulation in wild barley.

In this study, we measured TPC, FLC and AOA contents in grains of 156 Tibetan wild barley accessions and 67 cultivated barley genotypes, and performed association mappings for these traits with aims at (1) determining the genotypic variation of TPC, FLC and AOA contents in barley grains; (2) identifying the QTLs associated with phenolics and antioxidant activity in cultivated and wild barley grains. (3) evolution analysis of marker-associated gene (*HvUGT*) with *UGTs* in *Arabidopsis thaliana*.

Methods

Plant materials

In total 223 barley genotypes (Additional file 1: Table S1), including 156 wild types (planted in two seasons of 2013 and 2014) and 67 cultivars (they were widely planted in southern east China since 1960s, planted in 2014) were planted at Zijingang campus of Zhejiang University (Hangzhou, China, 30°22'N, 119°26'E). Each genotype was sown in a plot consisting of three rows (each row was 2 m length and row distance is 0.25 m). All plots were supplied with 150 kg/ha of N, including 40 kg/ha of N as compound fertilizer applied before seeding and 110 kg/ha of N as urea supplied at two-leaf stage and booting stage with equal amount, respectively. In addition, 180 kg/ha of potassium chloride was applied prior to seeding. The experiment was arranged in a completely randomized block with three replicates. All other field managements, including weed and disease control, were the same as those applied locally. At maturity, barley grains were harvested, dried and then stored in a cool room (4 °C) for further analysis. The grain samples were milled using grinder to pass through a 0.5 mm screen.

Extraction of phenolic compounds

All samples were defatted by blending hexane and extracted according to Shao et al. [24] with minor modifications.

Briefly, the mixture was shaken at 250 rpm for 1 h at room temperature and centrifuged at 5000 g for 15 min. After the residue was dried at 25 °C in a fume hood for 12 h, the samples were mixed with 4 ml of 80% methanol. Then the mixture was centrifuged at 5000 g for 20 min. The supernatant was collected for measurement of total phenolic content (TPC), total flavonoid content (FLC) and antioxidant activity (AOA).

TPC determination

Total phenolic content in barley extract was determined according to Zhao et al. [24] with minor modifications. Briefly, barley extract of 0.25 mL was mixed with 1.25 mL of 0.2 mol/L Folin-Ciocalteu's phenol reagent. After 5 min, 2 mL of 7.5% Na₂CO₃ solution (*w/v*) and 5 mL of deionized water were added. The mixture was incubated at room temperature under dark condition for 1 h on a shaking table. Then the absorbance was measured at 760 nm and a standard curve of gallic acid solution was made. TPC was expressed as micrograms of gallic acid equivalents (GAE) per gram of barley flour (μg GAE/g).

FLC determination

Total flavonoid content was assayed according to the modified method of Shao et al. [25]. Briefly, 0.5 ml of barley extract was mixed with 3 ml ddH₂O. Thereafter 150 μL of 5% NaNO₂ was added and then incubated for 5 min. After adding 150 μL of 10% AlCl₃·6H₂O and incubating for another 5 min, 1 ml NaOH (1 M) was added and thoroughly mixed, and reacted for 15 min, and then the absorbance of solution was measured at 510 nm. FLC was expressed as micrograms of catechin equivalent (CAE) per g of barley flour (μg CAE/g).

AOA determination

Antioxidant activity of barley extract was evaluated according to the procedure described by Saint-Cricq de Gaulejac et al. [26] with some modifications. Briefly, 0.1 ml of barley extract was mixed with 2.9 mL of 6 × 10⁻⁵ M methanolic solution of 1, 1-Diphenyl-2-picrylhydrazyl Radical (DPPH). After 60 min under dark condition on a shaking table, the absorbance of solution was measured at 517 nm. Inhibition of free radical DPPH in percent (*I* %) was calculated by using the following equation: $I \% = [(A_0 - A_e) / A_0] \times 100$, where *A*₀ is the absorbance of the blank sample and *A*_e is the absorbance of the tested sample.

Population structure, kinship and linkage disequilibrium (LD) analysis

The DArT markers used were derived from Diversity Arrays Technology Pty Ltd. in Australia (http://www.triarte.com.au/content/barley_diversity_analysis.html) [27,

28]. Totally, 780 barley DArT markers (MAF > 0.03) (Additional file 2: Table S2), were used for population structure analysis using STRUCTURE software (v2.3.4) [29], setting the number of clusters (k) from 1 to 10 with 100,000 MCMC (Markov Chain Monte Carlo) and eight independent iterations were performed in an admixture model. The largest value of Δk was used as the indicator of the most probable number of clusters (k) [30]. Kinship (K) was estimated using SPAGeDi (version 1.3d) software [30]. We calculated genetic distance and developed an UGPGA tree with NTSYSpc (version 2.10e). Tassel 3.0 was used to calculate linkage disequilibrium (LD) according to parameter r^2 , which is a measurement of the correlation between a pair of variables [31]. The genetic distance was derived from LD decay distance in the whole genome, when $r^2 = 0.1$ using the fitted equation.

GWAS for TPC, FLC and AOA

Association analysis of TPC, FLC, AOA was performed by TASSEL version 3.0 (<http://www.maizegenetics.net>) using 780 DArT markers. In order to acquire positive results in GWAS, four different models were performed [32]. Q model was applied to reduce the confounding caused by the sub-population membership. Q matrix using STRUCTURE software, and this model was expressed as $y = X\beta + Qv + e$, where X is the DArT marker matrix, Q is the sub-population membership matrix, e is the random error vector, and β and v are coefficient vectors for the DArT marker and sub-population membership, respectively [29]. K model includes the kinship matrix which contributes to reducing the confounding associated with false positive results. The K model was expressed as $y = X\beta + Z\mu + e$, where Z is the kinship matrix and μ is a vector of random genetic effects [$\mu \sim N(0, 2K)$] [33]. The third approach was the Q + K model, including both sub-

population membership and kinship, $y = X\beta + Qv + Z\mu + e$ [34]. Different models were evaluated for the fitness and efficiency by the quantile-quantile (Q-Q) plot using TASSEL v3.0. In the significance test $-\log_{10}(p) > 2.0$ was used as the lowest threshold.

Putative functions of *HvUGT* (HORVU1Hr1G020560) for phenolic compounds

The nucleotide and amino acid sequence of *HvUGT* was obtained from IPK barley BLAST Server (http://webblast.ipk-gatersleben.de/barley_ibsc/) [35]. SMART (<http://smart.embl-heidelberg.de/>) was used to predicate functional domain of *HvUGT*. The alignment protein sequences of this gene with *Arabidopsis thaliana* were collected by the BLAST result of EnsemblePlants (<http://plants.ensembl.org>) and then performed by CLUSTALW with default options [36]. FastTree version 2.1.10 was used Maximum Likelihood (ML) method with 1000 replications [37].

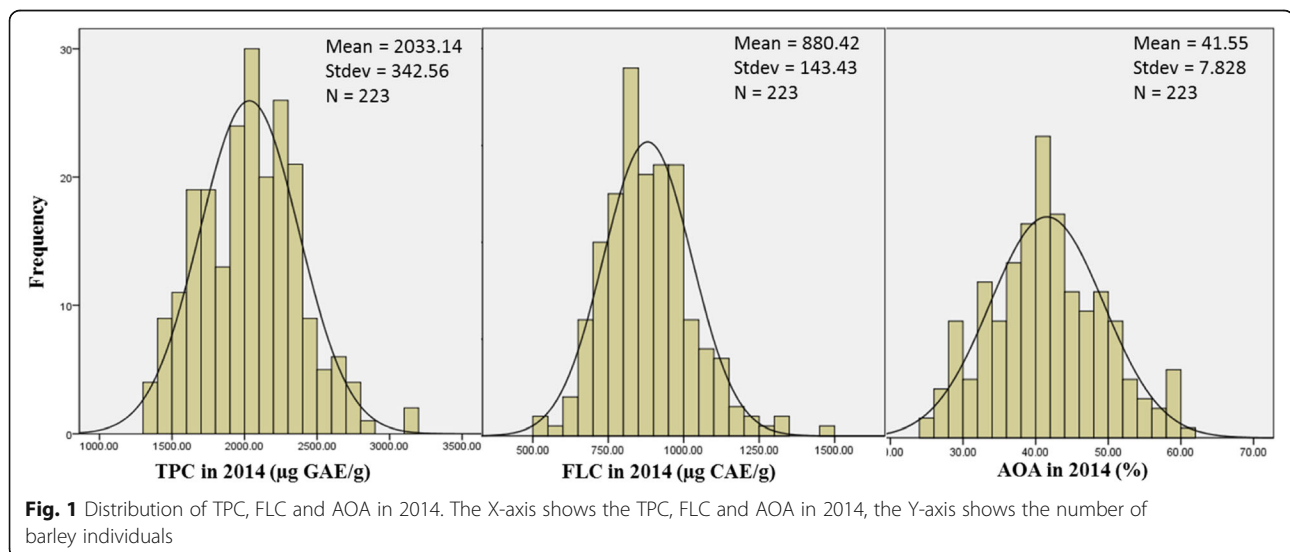
Statistical analysis

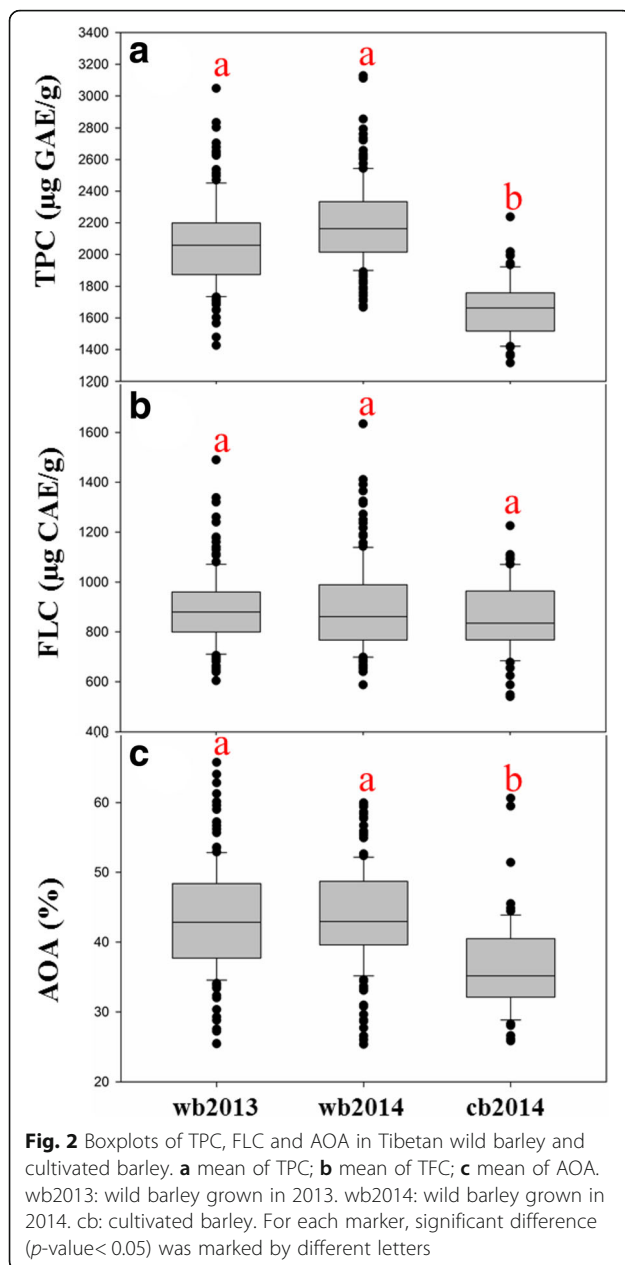
Data analysis for the distribution frequency of TPC, FLC and AOA were performed with the SPSS 20.0 (SPSS, Inc., Chicago, IL). Manhattan plot and Boxplot were made with the R version 3.4.0 (<http://www.r-project.org/>) and SigmaPlot version 12.5 (<http://www.sigmaplot.com/>) [38].

Results

Genotypic variations of TPC, FLC and AOA

There were large variations in the concentrations of TPC, FLC and AOA among the 223 barley genotypes in 2014, including 67 cultivated and 156 wild barley genotypes (Fig. 1). The mean value and variation of each parameter in wild barley had little difference in the two experimental years (2013 and 2014) (Fig. 2). There was





significantly positive correlation between the 2 years' data for each trait ($r = 0.81^{**}$, $r = 0.73^{**}$ and $r = 0.70^{**}$ for TPC, FLC and AOA, respectively), indicating that these traits might be mainly controlled by genetic factors (Additional file 3: Figure S1). In the year of 2014, compared with cultivated barley, wild barley had the much higher TPC and AOA concentrations, while the difference in FLC between wild and cultivated barleys was relatively small (Fig. 2).

Population structure

Linkage disequilibrium (LD) decay distance (r^2) was applied to determine the possibility if GWAS could be

used in association analysis of phenolic compounds and genetic markers. LD decay of genetic distance in the genome of all the 223 barley genotypes was 4.18 cM ($r^2 = 0.1$) (Additional file 4: Figure S2). Therefore, 801 DArT markers used in the present study were distributed over the whole genomic region with the average genetic distance of 0.70 cM and these markers were sufficient for genome wide association analysis.

The population structure was utilized in the model development in order to reduce nonfunctional and spurious associations caused by population stratification and an unequal distribution of alleles within these groups [39, 40]. The largest value of statistic index Δk was 64 when likelihood for sub-population (k values) calculated with STRUCTURE software were $k = 7$ (Additional file 5: Figure S3), indicating seven sub-populations ($k = 7$) were the most evident level of differentiation. These seven sub-populations consisted of 23, 31, 24, 8, 71, 21, 45 genotypes, respectively (Fig. 3; Additional file 6: Table S3). Interestingly, k7Q7 (the seventh population with k value = 7) consisted of most cultivated barleys and majority of wild barley was within other sub-populations, revealing the existence of considerable genetic diversity between Tibetan wild and cultivated barley. The result was consistent with the data from the cluster analysis (Additional file 7: Figure S4). The population structure of 223 barley genotypes was shown in Additional file 6: Table S3.

Association mapping of TPC, FLC and AOA

The contents of phenolic compounds in the wild and cultivated barleys were used to perform association mapping using 780 DArT markers. Three models, including Q model, K model and Q + K model, were tested to find the best fit model in association analysis, which was evaluated by Q-Q plot of P value distribution (Additional file 8: Figure S5). When Q or K model was applied, the observed P value was deviated from that expected, indicating the presence of abundant false positive results. Instead, the application of Q + K models reduced the number of false positive results. Thus, Q + K model was applied in this study. GWAS identified 13 unique loci for TPC, FLC and AOA in 223 barley genotypes in 2014 with the threshold $-\log_{10}(p) > 3$, and they were located on 6 of 7 chromosomes (except for 2H). Most of the loci were distributed on 3H, 5H and 6H, and each of them accounted for 4.0%–7.0% of phenotypic variation (Fig. 4; Additional file 9: Table S4).

Association analysis was also performed in subgroups. When the threshold of significant association was set at $p < 0.001$, 20 loci were detected in the 2 years' wild groups (Fig. 5; Additional file 9: Table S4). In the wild group, bPb-0572 (6H, 17.9 cM) was considered as the major locus controlling TPC ($-\log_{10}(p) = 5.61$ and 3.56 in 2 years), and contributed to 13.7 and 7.2% of

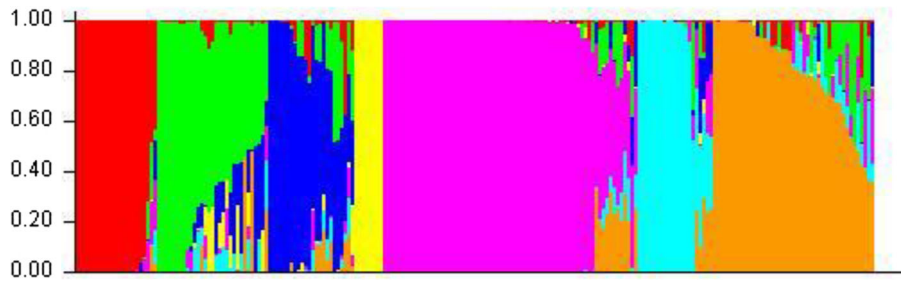


Fig. 3 Population structure of 223 barley genotypes. Population structure of all genotypes was divided based on genetic diversity detected by 801 DART markers with $k=7$. Seven subpopulations were represented by different colors

phenotypic variation, respectively. BPb-4531 on Chr. 1H explained 5.8% in 2013 and 10.7% in 2014 for the variation of TPC. Notably, bPb-4531 was closely associated with all three parameters in the wild genotypes in both years except for FLC in 2013. Furthermore, the locus was significantly associated with AOA ($-\log_{10}(p) = 4.72$ and 3.15 in 2 years, respectively), accounting for 9.8 and 7.1% of phenotypic variation. In the cultivated barley, 7 loci were associated with these three parameters with $-\log_{10}(p) > 2.5$. With the threshold of $-\log_{10}(p) > 3$, only one marker bPb-1068 on Chr.3H was detected for TPC with $-\log_{10}(p) = 3.61$ and explaining 25.7% of phenotypic variation. BPb-3227 was detected using 223 genotypes with the three models and 156 Tibetan wild accessions with Q + K model in the year of 2014. It may be assumed that this locus is powerful in genetic effect for

wild barley but not for cultivated barley, suggesting that Tibetan wild barley have wider genetic diversity than cultivated barley.

Haplotype analysis of bPb-0572 and bPb-4531 for TPC in cultivated and wild barleys

Haplotype analysis was performed in cultivated and Tibetan barleys for the markers bPb-0572 and bPb-4531, which showed the closest association with the examined traits (Figs. 5 and 6). For bPb-0572, the marker associated with TPC in wild barley, haplotypes ‘2013-0’ and ‘2014-0’ showed higher TPC than haplotypes ‘2013-1’ and ‘2014-1’, respectively. On the contrary, there was no distinct difference in TPC between haplotypes ‘c-0’ and ‘c-1’ in the cultivated barley. For another marker associated with TPC, bPb-4531, haplotypes ‘2013-0’, ‘2014-0’

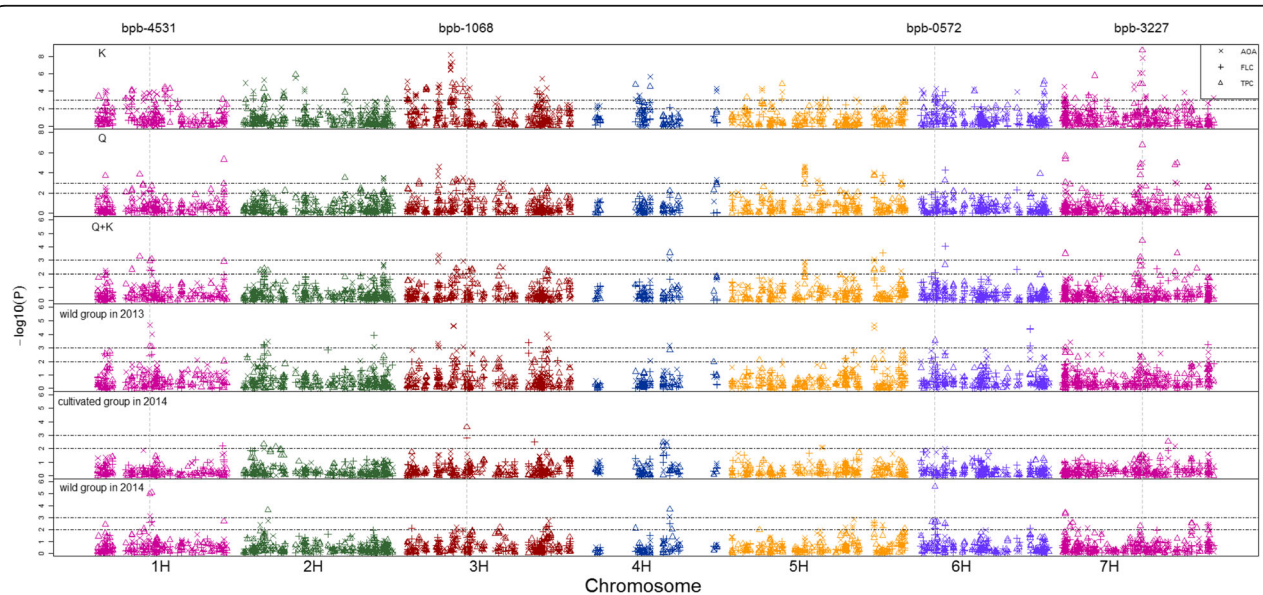
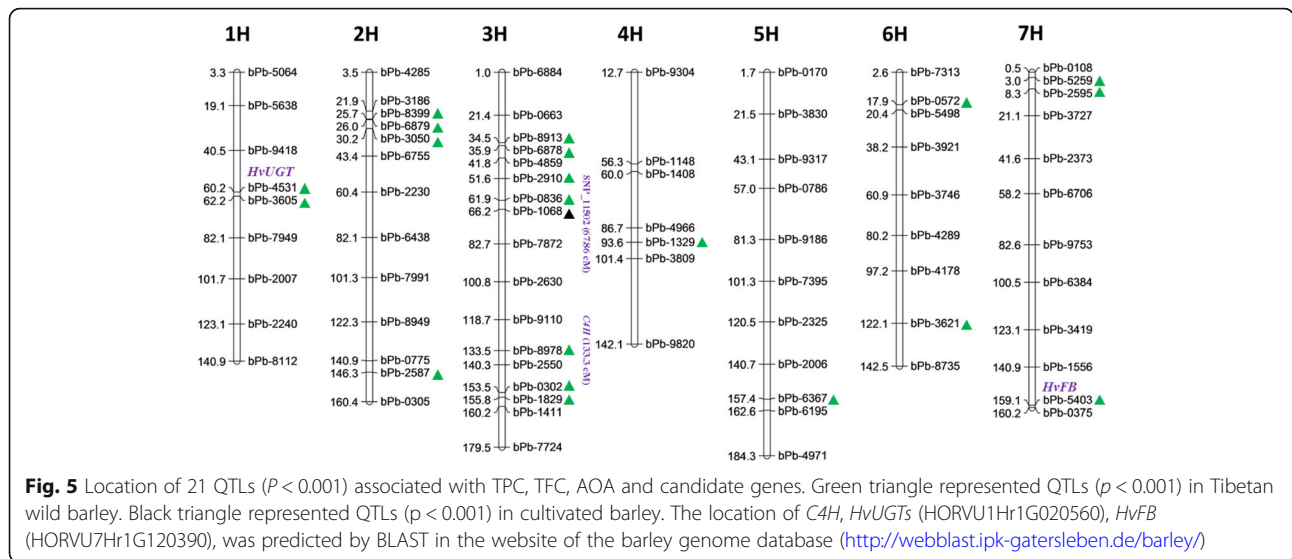


Fig. 4 GWAS of phenolic compounds within and across 67 cultivated and 156 Tibetan wild subgroups. Three grain parameters were applied to indicate content of phenolic compounds: TPC (Δ), FLC (+) and TPC (\times). GWA analysis was firstly conducted using three different methods: Q, K and Q + K methods. Then, Q + K method was selected to perform GWA analysis in cultivated and Tibetan wild subgroups individually. Dashed lines presented different p value of the four markers (bPb-4531, bPb-1068, bPb-0572 and bPb-3227) associated with these three parameters in wild and cultivated barley genotypes. Significant associations were identified using criterion of $-\log_{10}(P) > 2$ or 3



and ‘c-0’ showed higher TPC than haplotypes ‘2013-1’, ‘2014-1’ and ‘c-1’, respectively. However, in the cultivated groups, there was no significant difference.

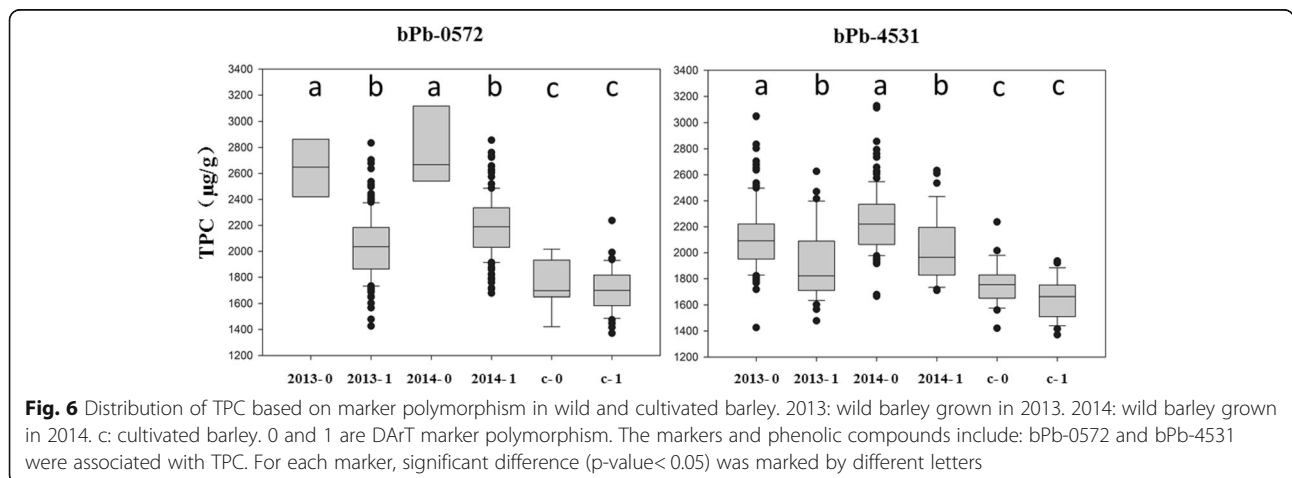
Candidate genes for phenolic compounds

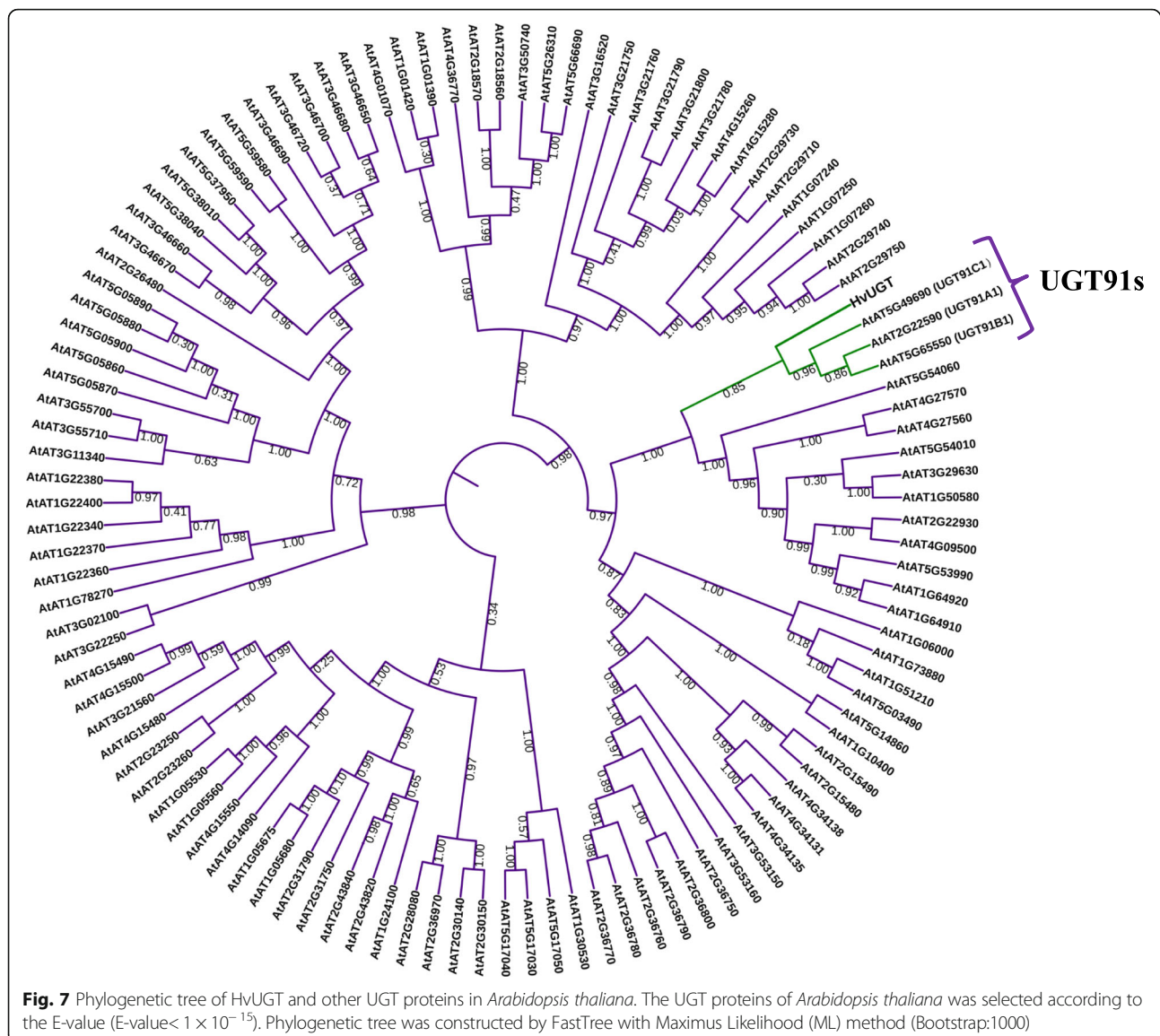
By BLAST analysis of the sequences of DArT markers using IPK barley Blast Server (<http://webblast.ipk-gatersleben.de/barley/>), we identified some candidate genes as shown in Fig. 5. Marker bPb-4531 was aligned to the location (Chromosome chr1H: 81,294,699-81,295,151) containing a gene (HORVU1Hr1G020560) encoding the UDP-Glycosyltransferase (UGTs) superfamily protein. Other genes were shown in Additional file 10: Table S5. Marker bPb-5403 was associated with the location (Chromosome chr7H: 651,255,911-651,256,221) containing a gene (HORVU7Hr1G120390) encoding F-box protein. In addition, the FLC-associated marker bPb-8978 (133.5 cM, 3H) was co-located with the gene encoding *C4H* (133.3 cM, 3H), which is the first Cyt-dependent

mono-oxygenase of phenylpropanoid pathway for flavonoid synthesis [41].

High similarity of HvUGT with UGT91s family in Arabidopsis thaliana

HvUGT protein sequence based on bPb-4531 was derived from IPK database (<http://webblast.ipk-gatersleben.de/barley/>) (Additional file 11: Figure S6). Then the amino acid sequence of HvUGT was analyzed by BLAST on genome of *Arabidopsis thaliana* (<http://www.arabidopsis.org/>) and 113 UGTs proteins ($E\text{-value} < 1 \times 10^{-15}$) were found (Additional file 12: Table S6). All these protein sequences were listed in Additional file 10: Figure S6. The phylogenetic analysis with *Arabidopsis thaliana*, showed that HvUGT had the most similarity with UGT91s family (UGT91A1, UGT91B1 and UGT91C1), which might include functional domain catalyzing glycosyl transfer to flavonoid glycosides [42] (Fig. 7). The function domain of HvUGT was also predicted by SMART (<http://>





smart.embl-heidelberg.de/), and the result showed that this gene consists of a similar UDPGT domain with UGT91s, indicating it is one of UDP-glucosyltransferase family members (Additional file 13: Figure S7).

Discussion

In this study, we determined TPC, FLC and AOA in grains of 156 Tibetan wild barley accessions and 68 cultivated barley genotypes. The results showed Tibetan wild barley has a wider variation in all three traits than cultivated barley. A previous study found that Tibetan wild barley had a significantly higher ferulic acid concentration than cultivated barley [43]. Total phenolics mainly consists of phenolic acids, flavonoids and anthocyanins. The higher TPC concentration in Tibetan wild barley may be partially attributed to more ferulic acid,

which is one of the major phenolic acids in barley grain. Flavonoids seem to have little contribution to the higher TPC in wild barley, as the difference in FLC between Tibetan wild and cultivated barleys was quite small. Currently, whether higher TPC in wild barley is also contributed by other phenolic compound is not clear. In this study, we found the high correlation between AOA and TPC ($r = 0.729^{**}$) and in fact AOA was an important component of phenolic compounds (Additional file 14: Table S7). In addition, Tibetan wild barley contained higher AOA than cultivated barley. Tibetan wild barley has the great potential for improving phenolic compounds of cultivated barley.

Currently, we identified 20 unique QTLs ($p < 0.001$) associated with TPC, FLC and AOA in Tibetan wild barley (Fig. 5). Most of the identified QTLs exist differently

in the wild and cultivated barleys, suggesting the distinct difference in genetic diversity between the two barley groups. Some QTLs associated significantly with TPC, such as bPb-0572 and bPb-4531 were identified in Tibetan wild barleys, but not in the cultivated barley suggesting some genes or alleles controlling phenolics remain in the Tibetan wild barley and have lost in the cultivated barley [44]. The Tibetan Plateau, so called “the roof of the world” because of its very high altitude, is well known for its extreme environment, such as low and variable temperature and strong UV irradiation [15, 16, 45]. According to previous studies, wild barley might develop unique mechanisms for adapting to such harsh environments [22, 46]. Under these abiotic stresses, excessive reactive oxygen species (ROS) could be formed in plant tissues, causing damage to plant cells [47]. On the other hand, the plants exposed to severe abiotic stress preferentially accumulate phenolic compounds in order to scavenge these ROS, and alleviate peroxidation of lipids [48, 49]. Thus Tibetan wild barley has developed its own mechanism of abiotic stress tolerance by producing more phenolics. In fact, phenolic compounds also play important roles in adjusting other plant growth and metabolisms, such as differentiation, pigmentation formation [50–52].

Although phenolic acids are components of phenolics and may partially contribute to TPC and AOA, only one common locus (bPb-0836) was found in this study (Fig. 5). A possible explanation would be that the individual phenolic acids only account for a small amount of total phenolics. Mohammadi et al. [53] reported three QTLs for total phenolics in barley genotypes collected from eight US breeding programs. One of them (SNP_11502, 3H, 67.86 cM) was located near the locus (bPb-1608, 3H, 66.16 cM), which was found in this study to be associated with TPC in the cultivated barley.

The release of barley genome sequence will facilitate the prediction of possible candidate genes based on the identified markers [35]. The UGT proteins are partly involved in the synthesis of anthocyanins and flavonoids [54]. Moreover, UGTs in plants participate in the formation of 3-O-glucoside and 4'-O-diglucoside, thus promoting flavonoid glycosylation [55]. In the current study, marker bPb-4531 was located on the Chr. 1H, which contains a gene encoding the UDP-glycosyltransferase (UGT) superfamily protein (HvUGT). According to the phylogenetic tree, *HvUGT* is closely related to gene family *UGT91s* in *Arabidopsis thaliana*. *UGT91A1*, one of the *UGT91* family members in *Arabidopsis thaliana* might be involved in the synthesis of flavonol glycosides [42, 56]. Without *UGT91A1* activity, there would be no galactose transferring into kaempferitrin, indicating that this protein specifically accepts and transfers galactose [57]. In fact this protein was found to have the similar specificity to

flavonoid as *UGT73C6* and *UGT78D1* did [58]. Thus, *HvUGT* may be involved in pathway of flavonoids. Markers bPb-5403 was associated with contig_41360 containing a gene (*HORVU7Hr1G120390*) encoding F-box protein. Kelch domain-containing F-box proteins (KFBs) negatively regulated naringenin chalcone accumulation, thus reducing the production of polyphenols [59, 60]. Kelch domain-containing F-box proteins (KFBs) have also been found to regulate the biosynthesis of phenylpropanoids, such as anthocyanins, flavonoids, phenolic ester and lignin [61]. In addition, the FLC-associated marker bPb-8978 (133.5 cM, 3H) was co-located with the gene encoding C4H (133.3 cM, 3H), which is the first Cyt-dependent mono-oxygenase of phenylpropanoid pathway for flavonoid synthesis, which have significant effect on lignin development [41]. In short, these identified genes are closely associated with phenolic metabolism, although their functions are still unclear in barley. Furthermore, the detected markers in this study should be helpful for better understanding the genetic control of TPC, FLC and AOA in barley.

Conclusions

The current results showed the wide variation among barley genotypes and obvious difference between Tibetan wild and cultivated barleys in grain TPC, FLC and AOA. Tibetan wild barley had higher concentration and wider genetic diversity of phenolic compounds than cultivated barley. Most QTLs were identified in the Tibetan wild barley and only one was detected in cultivated barley with $p < 0.001$, indicating Tibetan wild barley is potentially useful in barley breeding for improving phenolic compounds. The marker (bPb-4531) was co-located with *HvUGT*, which is a homolog to *Arabidopsis* UGT and may be responsible for flavonoid synthesis. This finding may serve as the foundation for further in-depth studies on molecular mechanism of natural variation in phenolic compounds.

Additional files

Additional file 1: Table S1. Cultivated and wild barley genotypes used in this study. C: cultivated barley; W: wild barley; XZ: Tibetan wild barley. (XLSX 21 kb)

Additional file 2: Table S2. MAF and PIC of 801 DArT markers. MAF, Minor allele frequency; PIC, polymorphic information content. (XLSX 40 kb)

Additional file 3: Figure S1. Correlation analysis of TPC, FLC and AOA in Tibetan wild barley. (a) correlation for TPC between 2013 and 2014; (b) correlation in FLC between 2013 and 2014; (c) correlation in AOA between 2013 and 2014. **, represents significant correlation at $P < 0.01$. (DOCX 80 kb)

Additional file 4: Figure S2. Decay of linkage disequilibrium of the population of 223 genotypes based on 801 DArT markers. The X-axis showed the genetic distance, the Y-axis showed the r^2 , the squared allele frequency correlations, which is a measurement of the correlation between a pair of variables. (DOCX 66 kb)

Additional file 5: Figure S3. Estimation of the most probable number of clusters (k). (DOCX 27 kb)

Additional file 6: Table S3. The value of population structure of 223 genotypes. Each genotype belongs to the population with the highest value calculated by STRUCTURE software. (XLSX 25 kb)

Additional file 7: Figure S4. Phylogenetic tree (UPGMA) of 223 barley genotypes based on 801 DArT markers. C, cultivated barley; X, Tibetan wild barley. (DOCX 831 kb)

Additional file 8: Figure S5. Quartile-quartile (Q-Q) plots of *P* value for 223 genotypes under three GWAS models. The black line is an expected one under the null distribution. Red symbol represents the observed *P* values for AOA, blue for FLC and green for TPC. (DOCX 425 kb)

Additional file 9: Table S4. List of DArT markers ($-\log_{10}(p) > 2.5$) in 223 genotypes and within subgroups. These markers were identified using Q + K model with significance threshold as $-\log_{10}(p) > 2.5$. R2 (marker) denotes the contribution of the marker for phenotypic variation. (DOCX 35 kb)

Additional file 10: Table S5. Genes around the region of marker bPb-4531. NOTE: Genomic location is showed in physical distance. (XLSX 12 kb)

Additional file 11: Figure S6. Protein sequences of HvUGT (HORVU1Hr1G020560) in barley and UGTs in *Arabidopsis thaliana*. (PDF 187 kb)

Additional file 12: Table S6. BLAST results of HvUGT with *Arabidopsis thaliana*. Note: These data are derived from The Arabidopsis Information Resource. (<http://plants.ensembl.org>). (XLSX 17 kb)

Additional file 13: Figure S7. Functional domain prediction of HvUGT. a: function domain in HvUGT; b: function domain in UGT91C1. Pink color represents low complexity sequence, the number marks amino acid sequence sites. (PPTX 90 kb)

Additional file 14: Table S7. Correlation analysis of TPC, FLC and AOA in 223 accessions. ** represent significant correlation at $P < 0.01$. (DOCX 15 kb)

Abbreviations

AOA: Antioxidant activity; C4H: Cinnamic acid 4-hydroxylase; CAE: Catechin equivalent; F3H: Flavanone 3-hydroxylase; FLC: Total flavonoid content; GAE: Gallic acid equivalent; GWAS: Genome-Wide Association Study; KFBs: Kelch domain-containing F-box proteins; TPC: Total phenolic content; UGT: UDP-Glycosyltransferase

Acknowledgements

Not applicable.

Funding

This study was supported by the National Natural Science Foundation of China (31620103912, 31471480, and 31401369), the China Agriculture Research System (CARS-05), the Jiangsu Collaborative Innovation Center for Modern Crop Production (JCIC-MCP), and the Fundamental Research Funds for the Central Universities. The funding bodies had no role in the design of the study and collection, analysis and interpretation of data and in writing the manuscript.

Availability of data and materials

The data supporting the results of this article are included within the article.

Author's contributions

GZ and ZH designed the experiments; ZH, JZ, XC and XQ did the experiments; ZH and SC analyzed data; ZH, SC and GZ wrote the paper. All authors read and approved the final manuscript.

Ethics approval

Tibetan wild barley used in this study were introduced from Huazhong Agricultural University (HZAU) and cultivated barley were collected from southern east China. All these samples were permitted to do scientific research. The experimental research on barley was complied with China's legislation and field studies was in accordance with legislation of Zhejiang University. All the experiments were not involved in any endangered or protected species.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Agronomy, Zhejiang Key Laboratory of Crop Germplasm, Zhejiang University, Hangzhou 310058, China. ²School of Biological Science and Technology, University of Jinan, Jinan 250022, China.

Received: 15 June 2017 Accepted: 16 January 2018

Published online: 25 January 2018

References

- Madhujith T, Izydorczyk M, Shahidi F. Antioxidant properties of pearled barley fractions. *J Agr Food Chem*. 2006;54(9):3283–9.
- Han Z, Cai S, Zhang X, Qian Q, Huang Y, Dai F, Zhang G. Development of predictive models for total phenolics and free p-coumaric acid contents in barley grain by near-infrared spectroscopy. *Food Chem*. 2017;227:342–8.
- Adom KK, Liu RH. Antioxidant activity of grains. *J Agr Food Chem*. 2002;50(21):6182–7.
- McMurrough I, Baert T. Identification of proanthocyanidins in beer and their direct measurement with a dual electrode electrochemical detector. *J I Brewing*. 1994;100(6):409–16.
- Bizjak J, Weber N, Mikulic-Petkovsek M, Alam Z, Thill J, Stich K, et al. Polyphenol gene expression and changes in anthocyanins and polyphenols in the skin of 'Braeburn' apples after the autumn application of prohexadione-calcium. *Plant Growth Regul*. 2013;71(3):225–33.
- Dvořáková M, Guido LF, Dostálek P, Skulilová Z, Moreira MM, Barros AA. Antioxidant properties of free, soluble Ester and Insoluble-bound Phenolic compounds in different barley varieties and corresponding malts. *J I Brewing*. 2008;114(1):27–33.
- Gupta M, Abu-Ghannam N, Gallagher E. Barley for brewing: characteristic changes during malting, brewing and applications of its by-products. *Compr Rev Food Sci F*. 2010;9(3):318–28.
- Cai SG, Han ZG, Huang YQ, Chen ZH, Zhang GP, Dai F. Genetic diversity of individual phenolic acids in barley and their correlation with barley malt quality. *J Agr Food Chem*. 2015;63(31):7051–7.
- Shao YF, Jin L, Zhang G, Lu Y, Shen Y, Bao JS. Association mapping of grain color, phenolic content, flavonoid content and antioxidant capacity in dehulled rice. *Theor Appl Genet*. 2011;122(5):1005–16.
- Rhodes DH, Hoffmann Jr L, Rooney WL, Ramu P, Morris GP, Kresovic T. Genome-wide association study of grain polyphenol concentrations in global sorghum [*Sorghum bicolor* (L.) Moench] germplasm. *J Agr Food Chem*. 2014;62(45):10916–27.
- Vogt T, Jones P. Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. *Trends Plant Sci*. 2000;5(9):380–6.
- McIntosh CA, Owens DK. Advances in flavonoid glycosyltransferase research: integrating recent findings with long-term citrus studies. *Phytochem Rev*. 2016;15(6):1075–91.
- Yonekura-Sakakibara K, Hanada K. An evolutionary view of functional diversity in family 1 glycosyltransferase. *Plant J*. 2011;66(1):182–93.
- Yin Q, Shen G, Di S, Fan C, Chang Z, Pang Y. Genome-wide identification and functional characterization of UDP-Glycosyltransferase genes involved in Flavonoid biosynthesis in *Glycine max*. *Plant Cell Physiol*. 2017;58(9):1558–72.
- Dai F, Nevo E, Wu DZ, Comadran J, Zhou MX, Qiu L, et al. Tibet is one of the centers of domestication of cultivated barley. *Proc Natl Acad Sci U S A*. 2012;109(42):16969–73.
- Dai F, Chen ZH, Wang XL, Li ZF, Jin GL, Wu DZ, et al. Transcriptome profiling reveals mosaic genomic origins of modern cultivated barley. *Proc Natl Acad Sci U S A*. 2014;111(37):13403–8.
- Cai SG, Yu G, Chen XH, Huang YC, Jiang XG, Zhang GP, et al. Grain protein content variation and its association analysis in barley. *BMC Plant Biol*. 2013;13(1):35.
- Koehler G, Wilson RC, Goodpaster JV, Sonstebay A, Lai XY, Witzmann FA, et al. Proteomic study of low-temperature responses in strawberry cultivars (*Fragaria x ananassa*) that differ in cold tolerance. *Plant Physiol*. 2012; 159(4):1787–805.

19. Diaz C, Saliba-Colombani V, Loudet O, Belluomo P, Moreau L, Daniel-Vedele F, et al. Leaf yellowing and anthocyanin accumulation are two genetically independent strategies in response to nitrogen limitation in *Arabidopsis thaliana*. *Plant Cell Physiol*. 2006;47(1):74–83.
20. Singh B, Kumar A, Malik AK. Flavonoids biosynthesis in plants and its further analysis by capillary electrophoresis. *Electrophoresis*. 2017;38(6):820–832.
21. Fabón G, Martínez-Abaigar J, Tomás R, Núñez-Olivera E. Effects of enhanced UV-B radiation on hydroxycinnamic acid derivatives extracted from different cell compartments in the aquatic liverwort *Junggermannia exsertifolia* subsp. *cordifolia*. *Physiol Plantarum*. 2010;140(3):269–79.
22. Nevo E. Evolution in action across phylogeny caused by microclimatic stresses at “evolution canyon”. *Theor Popul Biol*. 1997;52(3):231–43.
23. Bedada G, Westerbergh A, Müller T, Galkin E, Bdolach E, Moshelion M, et al. Transcriptome sequencing of two wild barley (*Hordeum Spontaneum* L.) ecotypes differentially adapted to drought stress reveals ecotype-specific transcripts. *BMC Genomics*. 2014;15(1):995.
24. Zhao HF, Dong JJ, Lu J, Chen J, Li Y, Shan LJ, et al. Effects of extraction solvent mixtures on antioxidant activity evaluation and their extraction capacity and selectivity for free phenolic compounds in barley (*Hordeum vulgare* L.). *J Agr Food Chem*. 2006;54(19):7277–86.
25. Shao YF, Tang FF, Huang Y, Xu FF, Chen YL, Tong C, et al. Analysis of genotype × environment interactions for polyphenols and antioxidant capacity of rice by association mapping. *J Agr Food Chem*. 2014;62(23):5361–8.
26. Saint-Cricq de Gaulejac N, Provost C, Vivas N. Comparative study of polyphenol scavenging activities assessed by different methods. *J Agr Food Chem*. 1999;47(2):425–31.
27. Wenzl P, Li H, Carling J, Zhou M, Raman H, Paul E, et al. A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. *BMC Genomics*. 2006;7(1):206.
28. Wenzl P, Carling J, Kudrna D, Jaccoud D, Huttner E, Kleinjohs A, et al. Diversity arrays technology (DArT) for whole-genome profiling of barley. *Proc Natl Acad Sci U S A*. 2004;101(26):9915–20.
29. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155(2):945–59.
30. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*. 2005;14(8):2611–20.
31. Hardy OJ, Vekemans X. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol*. 2002;2(4):618–20.
32. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*. 2007;23(19):2633–5.
33. Ritland K. Estimators for pairwise relatedness and individual inbreeding coefficients. *Genet Res*. 1996;67(02):175–85.
34. Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet*. 2006;38(2):203–8.
35. Mascher M, Gundlach H, Himmelbach A, Beier S, Twardziok SO, Wicker T, et al. A chromosome conformation capture ordered sequence of the barley genome. *Nature*. 2017;544(7651):427–33.
36. Yu J, Hu F, Dossa K, Wang Z, Ke T. Genome-wide analysis of UDP-glycosyltransferase super family in Brassica Rapa and Brassica Oleracea reveals its evolutionary history and functional characterization. *BMC Genomics*. 2017;18(1):474.
37. Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One*. 2010;5(3):e9490.
38. Hiersche M, Rühle F, Stoll M. PostGwas: advanced GWAS interpretation in R. *PLoS One*. 2013;8(8):e71775.
39. Knowler WC, Williams RC, Pettitt DJ, Steinberg AG. Gm3; 5, 13, 14 and type 2 diabetes mellitus: an association in American Indians with genetic admixture. *Am J Hum Genet*. 1988;43(4):520.
40. Sharbel TF, Haubold B, Mitchell-Olds T. Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization of Europe. *Mol Ecol*. 2000;9(12):2109–18.
41. Peukert M, Weise S, Röder MS, Matthies IE. Development of SNP markers for genes of the phenylpropanoid pathway and their association to kernel and malting traits in barley. *BMC Genet*. 2013;14(1):97.
42. Yonekura-Sakakibara K, Fukushima A, Nakabayashima R, Hanada K, Matsuda F, Sugawara S, et al. Two glycosyltransferases involved in anthocyanin modification delineated by transcriptome independent component analysis in *Arabidopsis thaliana*. *Plant J*. 2012;69(1):154–67.
43. Cai SG, Han ZG, Huang YQ, Hu HL, Dai F, Zhang GP. Identification of quantitative trait loci for the phenolic acid contents and their association with agronomic traits in Tibetan wild barley. *J Agr Food Chem*. 2016;64(4):980–7.
44. Purugganan MD, Fuller DQ. The nature of selection during plant domestication. *Nature*. 2009;457(7231):843–8.
45. Cai S, Wu D, Jabeen Z, Huang Y, Zhang G. Genome-wide association analysis of aluminum tolerance in cultivated and Tibetan wild barley. *PLoS One*. 2013;8(7):e69776.
46. Wu D, Qiu L, Xu L, Ye L, Chen M, Sun D, et al. Genetic variation of HvCBF genes and their association with salinity tolerance in Tibetan annual wild barley. *PLoS One*. 2011;6(7):e22938.
47. Mittler R, Vanderauwera S, Gollery M, et al. Reactive oxygen gene network of plants. *Trends Plant Sci*. 2004;9(10):490–8.
48. Agati G, Azzarello E, Pollastri S, Tattini M. Flavonoids as antioxidants in plants: location and functional significance. *Plant Sci*. 2012;196:67–76.
49. Nakabayashi R, Yonekura-Sakakibara K, Urano K, Suzuki M, Yamada Y, Nishizawa T, et al. Enhancement of oxidative and drought tolerance in *Arabidopsis* by overaccumulation of antioxidant flavonoids. *Plant J*. 2014;77(3):367–79.
50. Tohge T, de Souza LP, Fernie AR. Current understanding of the pathways of flavonoid biosynthesis in model and crop plants. *J Exp Bot*. 2017;68(15):4013–28.
51. Shao Y, Bao J. Polyphenols in whole rice grain: genetic diversity and health benefits. *Food Chem*. 2015;180:86–97.
52. Cheynier V, Comte G, Davies KM, Lattanzio V, Martens S. Plant phenolics: recent advances on their biosynthesis, genetics, and ecophysiology. *Plant Physiol Biochem*. 2013;72:1–20.
53. Mohammadi M, Endelman B, Nair S, Chao S, Jones SS, Muehlbauer GJ, et al. Association mapping of grain hardness, polyphenol oxidase, total phenolics, amylose content, and β -glucan in US barley breeding germplasm. *Mol Breed*. 2014;34(3):1229–43.
54. Bowles D, Isayenkova J, Lim EK, Poppenberger B. Glycosyltransferases: managers of small molecules. *Curr Opin Plant Biol*. 2005;8(3):254–63.
55. Yin R, Messner B, Faus-Kessler T, Hoffmann T, Schwab W, Hajirezaei MR. Feedback inhibition of the general phenylpropanoid and flavonol biosynthetic pathways upon a compromised flavonol-3-O-glycosylation. *J Exp Bot*. 2012;63(7):2465–78.
56. Stracke R, Ishihara H, Hupé G, Barsch A, Mehrtens F, Niehaus K, Weissshaar B. Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *Plant J*. 2007;50(4):660–77.
57. Mönchgesang S, Strehmel N, Schmidt S, Westphal L, Tarutti F, Müller E, et al. Natural variation of root exudates in *Arabidopsis thaliana*-linking metabolomic and genomic data. *Sci Rep*. 2016;6:29033.
58. Jones P, Messner B, Nakajima J, Schäffner AR, Saito K. UGT73C6 and UGT78D1, glycosyltransferases involved in flavonol glycoside biosynthesis in *Arabidopsis thaliana*. *J Biol Chem*. 2003;278(45):43910–8.
59. Zhang XB, Gou MY, Liu CJ. *Arabidopsis* Kelch repeat F-box proteins regulate phenylpropanoid biosynthesis via controlling the turnover of phenylalanine ammonia-lyase. *Plant Cell*. 2013;25(12):4994–5010.
60. Feder A, Burger J, Gao S, Lewinsohn E, Katzir N, Schaffer AA, et al. A Kelch domain-containing F-box coding gene negatively regulates flavonoid accumulation in *Cucumis melo* L. *Plant Physiol*. 2015; <https://doi.org/10.1104/pp.15.01008>.
61. ul Hassan MN, Zainal Z, Ismail I. Plant kelch containing F-box proteins: structure, evolution and functions. *RSC Adv*. 2015;5(53):42808–14.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

