BMC
Plant Biology

Construction of a high-density genetic map and QTLs mapping for sugars and acids in grape berries

Jie Chen[1,2†], Nian Wang[3†], Lin-Chuan Fang[3], Zhen-Chang Liang[1], Shao-Hua Li[1*] and Ben-Hong Wu[1*]

Abstract

Background: QTLs controlling individual sugars and acids (fructose, glucose, malic acid and tartaric acid) in grape berries have not yet been identified. The present study aimed to construct a high-density, high-quality genetic map of a winemaking grape cross with a complex parentage (*V. vinifera* × *V. amurensis*) × ((*V. labrusca* × *V. riparia*) × *V. vinifera*), using next-generation restriction site-associated DNA sequencing, and then to identify loci related to phenotypic variability over three years.

Results: In total, 1 826 SNP-based markers were developed. Of these, 621 markers were assembled into 19 linkage groups (LGs) for the maternal map, 696 for the paternal map, and 1 254 for the integrated map. Markers showed good linear agreement on most chromosomes between our genetic maps and the previously published *V. vinifera* reference sequence. However marker order was different in some chromosome regions, indicating both conservation and variation within the genome. Despite the identification of a range of QTLs controlling the traits of interest, these QTLs explained a relatively small percentage of the observed phenotypic variance. Although they exhibited a large degree of instability from year to year, QTLs were identified for all traits but tartaric acid and titratable acidity in the three years of the study; however only the QTLs for malic acid and β ratio (tartaric acid-to-malic acid ratio) were stable in two years. QTLs related to sugars were located within ten LGs (01, 02, 03, 04, 07, 09, 11, 14, 17, 18), and those related to acids within three LGs (06, 13, 18). Overlapping QTLs in LG14 were observed for fructose, glucose and total sugar. Malic acid, total acid and β ratio each had several QTLs in LG18, and malic acid also had a QTL in LG06. A set of 10 genes underlying these QTLs may be involved in determining the malic acid content of berries.

Conclusion: The genetic map constructed in this study is potentially a high-density, high-quality map, which could be used for QTL detection, genome comparison, and sequence assembly. It may also serve to broaden our understanding of the grape genome.

Keywords: Berry quality, Genetic map, Next-generation sequencing (NGS), QTL analysis, Quantitative trait loci, Restriction-site associated DNA (RAD), *Vitis*

Background

The organoleptic quality of table grapes and the flavor and stability of wine depend strongly on the types of sugars and acids, as well as the total sugar and acid concentration, in the grapes. Generally, fructose and glucose are predominant in berries at maturity, and sucrose is present in smaller quantities [1-3]. They have different levels of sweetness: if sucrose is rated 1, then fructose is 1.75 and glucose 0.75 [4-6]. The main organic acids in grape berries are tartaric and malic acids, which typically account for 90% of total acids [7-9]. Malic acid is involved in many processes that are essential for the health and sustainability of the vine, and tartaric acid plays an important role in maintaining the chemical stability and the color of the wine. Tartaric acid has a stronger acidic flavor than malic (pKa: 3.04 vs. 3.40), and is also more sour [10].

* Correspondence: shhli@ibcas.ac.cn; bhwu@ibcas.ac.cn
†Equal contributors
[1]Beijing Key Laboratory of Grape Science and Enology, and CAS Key Laboratory of Plant Resources, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, P. R. China
Full list of author information is available at the end of the article

BioMed Central

Chen *et al. BMC Plant Biology* (2015) 15:28

Page 2 of 14

Many studies have identified genomic loci that are linked to traits of interest in grapes. Modern strategies for the investigation of loci are based on the construction of genetic linkage maps, which was facilitated by the development of molecular markers. The first maps were constructed based mainly on RAPD [11] and AFLP [12] markers. Since then, a range of markers has been developed, and genetic maps of various grape cultivars and other *Vitis* species have been constructed [13-32]. One of these, a genetic map of a *V. vinifera* cross between Syrah and Pinot Noir, took into account most markers, including 483 SNP, 132 SSR and 379 AFLP markers [31]. Wang *et al.* [33] developed a genetic map with a total of 1 814 SNP markers. For a single SNP marker, the lowest integrity was ~85%. Of these 1 814 SNP markers, 1 545 were homozygous for one parent and heterozygous for the other (960 for lm×ll and 585 for nn×np), constituting 85.2% of all selected SNP markers. However, the other three types of markers that could be mapped on both female and male linkage maps amounted to 14.8% (ab×cd: 77, ef×eg: 171 and hk×hk: 21) [33]. Of these, 1 121 are on the female map, 759 are on the male map, and 1 646 are on the integrated map. This map was produced by combining next generation sequencing (NGS) and restriction-site associated DNA (RAD). Recently, Barba *et al.* [34] also used NGS to construct linkage maps for *V. rupestris* B38 and 'Chardonnay', with 1 146 and 1 215 SNPs each, covering 1 645 and 1 967 cM, respectively, and asserting that NGS was a powerful method for constructing a high-density, high-quality genetic map.

In grapes, quantitative trait loci (QTL) detection has mostly been used to investigate the genes related to resistance to diseases such as powdery and downy mildew and Pierce's disease [20,26,29,35-37], as well as pest resistance [19,20,38-41]. It has also been used to examine the genes related to a range of agronomic traits, e.g. berry size, seed number, mean and total seed fresh and dry weights, berry weight [14,17,20,27,39,42,43], inflorescence and flower morphology, number of inflorescences per shoot, flowering date [26], timing and duration of flowering and of veraison, veraison-ripening interval [14,44], architecture of the inflorescence [45], aroma profile [46], anthocyanin content [47], and number of clusters per vine [42]. In addition, the QTLs controlling sexual traits [26] and fertility [48] have been identified.

The genes controlling sugar and acid production in grapes are extremely complex, because of both the diverse chains of metabolic processes involved and the effect of environmental factors influencing these processes [49]. Viana *et al.* [50] have recently identified some QTLs involved in controlling soluble solid concentrations, pH, and titratable acidity in grape berries, but these explain a small amount of phenotypic variation in

these traits. To our knowledge, no QTLs controlling the production of individual sugars and acids in grape berries have yet been identified. Some analyses of QTLs controlling soluble solid concentrations, titratable acidity, pH and the production of individual sugars and acids have, however, been conducted for other fruit tree species, such as peach [51,52], apple [53-56], sour cherry [57] and melon [58].

The aim of this work was to investigate the genetic determination of soluble solid concentrations, titratable acidity, and individual sugars and acids in grape berries. A high-density genetic map was constructed for the population, as described in Wang *et al.* [33]. The map was used in combination with phenotypic data to identify marker-linked loci, after which we identified loci related to phenotypic variability observed over three years. This population was derived from the interspecies cross of cultivars 'Beihong' (BH) and 'E.S.7-11-49' (ES).

Methods
Plant material
The population, which comprised 1 200 individuals, was obtained by crossing BH (*Vitis vinifera* 'Muscat Hamburg' × *V. amurensis*) with ES ((Minnesota 78 (*V. labrusca* 'Beta' × Witt) × *V. riparia*) × *V. vinifera* 'Chenin Blanc') in 2007. We randomly selected 249 individuals for our experiment, and used these to construct the genetic map. Due to plant mortality, poor fruit setting, and environmental factors (e.g. rainfall, hail storms), the number of individuals bearing fruits varied from year to year. Vines were planted in 2008, without replicate, in the vineyard at the Institute of Botany, Chinese Academy of Sciences, Beijing (39°90' N 116°30' E). They were trained to fan-shaped trellises and had single trunks, which facilitated protection during winter. The vines were spaced 1.0 m apart within the row and 2.5 m apart between rows, and rows were north–south oriented. They were maintained under routine cultivation conditions, including irrigation, fertilization, soil management, pruning and disease control.

A random set of fruiting genotypes and the two parents were used in each of the three years of the study (2011–2013). In total, 241 genotypes were used in 2011, 225 in 2012, and 197 in 2013 for phenotypic measurement. Of these, 187 were common to all three years. Three replicates of one or two berry clusters were harvested from each genotype and parent at maturity. Maturity date was estimated primarily by assessing the physical properties of the berries, the ease of removal of berries from pedicels (without berry tissue shriveling because of loss of water), and the change of seed color from bright green to tan-brown [59]. Date of maturity was also estimated partly based on previous records. In addition, by the same person was responsible for berry

Chen *et al. BMC Plant Biology* (2015) 15:28

Page 3 of 14

harvesting for the duration of the study, to ensure consistency in the estimation of date of maturity. Maturity date ranged from 15 August to 15 September in 2011, and from 20 August to 20 September in both 2012 and 2013, depending on the genotype. Harvested clusters were placed in plastic bags on ice and transported immediately to the laboratory, which took ~10 min. This mode of transportation did not result in significant change in tartaric acid concentration relative to normal transportation.

Measurement of sugars and acids

Each replicate was pressed using a hand juicer to extract berry juice. Soluble solids concentration (SSC, °Brix) of the juice was measured with a digital hand-held refractometer (Atago, Tokyo, Japan). A 2 mL sample of juice was diluted to 10 mL with deionized water, and titratable acidity was measured by titration up to pH 8.2 with 0.1 mol·L^{-1} NaOH, and expressed as g·L^{-1} of tartaric acid.

The remaining juice was centrifuged at 5 000 g for 15 min. The supernatants were decanted, passed through a SEP-C18 cartridge (Superclean ENVI C18 SPE), and filtered through a 0.22 μm Sep-Pak filter. The sugar and acid concentrations of the filtered supernatants were measured using a Dionex P680 HPLC system (Dionex Corporation, CA, USA).

Fructose and glucose concentrations were measured using a Shodex RI-101 refractive index detector with a Waters Sugar-Pak I column (300 mm × 6.5 mmI.D., 10 μm particle size) and a guard column cartridge (Sugar-Pak I Guard-Pak Insert, 10 μm particle size). The reference cell was maintained at 40°C. The column was maintained at 90°C using a Dionex TCC-100 thermostated column compartment. Degassed, distilled, deionized water at a flow rate of 0.6 mL·min^{-1} was used as the mobile phase. The injection volume was 10 μL.

Malic and tartaric acid concentrations were measured using a Dionex UltiMate3000 detector, with a Dikma PLATISIL ODS column (250 mm × 4.6 mmI.D., 5 μm particle size) and a guard column cartridge (DikmaSpursil C18 Guard Cartridge 3μm, 10 mm × 2.1 mm). The column was maintained at 40°C. Samples were eluted with 0.02 mol·L^{-1} KH$_2$PO$_4$ solution with pH 2.4, at a flow rate of 0.8 mL·min^{-1}. Eluted compounds were detected using UV absorbance at 210 nm.

The Chromeleon chromatography data system was used to integrate peak areas according to external standard solution calibrations [60] (reagents from Sigma Chemical Co. Castle Hill, NSW, Australia). Sugar and acid concentrations were expressed in mg·mL^{-1} juice.

DNA extraction

Young leaves (the second and third leaf from the apex) were harvested from each genotype and the two parents

at the beginning of the vegetative period (late spring). The samples were immediately stored in liquid nitrogen and transferred to a freezer maintained at −80°C. Samples, weighing 0.5 g were ground in liquid nitrogen and genomic DNA was extracted using DNeasy plant mini prep kit (Qiagen). Briefly, 2 μg genomic DNA from each sample (249 F1 progeny and both parents) was treated with 20 units (U) MseI (New England Biolabs [NEB]) for 60 min at 37°C in a 50 μL reaction. A quick blunting kit (NEB) was used to convert 30 μL of the digested sample to 5'-phosphorylated, blunt-ended DNA in a 50 μL reaction mixture; the reaction was performed with 30 μL of digested sample, 5 μL 10× blunting buffer, 5 μL 1 mM dNTP mix, 2 μL blunting enzyme mix and 8 μL sterile dH$_2$O at room temperature for 30 min. A 3'-adenine overhang was added to the resulting samples in a 50 μL reaction with 32 μL blunt-ended DNA sample, 5 μL Klenow buffer (10×), 10 μL dATP (1 mM), 3 μL Klenow fragments (3'→5'exo-, 5 U·μL^{-1}) and sterile dH$_2$O to the final volume at 37°C for 1 h. Then 2 μL of 100 nM P1 and P2 adapter with a 3- to 5- bp plant-specific index (barcode) at the 5' end and a thymine overhang at the 3' end was added to each sample in a 50 μL reaction. A ligation reaction was carried out overnight at 16°C with T4 DNA ligase and 16 samples with different plant indices pooled into one. DNA fragments of 400–500 bp (including the ~120 bp adaptor) were separated on a 1.5% agarose gel and purified using a MiniElute gel extraction kit (Qiagen). Finally, all pooled samples were amplified with Phusion High-Fidelity PCR Master Mix (NEB) for 18 cycles in a 100 μL reaction including 20 μL Phusion master mix, 5 μL of 10 μM modified Solexa amplification primer mix (AP1 and AP1; 2006 Illumina, Inc., allright reserved) and sterile dH$_2$O to the final volume. The AP1 and AP2 primers contained Illumina paired end sequencing primer sites. DNA concentration was measured using a 2.0 fluorometer at BGI (Beijing Genomics Institute, China) [33].

High-throughput genotyping and map construction

High-density genetic maps for the two parents, BH and ES, were constructed using a slightly altered version of the method described by Wang *et al.* [33]. All experiments were performed at BGI. RAD-seq libraries for all 249 genotypes and the two parents were constructed according to Etter *et al.* (2011) [61], and sequenced using the Illumina HiSeq 2000 platform. The raw data produced were filtered to remove adaptors, indices and low-quality data (reads with > 15% of bases with quality score < 30). The cleaned data were analyzed using a standard RAD-seq analysis pipeline in the software package Stacks [62]. Genotypes for each plant in the population were assigned according to these results. Representative sequences for each SNP marker were obtained based on sequence clustering during

Chen *et al. BMC Plant Biology* (2015) 15:28

Page 4 of 14

the RAD-seq analysis pipeline. To manage the large quantity of data, a number of custom-programmed Perl scripts were also used to conduct the analysis.

To identify anchor markers for this study, we first identified a set of SNP markers, which we used to assign the 19 grapevine chromosomes to 19 linkage groups (LGs). This was done in two steps. Firstly, we marked the segregation patterns of all identified SNP markers as ab × cd, ef × eg, hk × hk, lm × ll, and nn × np. The first three of these pairs, which appeared in both parental linkage maps, were treated as candidate anchor markers. Secondly, because all alleles of each SNP marker had two nearly identical 100 bp sequences, the sequences from any allele could be taken as representative of the genotype of this SNP marker. These two representative sequences from the candidate anchor markers were aligned with the sequence of the 12× genomic assembly for *V. vinifera* PN40024, using local BLAST software with parameters set to −m 8 and −e 1E-5. The positions of each sequence for one SNP marker on the genome were identified based on their top hit. Three strict criteria were used to select anchor markers: 1) the marker had to show no significant segregation distortion among the 249 progeny genotypes in our population ($P < 0.001$); 2) both of the marker's end sequences had to align with the same chromosome position on the physical map for the reference PN40024 genome; and 3) the distance between the positions for the two end sequences on the reference genome had to fall between 200 and 500 bp (the expected size of the digested fragments was ~300−400 bp).

In constructing the map, the double pseudo-test cross strategy of Grattaglia and Sederoff [63] was applied, using JoinMap4.0 (Kyazma). After data had been imported, a cross pollination (CP) model was used for data mining. The ratio of marker segregation was calculated using Chi-squared tests. Firstly, markers that showed significantly distorted segregation ($P < 0.001$) were excluded from further analyses; secondly, marker order on each linkage group was optimized by excluding markers with $\chi^2 > 3.0$. The genotypes of 1 826 SNP markers were analyzed for linkage and recombination, using the Kosambi function to estimate genetic map distances. Logarithm of odds (LOD) score thresholds ≥ 7 was used to group the markers. After the LGs had been computed, their number was assigned according to the anchor markers mapped on them.

QTL analysis

All trait data were Box-Cox transformed to unskew their distributions, and the normality of the distributions was tested using the Shapiro-Wilks test. The detection of QTLs using both the transformed and the original data yielded similar results in terms of number, location and

contribution of QTLs, so the original data were henceforward used and reported.

QTLs for all traits in the population in the three separate years were analyzed for the parents only using the composite interval mapping (CIM) method in WinQTL Cartographer 2.5 [64,65]. CIM was used to scan the genetic map and estimate the likelihood of a QTL and its corresponding effect for every 1 cM. The forward regression algorithm was used to identify cofactors. A thousand permutations were performed using the CIM model within, and the thresholds for each environment were identified (almost all environments had thresholds at LOD ~3.0; $P ≤ 0.05$). The 1-LOD confidence interval within the CIM model corresponded to the 95% confidence interval calculated by WinQTL Cartographer 2.5 for each QTL. The results showed that when LOD values were 3–3.2, the error rate was 5%. Threshold LOD value was therefore set to 3 for all traits. QTLs with peaks close to 5 cM were merged into one QTL, and each significant QTL was characterized by its maximum LOD score, the percentage of variation it explained and its confidence intervals in cM, corresponding to the maximum LOD score withinone unit's width either side of the LOD peak.

Search for candidate genes

For each QTL, the search for candidate genes was conducted in the genomic region corresponding to the confidence interval determined on the consensus map. The scrutinized sequence was limited by the most proximal SNP markers that were present in both the reference genome and the consensus map. The genes were selected based on the information available for the annotated reference genome (Genoscope 12×) of the quasi-homozygous line 40024 derived from Pinot noir (http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/) [66]. They were classified according to their biological function as registered in the database. The genes catalogued as "unknown function" or equivalent were not considered in further analyses. In addition, a gene ontology (GO) enrichment analysis was performed, considering the genes identified in the physical genomic region that was associated with the confidence interval for each QTL. We also compared the frequency of each QTL vs. the complete reference genome, and searched for possible enrichment in gene functions. All enrichment analyses were done with the agriGO tool (http://bioinfo.cau.edu.cn/agriGO), using the options "singular enrichment analysis" and "complete GO". Significant GO terms ($P < 0.05$) were calculated using a hypergeometric distribution and the Yekutieli multi-test adjustment method [67].

Statistical analysis

Glucose-to-fructose ratio and β ratio (tartaric acid-to-malic acid ratio) were calculated, as these have

Chen *et al. BMC Plant Biology* (2015) 15:28

Page 5 of 14

been proposed as useful descriptors for evaluating the sugar and acid composition of grape berries [3,68]. For all further analyses, the means of the three replicates for each genotype and the parents were used.

All statistical analyses were performed using S-Plus (MathSoft Inc.). The frequency distribution of each trait was analyzed using the function 'hist', and the number of classes was determined using the Sturges method. Phenotypic correlations between traits within years and between years for each trait were calculated using the non-parametric Spearman correlation coefficient.

Results

Phenotypic characterization of parents and individuals
Averaged correlation coefficients between each pair of years were significant at $P < 0.001$ for almost all traits, ranging from 0.52 for the glucose-to-fructose ratio, to 0.74 for titratable acidity (Table 1).

Fructose, glucose, total sugar and SSC were positively correlated with each other. Fructose and glucose were strongly positively correlated, with a correlation coefficient of 0.93 ($P < 0.001$). The glucose-to-fructose ratio, however, was inconsistently correlated with fructose and glucose over the three years, and was not significantly correlated with total sugar, SSC or the acid-related traits. There were significant positive correlations between tartaric acid, malic acid, total acid and titratable acidity, from 0.36 between tartaric acid and malic acid to 0.88 between total acid and titratable acidity. The β ratio was significantly negatively correlated with malic acid and titratable acidity, but did not have consistent relationships

with tartaric or total acid. The sugar-related and acid-related traits were, in general, negatively correlated, but the sugar-related traits were weakly positively correlated with the β ratio.

The traits examined showed approximately the same phenotypic data distributions for all three years (Figures 1 and Additional file 1: Figure S1). All traits exhibited continuous variation, which is typical of quantitatively inherited traits. Transgressive segregation was apparent in fructose, glucose, total sugar, SSC, glucose-to-fructose ratio and β ratio traits. For these traits, fewer than 12% of the genotypes had higher phenotypic values than the high-value parent (indeed only one genotype exceeded the parents' phenotypic value in 2011), and fewer than 29% of genotypes had lower phenotypic values than the low-value parent. Transgressive segregation was more apparent in the tartaric acid, malic acid, total acid and titratable acidity traits; for these traits, 37–88% of genotypes exceeded the high-value parent's phenotypic value, and 25–58% of genotypes were below the low-value parent.

Construction of genetic maps
A total of 1 826 SNP-based markers were used to construct the genetic maps. The lowest integrity for a single SNP marker was ~83.0%. Of the 1 826 SNP markers, 1 515 were homozygous for one parent and heterozygous for the other (803 for lm × ll and 712 for nn × np), constituting 83.0% of all selected SNP markers. The remaining 17.0% constituted the other three types of markers that could be mapped on both female and male linkage maps (ab × cd: 1, ef × eg: 109 and hk × hk: 201).

Table 1 Phenotypic correlation coefficients between the traits of grape berries produced by crossing 'Beihong' with 'E.S.7-11-49'

	Fructose	Glucose	Total sugar	SSC	G/F	Tartaric	Malic	Total acid	TA	β ratio
Fructose	**0.56*****	0.93***	0.98***	0.86***	−0.27(ns)	−0.32(ns)	−0.52***	−0.49***	−0.57***	0.25(ns)
Glucose		**0.57*****	0.98***	0.86***	ns(+)	−0.32 (ns)	−0.48***	−0.47***	−0.55***	0.20 (ns)
Total sugar			**0.56*****	0.87***	ns	−0.32 (ns)	−0.52***	−0.49***	−0.57***	0.23 (ns)
SSC				**0.63*****	ns	−0.32***	−0.54***	−0.54***	−0.56***	0.20 (ns)
G/F					**0.53*****	ns	ns(+)	ns	ns	ns(−)
Tartaric						**0.63*****	0.36***	0.76***	0.59***	0.39(ns)
Malic							**0.71*****	0.88***	0.82***	−0.56***
Total acid								**0.68*****	0.88***	−0.27(ns)
TA									**0.74*****	−0.30***
β ratio										**0.64*****

Correlation coefficients were averaged over three years, and over 241 genotypes in 2011, 225 in 2012, and 197 in 2013 (except for TA in 2013, for which there were 189 genotypes). The averages of the correlation coefficients between each two-year combination (2011 and 2012, 2011 and 2013, 2012 and 2013) for each trait are shown in the diagonal. SSC is the soluble solids content, G/F is the glucose-to-fructose ratio, TA is titratable acidity, and β ratio is the tartaric acid-to-malic acid ratio.
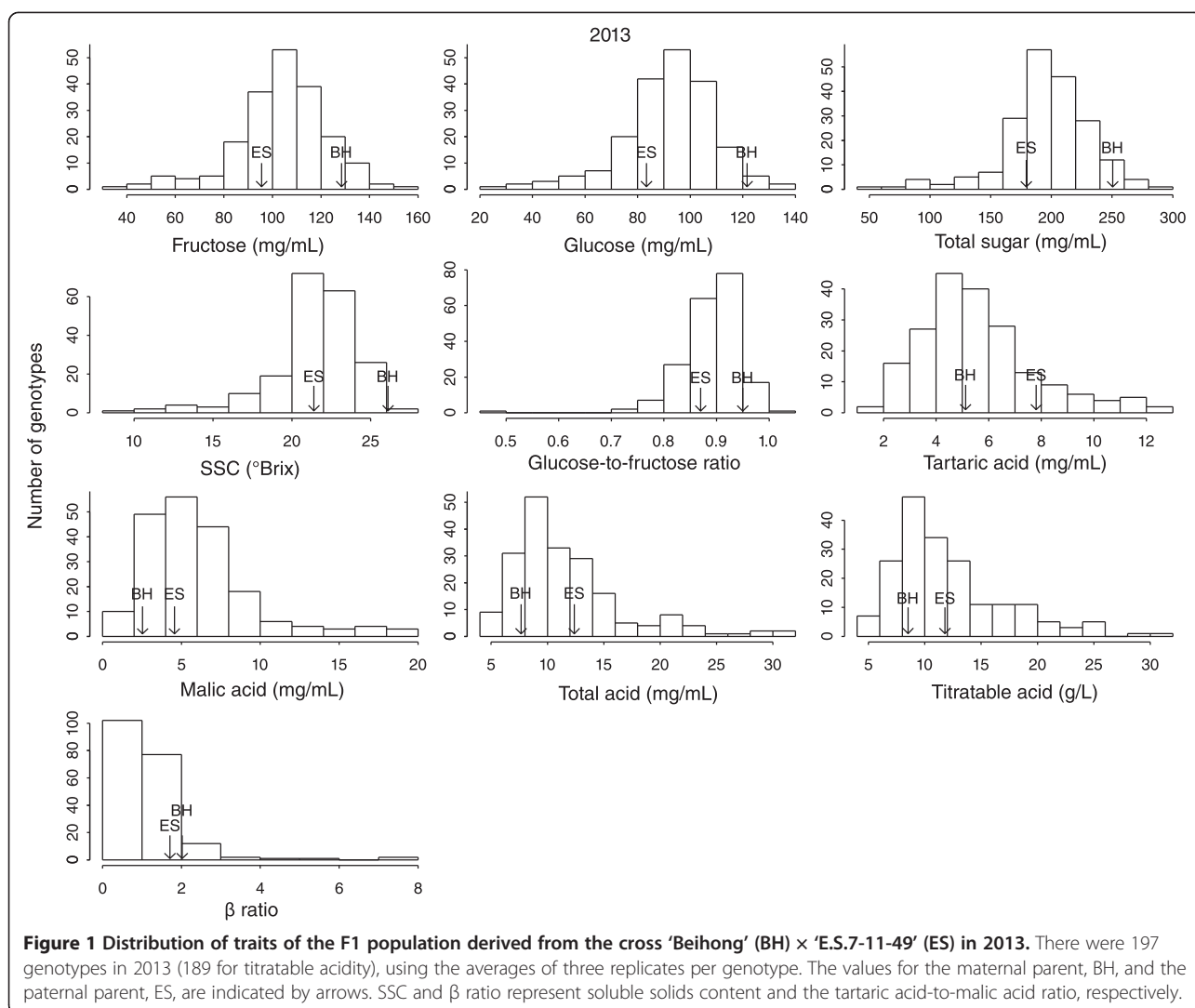***Significant at $P<0.001$ in all three years.
ns: not significant and/or significant at $P<0.05$ in all three years.
ns (+/−): significant (+ = positive, − = negative) only in one year at $P< 0.001$ or $P< 0.01$.
ns (+) in diagonal: significant only between 2011 and 2012.
(ns): not significant only in one year; the correlation coefficients significant at $P< 0.001$ or $P< 0.01$ for the other two years were averaged.

Chen *et al. BMC Plant Biology* (2015) 15:28

Page 6 of 14

Figure 1 Distribution of traits of the F1 population derived from the cross 'Beihong' (BH) × 'E.S.7-11-49' (ES) in 2013. There were 197 genotypes in 2013 (189 for titratable acidity), using the averages of three replicates per genotype. The values for the maternal parent, BH, and the paternal parent, ES, are indicated by arrows. SSC and β ratio represent soluble solids content and the tartaric acid-to-malic acid ratio, respectively.

The minimum number of reads for an SNP marker to be accepted was five per allele; 181 distorted markers were removed. For the BH map, 621 markers were assembled into 19 LGs spanning 1 553.43 cM of map distance, with an average interval length of 2.50 cM. The ES map was based on 696 markers positioned in 19 LGs, and covered 1 381.02 cM, with an average interval length of 1.98 cM (Table 2, Additional file 2: Figure S2, Additional file 3: Table S1). The integrated map of maternal and paternal LGs included 1 254 markers, unevenly distributed between LGs. The total number of markers per LG ranged from 11 (LG16) to 66 (LG18) for the BH map, and from seven (LG05) to 63 (LG07) for the ES map. Each 1 000 kb of DNA sequence occupied an average of ~3.68 cM on the BH map and ~3.27 cM on the ES map. The average interval between two adjacent mapped markers was estimated at ~679 kb (2.50/3.68 × 1 000) for the BH map, and ~606 kb (1.98/3.27 × 1 000) for the ES map.

Comparison of genetic and reference sequences

Of the 1 254 markers used in the integrated genetic map, 1 055 were on the physical map for the reference PN40024 genome (Table 2), which suggests our genetic maps cover 84.1% of the reference genome. Of the 621 markers on the BH map, 480 (77.3%) were common to both the genetic and physical maps, and of the 696 markers on the ES map, 625 (89.8%) were shared (Table 2). The physical size of the corresponding chromosomes ranged from 16.5 Mb (LG17) to 30.1 Mb (LG14). In individual LGs, the number of markers common to both the genetic and physical maps ranged from eight (LG16) to 54 (LG18) for BH, and from seven (LG05) to 52 (LG07) for ES. The positions of the common markers on the genetic maps were compared with their physical positions on the reference genome (Additional file 4: Figure S3, Additional file 3: Table S1). Most of the markers showed good linear agreement

Chen *et al. BMC Plant Biology* (2015) 15:28

Page 7 of 14

Table 2 Genetic map and number of common markers between genetic and physical maps for linkage groups

	Number of markers			Genetic size (cM)		Number of common markers		Chromosome size (Mb)
	BH	ES	Integrated	BH	ES	BH	ES	
LG01	25	43	66	81.688	78.187	23	41	22.8
LG02	49	30	75	106.868	67.662	37	25	18.6
LG03	16	30	45	96.402	69.187	15	27	19.2
LG04	46	41	83	106.634	104.325	40	39	23.7
LG05	39	7	46	79.219	22.831	13	7	24.8
LG06	30	47	75	79.630	71.633	27	45	21.3
LG07	13	63	73	25.035	98.122	9	52	20.9
LG08	53	41	88	94.558	73.718	41	41	22.3
LG09	24	33	54	86.867	71.318	23	32	23.0
LG10	45	20	60	81.571	79.890	23	10	17.7
LG11	24	40	60	79.699	74.507	23	40	19.3
LG12	34	53	75	66.618	52.297	29	47	22.3
LG13	33	54	81	79.181	82.256	30	41	24.4
LG14	28	38	65	83.639	90.139	26	36	30.1
LG15	29	22	49	87.052	58.515	14	21	20.1
LG16	11	19	27	50.544	56.048	8	18	21.7
LG17	27	34	60	63.685	58.969	26	34	16.5
LG18	66	52	117	92.401	101.306	54	44	29.3
LG19	29	29	54	112.139	70.110	19	25	23.8
Total	621	696	1254	1553.430	1381.020	480	625	421.8

The number of markers on the 19 linkage groups (LGs) of the 'Beihong' (BH) and 'E.S.7-11-49' (ES) genetic maps, their genetic sizes, and the number of markers common to both the genetic maps and the physical map for the reference PN40024 genome.

between the genetic and physical maps, with exceptions found on a few specific chromosomes (e.g. Chr05 and Chr16).

QTL identification

QTLs were analyzed separately on the parental maps for each of the three years (Table 3, Additional file 2: Figure S2). The CIM procedure detected 19 QTLs on the BH map, on LG02, LG03, LG06, LG09 and LG18, with 1, 1, 2, 1 and 14 QTLs, respectively. The average LOD value of the QTLs was 4.0, ranging from 3.0–8.1. On the ES map, 19 QTLs were detected on LG01, LG04, LG07, LG11, LG13, LG14, LG17 and LG18, with 1, 1, 1, 1, 1, 10, 3 and 1 QTLs, respectively. Here the average LOD value of the QTLs was 3.9, ranging from 3.0–6.1. The genomic threshold for both maps was ~3.0.

The number of QTLs identified for each trait varied between one and six, reflecting the quantitative nature of these traits, although no QTLs were detected for tartaric acid or titratable acidity. The QTLs that were identified were located within 13 of the 19 LGs. They each accounted for 5.28–17.31% of the total phenotypic variance in each trait.

Five QTLs for fructose were found in LG04, LG11, LG14, and LG17 of the ES map, each accounting for 5.58–9.71% of total variance. Three QTLs controlling glucose were found in LG14 of the ES map, contributing 6.04–8.16% of the variance. The QTLs for total sugar overlapped with those for fructose and/or glucose in LG14 of the ES map, individually contributing 5.85–8.37% of the variance. The QTL for SSC in LG14 of the ES map was the same as that for fructose, glucose and total sugar. A QTL for SSC was also identified in LG18 of the BH map, which explained 6.03% of the variance. A QTL for the glucose-to-fructose ratio and the α ratio, which was identified in LG03 and LG09 of the BH map, did not overlap with any of those for the individual sugars. Another QTL for the glucose-to-fructose ratio was found in LG02. QTLs for glucose-to-fructose ratio were also found in LG07 and LG17, on the ES map.

Malic acid, total acid and β ratio each had two to six QTLs in LG18 of the BH map, and there was another QTL for the β ratio in LG13 and LG18 of the ES map. There was also one QTL for both malic acid and total acid in LG06 of the BH map, which contributed 16.77–17.31% of the variance. However, no QTL could be identified for tartaric acid or titratable acidity.

Chen *et al. BMC Plant Biology* (2015) 15:28

Page 8 of 14

Table 3 Summary of QTLs in F1 population derived from the cross 'Beihong' (BH) × 'E.S.7-11-49' (ES)

Trait	Year	LG	Parent	Marker name	Peak location	LOD score	R^2 (%)	95% confidence interval left	95% confidence interval right
Fructose	2011	14	ES	186084	17.51	3.52	8.24	16.90	20.90
	2011	14	ES	66968	23.11	4.20	7.63	22.40	24.50
	2011	17	ES	275393	30.71	3.47	6.27	29.90	32.50
	2012	4	ES	200017	85.31	3.11	5.58	81.90	88.90
	2013	11	ES	201825	72.71	4.21	9.71	72.20	73.90
Glucose	2011	14	ES	17727	16.51	3.00	6.18	16.10	18.00
	2011	14	ES	66968	23.11	4.44	8.16	22.40	24.30
	2011	14	ES	149275	46.21	3.31	6.04	42.90	49.30
Total sugar	2011	14	ES	17727	16.51	3.26	6.72	16.20	18.00
	2011	14	ES	66968	23.11	4.52	8.37	22.40	24.30
	2011	14	ES	149275	46.21	3.19	5.85	43.00	49.10
SSC	2011	14	ES	66968	23.11	6.13	11.42	22.80	25.70
	2011	14	ES	167058	29.81	3.58	7.17	28.10	31.90
	2012	18	BH	120619	47.41	3.08	6.03	45.40	47.80
	2013	1	ES	32402	61.01	3.58	6.96	54.60	65.00
Glucose-to-fructose ratio	2011	2	BH	125072	103.91	3.08	6.33	102.90	104.90
	2011	3	BH	182932	28.41	3.92	6.77	17.00	29.90
	2012	9	BH	117666	40.41	3.18	5.71	39.60	44.70
	2013	7	ES	296046	48.41	4.58	10.87	47.60	52.30
	2013	17	ES	250686	29.91	3.01	5.94	29.10	30.80
	2013	17	ES	76880	41.41	4.91	9.92	40.90	41.60
Malic acid	2011	6	BH	248487	74.11	3.68	17.31	72.00	76.90
	2011	18	BH	215808	24.31	5.45	9.49	23.90	26.40
	2011	18	BH	120619	47.41	5.07	9.63	45.50	47.60
	2012	18	BH	280521	23.21	3.68	7.13	22.80	26.80
	2012	18	BH	120619	47.41	6.09	11.83	45.60	47.60
	2012	18	BH	15694	53.51	4.78	8.49	53.20	55.40
Total acid	2011	6	BH	248487	74.11	3.20	16.77	72.10	75.30
	2012	13	ES	52766	54.81	3.31	6.25	51.20	54.90
	2012	18	BH	233088	38.21	3.00	16.05	37.10	39.40
	2012	18	BH	120619	46.91	3.75	7.90	44.90	47.40
	2012	18	ES	254573	34.91	3.29	6.04	32.70	37.60
β ratio	2011	18	BH	166745	23.61	8.10	16.06	23.20	24.40
	2011	18	BH	120619	47.41	4.20	7.91	44.90	47.60
	2011	18	BH	15694	53.51	3.03	5.29	53.20	55.80
	2012	18	BH	215808	24.31	3.10	5.69	23.90	27.10
	2012	18	BH	120619	47.41	4.22	8.34	45.10	47.60
	2012	18	BH	15694	53.51	3.23	6.09	53.20	55.40

Locations on linkage groups (LGs) of the BH and ES genetic maps, and contributions of the putative QTLs that control sugar- and acid-related traits, which were identified in at least two of three successive years (2011, 2012 and 2013). The locus is the marker showing the strongest association with the trait. The location of markers is given in cM, quoted from the top of each linkage group. R^2 represents the individual contribution of one QTL to the variation in a trait, and LOD is the logarithm of the odds ratio. SSC and β ratio represent soluble solids content and the tartaric acid-to-malic acid ratio, respectively. Traits in bold had QTLs detected in two years.

Chen *et al. BMC Plant Biology* (2015) 15:28

Page 9 of 14

Candidate gene identification

In total, we identified 499 genes underlying the 19 QTLs of the BH map, and 724 genes underlying the 19 QTLs of the ES map. Of these, 835 (68.3%) were annotated and classified. However, only two QTLs (for malic acid, total acid and β ratio on LG18 of BH map) were stable across years (having been observed in two years). We therefore henceforward focused only on the candidate genes located within the confidence intervals of these two QTLs. For these two QTLs, 134 candidate genes were found. They were unevenly distributed, with 106 (22.8–26.8 cM) for one and 28 (45.5–47.6 cM) for the other QTL. Fifty of these genes were catalogued as having an "unknown protein function", and the others were classified into six major groups, namely cell, glycolysis, protein, RNA, TCA/org transformation, and transport. Of the 134 candidate genes, 10 that were probably related to TCA, acid metabolism or transport were listed, mainly (but not exclusively) based on their biological function as described in model plant species such as *Arabidopsis*, rice and poplar (Table 4).

Discussion

Phenotypic evaluation

The grape berries we analyzed displayed similar substantial variation in sugar and acid concentration across three successive years, which supports previous results showing that sugar and acid concentrations of grapes vary significantly by year [1]. For the study period, fructose, glucose, total sugar, tartaric acid, malic acid, and total acid concentration ranges were 7.5–136.7, 7.8–154.4, 15.3–291.9, 1.5–17.2, 0.8–21.3, and 4.9–34.7 mg·mL^{-1}, respectively. These ranges were greater than those found in other *Vitis* populations [49,69]. The reported ranges for fructose, glucose, and total sugar concentrations for these populations are 36.2–111.9, 38.5–104.4, and 78.9–216.3 mg·mL^{-1}, respectively, and those for tartaric acid, malic acid, and total acid

concentrations are 1.1–6.0, 0.6–8.3 and 2.1–11.8 mg·mL^{-1}, respectively. Phenotypic correlations between sugar and acid concentrations were also relatively stable across the three years, although these may be affected by environmental factors.

Genetic map

Although genetic maps for grape cultivars have developed greatly in recent years, the number of markers in the LGs in existing maps is still generally less than 1 000, and some of the mapped markers have no sequence information. We recently identified 1 814 high-quality SNP markers for a population of 'Z180' (1 212 markers) × 'Beihong' (759 markers) [33]. In this study we used the same procedure to construct the genetic map, and the density of the resultant linkage map was similarly high. In total we identified 1 826 SNP markers, 621 of which were mapped on the female BH genetic map, and 696 on the male ES map. The difference between the number of markers we identified in this study and in the earlier one may be related to the different F1 population. On the BH map, the average size of LGs was 81.76 cM, ranging from 25.04 cM (LG07) to 112.14 cM (LG19). On the ES map, the average size was 72.69 cM, ranging from 22.83 cM (LG05) to 104.33 cM (LG04). There were 17 and 12 marker-free regions longer than 10 cM on the BH map (LG02, 03, 04, 05, 09, 11, 14, 16, 19) and the ES map (LG02, 04, 06, 09, 10, 11, 14, 15, 18), respectively.

The total physical size of the grape genome is ~470 Mb [66,70]. In most regions of the parental genetic and physical maps (for *V. vinifera*), the markers occurred in the same order, but not in all the chromosome regions. This indicates that on one hand the genome is well conserved among grape species, but that some changes in marker order have occurred during speciation. In this study, the maternal parent, BH, is bred from *V. vinifera* and *V. amurensis*, and the paternal parent, ES, is bred from *V. labrusca* × *V. riparia* and *V. vinifera*. Differences

Table 4 Genes in LG18 that may participate in acid regulation

Groups	Position	Gene ID	Gene symbol	Description	References
Cell	5462823-5465587	GSVIVT01009139001	CYCD4;1	CYCLIN D4;1	[87]
Glycolysis	5505643-5516683	GSVIVT01009147001	PGI,PGI1	phosphoglucose isomerase 1	[88]
protein	6442873-6452788	GSVIVT01009228001	ATCIPK8,CIPK8,PKS11,SnRK3.13	CBL-interacting protein kinase 8	[89]
protein	6652256-6660443	GSVIVT01009251001	ATDBR1,DBR1	debranching enzyme 1	[90]
RNA	6526891-6532321	GSVIVT01009238001	IAA9	indole-3-acetic acid inducible 9	[91,92]
TCA/org transformation	5663288-5667315	GSVIVT01009165001	ATBCA5,BCA5	beta carbonic anhydrase 5	[93]
transport	6771378-6774993	GSVIVT01009260001	AAP6	amino acid permease 6	[94]
transport	5605382-5606712	GSVIVT01009152001	ATPUMP5,DIC1,UCP5	uncoupling protein 5	[86]
transport	6688299-6690573	GSVIVT01009253001	ZF14	MATE efflux family protein	[80-85]
transport	10039446-10043834	GSVIVT01009629001		MATE efflux family protein	[80-85]

Chen *et al. BMC Plant Biology* (2015) 15:28

Page 10 of 14

in the order of markers on some chromosomes have presumably resulted from different micro-structures on chromosomes in the various species. Alternatively, these differences might have arisen because of possible errors in mapping causing small inversions in marker order.

QTL detection

QTLs were analyzed separately for each of the traits on the parental maps for each of the three years, but were inconsistently detected. Some minor QTLs were detected only in a single year, such as a glucose QTL in LG14, for which R^2 = 6.04–6.18%, and a total sugar QTL in LG14, with R^2 = 5.85–6.72%. Other QTLs that contributed strongly to total variance were also detected in only one year, e.g. an SSC QTL in LG14 in 2011, for which R^2 = 11.42%, and a malic acid QTL in LG06 in 2011, with R^2 = 17.31%. In some cases, no QTLs were detected for a trait in a specific year, e.g. malic acid and total acid in 2013. Similar instability of QTLs across years has been widely reported for grapes [14,42,71], and also for other fruit tree species [52,57,72-74]. In contrast to crops such as maize, soya and rice, in which there may be many plants and biological replications of each genotype in one growth environment, there was only one vine per genotype in our trial. Phenotypic value assessment is potentially subject to bias which would increase the likelihood of error and affect the QTL analysis, with possible results including underestimated LOD values and overlap between QTLs across years. Furthermore, variation in climatic factors (such as rainfall and temperature) between years could bias assessment of fruit maturity, which would affect phenotypic evaluations. The observed low repeatability of QTL detection may thus have been exacerbated by the lack of replicate vines and the potential inconsistency in assessment of maturity. Addressing these problems in QTL studies on fruit species is difficult, however.

The percentage of variation explained by each QTL was small, and varied between 5.29% and 17.31%. This is consistent with the low R^2 values previously reported for some grape agronomic traits. Fanizza *et al.* [42] found that a QTL controlling berry weight had R^2 = 19%, but QTLs controlling the number of clusters per vine, cluster weight, number of berries per cluster, and berry weight had substantially lower R^2 values (1.2–10%). Similarly, Viana *et al.* [50] found that most QTLs accounted for less than 5.5% of the variance. QTLs with high R^2 values have generally been found to be related to properties such as veraison time/period, anthocyanin content (up to 48–62%) [47], and seed dry/fresh weight (up to 91.4%) [14,71]. It seems that agronomic traits, including sugar and acid concentration, are generally controlled by numerous QTLs, each with small effects. This might be due to the quantitative nature of these traits,

as well as complicated metabolic pathways and regulatory networks.

Heritability for most traits is generally less than 50%, so the heritability associated with each QTL is a small fraction of this [75]. The more QTLs there are in the population, the smaller their individual contribution and the more difficult they are to detect [75]. As a result, precise map construction may be challenging, and maps may include some QTLs with very small R^2 values. Furthermore, the number of QTLs detected and the phenotypic variance they explain might be biased because of the limitations of the experiment itself, such as small sample size (as the effectiveness of marker loci increases with the number of individuals in a population) [76].

To our knowledge, no QTLs controlling the production of individual sugars and acids in grape berries have previously been identified. Viana *et al.* [50] reported one QTL in LG03 for SSC, one QTL in each of LG06, 13 and 19 for titratable acidity (% tartaric acid), and one QTL in each of LG01, 06, 11, 13 and 16 for pH, based on results for one year. We did not detect a QTL in LG03 for SSC in any of the three years of our study, or any QTLs for titratable acidity. This discrepancy between results may result from different genetic determinants of trait variation in the populations studied. Another cause might be differences in sampling strategies and the methods used for measuring traits. In this study, a QTL for malic acid in LG06 positioned at 74.11 cM explained a relatively large amount of variance (17.31%). Viana *et al.* [50] reported a QTL in LG06 positioned at 0.00 cM for pH, which explained 10.34% of variance. Although they were not the same QTL, these two regions might be worth exploring for genes controlling the quality of fruit acidity.

QTLs co-location

For breeding purposes, it is worth examining QTLs that are co-located. With respect to individual sugars, a QTL in LG14 affected both fructose and glucose, which explained the high correlation between them (r = 0.93). However, this QTL was detected only in 2011. In peaches, three QTLs, located in three different LGs are related to both glucose and fructose concentration [52]. The co-location of QTLs controlling fructose and glucose probably indicates a unique gene with a pleiotropic effect, or genes with close linkage, because glucose and fructose are absent from phloem sap and in grape berries are synthesized concurrently by sucrose hydrolysis [77]. The QTL in LG14 is potentially promising to work with to increase sugar concentrations, which may be beneficial for wine-making. Further study on candidate functional genes within the confidence intervals of this QTL may help to assess the mechanism for controlling hexose metabolism. The QTL for total sugar in LG14

Chen *et al. BMC Plant Biology* (2015) 15:28

Page 11 of 14

overlapped with the QTL for fructose and glucose, which is probably related to the fact that these individual sugars contribute to total sugar. Similarly, a QTL for SSC in LG14 overlapped with one for fructose and glucose. However, the fact that the QTLs for the glucose-to-fructose ratio were not co-located with QTLs for either glucose or fructose is difficult to explain. With respect to the acid-related traits, QTLs for total acid and β ratio were co-located with those for malic acid, suggesting that malic acid contributes greatly to total acid and β ratio.

An important problem encountered when breeding for improved traits is negative correlations between favorable traits. For instance, as our results confirmed, malic acid and fructose are negatively correlated in 98 grape cultivars [1]. In some cases, this was caused by co-located QTLs with opposite horticultural effects. In tomatoes, fruit size and soluble sugar concentration are often negatively correlated, and the QTLs controlling them are located in the same LG [78,79]. In this study, although fructose, glucose, total sugar and SSC were negatively correlated with tartaric acid, malic acid, total acid and titratable acidity, in six LGs (01, 04, 11, 14, 17, 18) there were 15 QTLs related to sugars, and in three LGs (06, 13, 18) there were 11 QTLs related to acids. Viana *et al.* [50] also reported different LGs for sugar- and acid-related traits: one QTL in LG3 for SSC, and one QTL for pH and titratable acidity in each of LG01, 06, 11, 13, 16 and 19. The non-co-location of QTLs controlling sugar and acid concentration is favorable from the perspective of breeding. However, it is likely that some QTLs that have not yet been detected contribute to the remaining unexplained variance, and these should not be ignored.

Candidate genes for acid regulation

Of 134 genes located within the confidence intervals of the two QTLs controlling malic acid, total acid, and β ratio in LG18, 11 were probably involved in acid metabolism [80-94]. For example, beta carbonic anhydrase 5 is related to TCA/org transformation, and malic acid is involved in this process. Numerous studies reveal that the MATE family plays a key role in malate efflux from root apices in many plant species, including *Arabidopsis*, maize (*Zea mays*), wheat, rice (*Oryza sativa*), and rice bean (*Vigna umbellata*) [80-85]. Uncoupling protein 5 (DIC1) belongs to the mitochondrial carrier protein family. The *Arabidopsis* DIC proteins transport a wide range of dicarboxylic acids including malate, oxaloacetate and succinate [86], and these proteins might function as malate/oxaloacetate shuttles, which provide other cell components with reducing equivalents [86]. Expression levels of or changes in these candidate genes could be evaluated in further transcriptomic and gene-directed

studies. The identification of the most relevant genes would help to reveal the molecular mechanisms operating in grape cultivars and could have a large impact on future breeding efforts.

Conclusions

The genetic map we have constructed for a winemaking grape cross is potentially a high-density, high-quality map, which could be used for QTL detection, genome comparison, and sequence assembly. In total, we mapped 1 254 markers with 60 bp sequences on the integrated map. These markers can be used as anchors to compare genetic and physical maps. This may facilitate the improved use of grape genomic resources. Moreover, this genetic map of a cross with a complex parentage, (*V. vinifera* × *V. amurensis*) × ((*V. labrusca* × *V. riparia*) × *V. vinifera*), will help to broaden our understanding of the grape genome. However, the stability and accuracy of QTLs are also affected by environmental factors. In our study, year and climatic conditions were the most important source of variability. Thus, controlling environmental factors can increase the likelihood that an observed phenotypic value is both accurate and repeatable.

Several QTLs controlling berry sugar and acid traits were detected in different LGs, suggesting that these traits are influenced by several genes that control different aspects of complex metabolic pathways. For example, we have identified a set of 10 candidate genes underlying the QTLs that are potentially related to malic acid. We anticipate that we will soon be able to narrow down these regions to the point where effects can be ascribed to specific genes. In addition, studies based on transcriptomics, proteomics, and metabolomics would help us to achieve a more accurate understanding of the molecular parameters involved in berry sugar and acid regulation. The long-term objective of this research is to provide information on the genetic basis of these traits, and to facilitate the selection of varieties to improve sugar and acid quality.

Additional files

Additional file 1: Figure S1. Distribution of traits of F1 population derived from the cross 'Beihong' (BH) × 'E.S.7-11-49' (ES) in 2011 and 2012.

Additional file 2: Figure S2. Genetic maps and QTL locations for 'Beihong' (BH, maternal parent) and 'E.S.7-11-49' (ES, paternal parent).

Additional file 3: Table S1. SNP-based markers and genetic positions on the 19 linkage groups (LGs).

Additional file 4: Figure S3. Positions of SNP-based markers on genetic maps, and their physical positions on the reference genome.

Competing interests
The authors declare that they have no competing interests.

Chen *et al. BMC Plant Biology* (2015) 15:28

Page 12 of 14

Author details
[1]Beijing Key Laboratory of Grape Science and Enology, and CAS Key Laboratory of Plant Resources, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, P. R. China. [2]University of Chinese Academy of Sciences, Beijing 100049, P. R. China. [3]Key Laboratory of Plant Germplasm Enhancement and Speciality Agriculture, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China.

References
1. Liu HF, Wu BH, Fan PG, Li SH, Li LS. Sugar and acid concentrations in 98 grape cultivars analyzed by principal component analysis. J Sci Food Agric. 2006;86:1526–36.
2. Kliewer WM. Changes in concentration of glucose, fructose, and total soluble solids in flowers and berries of *Vitis vinifera*. Am J Enol Vitic. 1965;16:101–10.
3. Shiraishi M. Three descriptors for sugars to evaluate grape germplasm. Euphytica. 1993;71:99–106.
4. Doty TE. Fructose sweetness: a new dimension. Cereal Foods World. 1976;21:62–3.
5. Kulp K, Lorenz K, Stone M. Functionality of carbohydrates ingredients in bakery products. Food Technol. 1991;45:136–42.
6. Pangborn RM. Relative taste intensities of selected sugars and organic acids. J Food Sci. 1963;28:726–33.
7. Conde C, Silva P, Fontes N, Dias ACP, Tavares RM, Sousa MJ, et al. Biochemical changes throughout grape berry development and fruit and wine quality. Food. 2007;1:1–22.
8. Kliewer WM. Concentration of tartrates, malic acids, glucose and fructose in the fruits of genus *Vitis*. Am J Enol Vitic. 1967;18:87–96.
9. Kliewer WM, Howarth L, Omori M. Concentrations of tartaric acid and malic acids and their salts in *Vitis vinifera* grapes. Am J Enol Vitic. 1967;18:42–54.
10. Amerine MA, Roessler EB, Ough CS. Acids and the taste. I. The effect of pH and titratable acidity. Am J Enol Vitic. 1965;16:29–37.
11. Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl Acids Res. 1990;18:6531–5.
12. Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, et al. AFLP: a new technique for DNA fingerprinting. Nucl Acids Res. 1995;23:4407–14.
13. Adam-Blondon AF, Roux C, Claux D, Butterlin G, Merdinoglu D, This P. Mapping 245 SSR markers on the V vinifera genome: a tool for grape genetics. Theor Appl Genet. 2004;109:1017–27.
14. Costantini L, Battilana J, Lamaj F, Fanizza G, Grando MS. Berry and phenology-related traits in grapevine (Vitis vinifera L.): from quantitative trait loci to underlying genes. BMC Plant Biol. 2008;8:38.
15. Dalbó MA, Ye GN, Weeden NF, Steinkellner H, Sefc KM, Reisch BI. A gene controlling sex in grapevines placed on a molecular marker-based genetic map. Genome. 2000;43:333–40.
16. Di Gaspero G, Cipriani G, Adam-Blondon AF, Testolin R. Linkage maps of grapevine displaying the chromosomal locations of 420 microsatellite markers and 82 markers for *R*-gene candidates. Theor Appl Genet. 2007;114:1249–63.
17. Doligez A, Bouquet A, Danglot Y, Lahogue F, Riaz S, Meredith CP, et al. Genetic mapping of grapevine (*Vitis vinifera* L.) applied to the detection of QTLs for seedlessness and berry weight. Theor Appl Genet. 2002;105:780–95.
18. Doligez A, Adam-Blondon AF, Cipriani G, Di Gaspero G, Laucou V, Merdinoglu D, et al. An integrated SSR map of grapevine based on five mapping populations. Theor Appl Genet. 2006;113:369–82.
19. Doucleff M, Jin Y, Gao F, Riaz S, Krivanek AF, Walker MA. A genetic linkage map of grape, utilizing *Vitis rupestris* and *Vitis arizonica*. Theor Appl Genet. 2004;109:1178–87.
20. Fischer BM, Salakhutdinov I, Akkurt M, Eibach R, Edwards KJ, Töpfer R, et al. Quantitative trait locus analysis of fungal disease resistance factors on a molecular map of grapevine. Theor Appl Genet. 2004;108:501–15.
21. Grando MS, Bellin D, Edwards KJ, Pozzi C, Stefanini M, Velasco R. Molecular linkage maps of Vitis vinifera L and V riparia Mchx. Theor Appl Genet. 2003;106:1213–24.
22. Lamoureux D, Bernole A, Le Clainche I, Tual S, Thareau V, Paillard S, et al. Anchoring of a large set of markers onto a BAC library for the development of a draft physical map of the grapevine genome. Theor Appl Genet. 2006;113:344–56.
23. Lodhi MA, Daly MJ, Ye GN, Weeden NF, Reisch BI. A molecular marker based linkage map of *Vitis*. Genome. 1995;38:786–94.
24. Lowe KM, Walker MA. Genetic linkage map of the interspecific grape rootstock cross Ramsey (*Vitis champinii*) × Riparia Gloire (*Vitis riparia*). Theor Appl Genet. 2006;112:1582–92.
25. Mandl K, Santiago JL, Hack R, Fardossi A, Regner F. A genetic map of Welschriesling × Sirius for the identification of magnesium-deficiency by QTL analysis. Euphytica. 2006;149:133–44.
26. Marguerit E, Boury C, Manicki A, Donnart M, Butterlin G, Némorin A, et al. Genetic dissection of sex determinism, inflorescence morphology and downy mildew resistance in grapevine. Theor Appl Genet. 2009;118:1261–78.
27. Mejía N, Gebauer M, Muñoz L, Hewstone N, Muñoz C, Hinrichsen P. Identification of QTLs for seedlessness, berry size, and ripening date in a seedless × seedless table grape progeny. Amer J Enol Viticult. 2007;58:499–507.
28. Riaz S, Dangl GS, Edwards KJ, Meredith CP. A microsatellite marker based framework linkage map of *Vitis vinifera* L. Theor Appl Genet. 2004;108:864–72.
29. Riaz S, Krivanek AF, Xu K, Walker MA. Refined mapping of the Pierce's disease resistance locus, *PdR1*, and *Sex* on an extended genetic map of *Vitis rupestris* × *V. arizonica*. Theor Appl Genet. 2006;113:1317–29.
30. Salmaso M, Malacarne G, Troggio M, Faes G, Stefanini M, Grando MS, et al. A grapevine (*Vitis vinifera* L.) genetic map integrating the position of 139 expressed genes. Theor Appl Genet. 2008;116:1129–43.
31. Troggio M, Malacarne G, Coppola G, Segala C, Cartwright DA, Pindo M, et al. A dense single-nucleotide polymorphism-based genetic linkage map of grapevine (Vitis vinifera L.) anchoring Pinot Noir bacterial artificial chromosome contigs. Genetics. 2007;176:2637–50.
32. Zhang J, Hausmann L, Eibach R, Welter LJ, Töpfer R, Zyprian EM. A framework map from grapevine V3125 (Vitis vinifera 'Schiava grossa' × 'Riesling') × rootstock cultivar 'Börner' (*Vitis riparia* × *Vitis cinerea*) to localize genetic determinants of phylloxera root resistance. Theor Appl Genet. 2009;119:1039–51.
33. Wang N, Fang LC, Xin HP, Wang LJ, Li SH. Construction of a high-density genetic map for grape using next generation restriction-site associated DNA sequencing. BMC Plant Biol. 2012;12:148.
34. Barba P, Cadle-Davidson L, Harriman J, Glaubitz JC, Brooks S, Hyma K, et al. Grapevine powdery mildew resistance and susceptibility loci identified on a high-resolution SNP map. Theor Appl Genet. 2014;127:73–84.
35. Barker CL, Donald T, Pauquet J, Ratnaparkhe MB, Bouquet A, Adam-Blondon AF, et al. Genetic and physical mapping of the grapevine powdery mildew resistance gene, *Run1*, using a bacterial artificial chromosome library. Theor Appl Genet. 2005;111:370–7.
36. Coleman C, Copetti D, Capriani G, Hoffmann S, Kozma P, Kovács L, et al. The powdery mildew resistance gene *REN1* co-segregates with an NBS-LRR gene cluster in two Central Asian grapevines. BMC Genet. 2009;10:89.
37. Riaz S, Tenscher AC, Rubin J, Graziani R, Pao SS, Walker MA. Fine-scale genetic mapping of two Pierce's disease resistance loci and a major segregation distortion region on chromosome 14 of grape. Theor Appl Genet. 2008;117:671–81.
38. Marino R, Sevini F, Mandini A, Vecchione A, Pertot I, Serra AD, et al. QTL mapping for disease resistance and fruit quality in grape. Acta Hort. 2003;603:527–33.
39. Zyprian E, Eibach R, Töpfer R. Comparative molecular mapping in segregating populations of grapevine. Acta Hort. 2003;603:73–7.
40. Krivanek AF, Riaz S, Walker MA. Identification and molecular mapping of PdR1, a primary resistance gene to Pierce's disease in *Vitis*. Theor Appl Genet. 2006;112:1125–31.

Chen *et al. BMC Plant Biology* (2015) 15:28

Page 13 of 14

41. Xu K, Riaz S, Roncoroni NC, Jin Y, Hu R, Zhou R, et al. Genetic and QTL analysis of resistance to *Xiphinema index* in a grapevine cross. Theor Appl Genet. 2008;116:305–11.

42. Fanizza G, Lamaj F, Costantini L, Chaabane R, Grando MS. QTL analysis for fruit yield components in table grapes (*Vitis vinifera*). Theor Appl Genet. 2005;111:658–64.

43. Cabezas JA, Cervera MT, Ruiz-Garcia L, Carreño J, Martínez-Zapater JM. A genetic analysis of seed and berry weight in grapevine. Genome. 2006;49:1572–85.

44. Duchêne E, Butterlin G, Dumas V, Merdinoglu D. Towards the adaptation of grapevine varieties to climate change: QTLs and candidate genes for developmental stages. Theor Appl Genet. 2012;124:623–35.

45. Correa J, Mamani M, Muñoz-Espinoza C, Laborie D, Muñoz C, Pinto M, et al. Heritability and identification of QTLs and underlying candidate genes associated with the architecture of the grapevine cluster (*Vitis vinifera* L.). Theor Appl Gene t. 2014;127:1143–62.

46. Eibach R, Hastrich H, Töpfer R. Inheritance of aroma compounds. Acta Hort. 2003;603:337–44.

47. Fournier-level A, Le Cunff L, Gomez C, Doligez A, Ageorges A, Roux C, et al. Quantitative genetic bases of anthocyanin variation in grape (Vitis vinifera L. ssp. sativa) berry: a quantitative trait locus to quantitative trait nucleotide integrated study. Genetics. 2009;183:1127–39.

48. Doligez A, Bertrand Y, Dias S, Grolier M, Ballester JF, Bouquet A, et al. QTLs for fertility in table grapes (*Vitis vinifira* L.). Tree Genet Genomes. 2010;6:413–22.

49. Liu HF, Wu BH, Fan PG, Xu HY, Li SH. Inheritance of sugars and acids in berries of grape (Vitis vinifera L.). Euphytica. 2007;153:99–107.

50. Viana AP, Riaz S, Walker MA. Genetic dissection of agronomic traits within a segregating population of breeding table grapes. Genet Mol Res. 2013;12:951–64.

51. Dirlewanger E, Moing A, Rothan C, Svanella L, Pronier V, Guye A, et al. Mapping QTLs controlling fruit quality in peach (*Prunus persica* (L.) Batsch). Theor Appl Genet. 1999;98:18–31.

52. Quilot B, Wu BH, Kervella J, Génard M, Foulongne M, Moreau K. QTL analysis of quality traits in an advanced backcross between *Prunus persica* cultivars and the wild relative species *P. davidiana*. Theor Appl Genet. 2004;109:884–97.

53. Kenis K, Keulemans J, Davey MW. Identification and stability of QTLs for fruit quality traits in apple. Theor Appl Genet. 2008;4:647–61.

54. King GJ, Maliepaard C, Lynn JR, Alston FH, Durel CE, Evans KM, et al. Quantitative genetic analysis and comparison of physical and sensory descriptors relating to fruit flesh firmness in apple (*Malus pumila* Mill.). Theor Appl Genet. 2000;100:1074–84.

55. Liebhard R, Kellerhals M, Pfammatter W, Jertmini M, Gessler C. Mapping quantitative physiological traits in apple (*Malus × domestica* Borkh.). Plant Mol Biol. 2003;52:511–26.

56. Maliepaard C, Alston FH, Van Arkel G, Brown LM, Chevreau E, Dunemann F, et al. Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. Theor Appl Genet. 1998;97:60–73.

57. Wang D, Karle R, Iezzoni AF. QTL analysis of flower and fruit traits in sour cherry. Theor Appl Genet. 2000;100:535–44.

58. Cohen S, Tzuri G, Harel-Beja R, Itkin M, Portnoy V, Sa'ar U, et al. Co-mapping studies of QTLs for fruit acidity and candidate genes of organic acid metabolism and proton transport in sweet melon (*Cucumis melo* L.). Theor Appl Genet. 2012;125:343–53.

59. Bisson L: In search of optimal grape maturity. Pract Winery Vineyard 2001, 23: 32, 34, 36, 38, 40, 42–43.

60. Kupiec T. Quality-Control Analytical Methods: High-Performance Liquid Chromatography. Int J Pharm Compound. 2004;8:223–7.

61. Etter PD, Bassham S, Hohenlohe PA, Johnson EA, Cresko WA. SNP Discovery and Genotyping for Evolutionary Genetics Using RAD Sequencing. Methods Mol Biol. 2011;772:157–78.

62. Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH. *Stacks*: building and genotyping Loci *de novo* from short-read sequences. G3. 2011;1:171–82.

63. Grattapaglia D, Sederoff R. Genetic linkage maps of Eucalyptus grandis and Eucalyptus urophylla using a pseudo-testcross: mapping strategy and RAPD markers. Genetics. 1994;137:1121–37.

64. Zeng ZB. Precision mapping of quantitative trait loci. Genetics. 1994;136:1457–68.

65. Wang SC, Basternand J, Zeng ZB (2006): Windows QTL Cartographer 2.5. North Carolina State University, Raleigh, NC [http://statgen.ncsu.edu/qtlcart/WQTLCart.htm].

66. Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 2007;449:463–7.

67. Du Z, Zhou X, Ling Y, Zhang Z, Su Z. agriGO: a GO analysis toolkit for the agricultural community. Nucleic Acids Res. 2010;38:W64–70.

68. Shiraishi M. Proposed descriptors for organic acids to evaluate grape germplasm. Euphytica. 1995;81:13–20.

69. Liang ZC, Sang M, Ma AH, Zhao SJ, Zhong GY, Li SH. Inheritance of sugar and acid contents in the ripe berries of a tetraploid diploid grape cross population. Euphytica. 2011;182:251–9.

70. Velasco R, Zharkikh A, Troggio M, Cartwright DA, Pruss D, et al. A High Quality Draft Consensus Sequence of the Genome of a Heterozygous Grapevine Variety. PLoS One. 2007;12:e1326.

71. Grzeskowiak L, Costantini L, Lorenzi S, Stella Grando M. Candidate loci for phenology and fruitfulness contributing to the phenotypic variability observed in grapevine. Theor Appl Genet. 2013;126:2763–76.

72. Conner PJ, Brown SK, Weeden NF. Molecular-marker analysis of quantitative traits for growth and development in juvenile apple trees. Theor Appl Genet. 1998;96:1027–35.

73. Costa F, Peace CP, Stella S, Serra S, Musacchi S, Bazzani M, et al. QTL dynamics for fruit firmness and softening around an ethylene-dependent polygalacturonase gene in apple (*Malus × domestica* Borkh.). J Exp Bot. 2010;61:3029–39.

74. García MR, Asíns MJ, Carbonell EA. QTL analysis if yield and seed number in *Citrus*. Theor Appl Genet. 2000;101:487–93.

75. Kearsey MJ. The principles of QTL analysis (a minimal mathematics approach). J Exp Bot. 1998;49:1619–23.

76. Staub JE, Serquen FC, Gupta M. Genetic markers, map construction, and their application in plant breeding. HortSci. 1996;31:729–41.

77. Davies C, Robinson SP. Sugar accumulation in grape Berries (cloning of two putative vacuolar invertase cDNAs and their expression in grapevine tissues). Plant Physiol. 1996;111:275–83.

78. Causse M, Saliba-Colombani V, Lecomte L, Duffé P, Rousselle P, Buret M. QTL analysis of fruit quality in fresh market tomato: a few chromosome regions control the variation of sensory and instrumental traits. J Exp Bot. 2002;53:2089–98.

79. Fulton TM, Beck-Bunn T, Emmatty D, Eshed Y, Lopez J, Petiard V, et al. QTL analysis of an advanced backcross of *Lycopersicon peruvianum* to the cultivated tomato and comparisons with QTLs found in other wild species. Theor Appl Genet. 1997;95:881–94.

80. Liu JP, Magalhaes JV, Shaff J, Kochian LV. Aluminum-activated citrate and malate transporters from the MATE and ALMT families function independently to confer *Arabidopsis* aluminum tolerance. Plant J. 2009;57:389–99.

81. Ryan PR, Raman H, Gupta S, Horst WJ, Delhaize E. A second mechanism for aluminum resistance in wheat relies on the constitutive efflux of citrate from roots. Plant Physiol. 2009;149:340–51.

82. Yang XY, Yang JL, Zhou Y, Pineros MA, Kochian LV, Li GX, et al. A *de novo* synthesis citrate transporter, *Vigna umbellate* multidrug and toxic compound extrusion, implicates in Al-activated citrate efflux in rice bean (*Vigna umbellata*) root apex. Plant Cell Environ. 2011;34:2138–48.

83. Yokosho K, Yamaji N, Ma JF. An Al-inducible MATE gene is involved in external detoxification of Al in rice. Plant J. 2011;68:1061–9.

84. Maron LG, Guimarães CT, Kirst M, Albert PS, Birchler JA, Bradbury PJ, et al. Aluminum tolerance in maize is associated with higher *MATE1* gene copy number. Proc Natl Acad Sci U S A. 2013;110:5241–6.

85. Tovkach A, Ryan PR, Richardson AE, Lewis DC, Rathjen TM, Ramesh S, et al. Transposon-mediated alteration of *TaMATE1B* expression in wheat confers constitutive citrate efflux from root apices. Plant Physiol. 2013;161:880–92.

86. Palmieri L, Picault N, Arrigoni R, Besin E, Palmieri F, Hodges M. Molecular identification of three *Arabidopsis thaliana* mitochondrial dicarboxylate carrier isoforms: organ distribution, bacterial expression, reconstitution into liposomes and functional characterization. Biochem J. 2008;410:621–9.

87. Nieuwland J, Maughan S, Dewitte W, Scofield S, Sanz L, Murray James AH. The D-type cyclin CYCD4;1 modulates lateral root density in Arabidopsis by affecting the basal meristem region. PNAS. 2009;106:22528–33.

88. Sun JR, Wu Q, Zhou HX, Xu H, Cao Y, Xie L, et al. Correlation Between Polysaccharide Biosynthesis and Glycometabolism Related Enzymes of

Chen *et al. BMC Plant Biology* (2015) 15:28

Page 14 of 14

Chlamydomonas sp. YB-204 in Different Conditions. J Pure Microbio. 2014;8:1923–8.

89. Hu HC, Wang YY, Tsay YF. AtCIPK8, a CBL-interacting protein kinase, regulates the low-affinity phase of the primary nitrate response. Plant J. 2009;57:264–78.

90. Khalid MF, Damha MJ, Shuman S, Schwer B. Structure–function analysis of yeast RNA debranching enzyme (Dbr1), a manganese-dependent phosphodiesterase. Nucleic Acids Res. 2005;33:6349–60.

91. Fujita K, Horiuchi H, Takato H, Kohno M, Suzuki S. Fujita. Auxin-responsive grape *Aux/IAA9* regulates transgenic Arabidopsis plant growth. Mol Biol Rep. 2012;39:7823–9.

92. Wang H, Jones B, Li ZG, Frasse P, Delalande C, Regad F, et al. The Tomato *Aux/IAA* Transcription Factor *IAA9* Is Involved in Fruit Development and Leaf Morphogenesis. Plant Cell. 2005;17:2676–92.

93. Wang M, Zhang Q, Liu FC, Xie WF, Wang GD, Wang J, et al. Family-wide expression characterization of Arabidopsis beta-carbonic anhydrase genes using qRT-PCR and Promoter::GUS fusions. Biochimie. 2014;97:219–27.

94. Okumoto S, Schmidt R, Tegeder M, Fischer WN, Rentsch D, Frommer WB, et al. High affinity amino acid transporters specifically expressed in xylem parenchyma and developing seeds of arabidopsis. J Biol Chem. 2002;277:5338–45346.