Plant Mol Biol (2013) 81:161–174
DOI 10.1007/s11103-012-9989-8

Effect prediction of identified SNPs linked to fruit quality and chilling injury in peach [*Prunus persica* (L.) Batsch]

Pedro J. Martínez-García · Jonathan Fresnedo-Ramírez · Dan E. Parfitt · Thomas M. Gradziel · Carlos H. Crisosto

Received: 15 August 2012 / Accepted: 9 November 2012 / Published online: 25 November 2012

Abstract Single nucleotide polymorphisms (SNPs) are a fundamental source of genomic variation. Large SNP panels have been developed for *Prunus* species. Fruit quality traits are essential peach breeding program objectives since they determine consumer acceptance, fruit consumption, industry trends and cultivar adoption. For many cultivars, these traits are negatively impacted by cold storage, used to extend fruit market life. The major symptoms of chilling injury are lack of flavor, off flavor, mealiness, flesh browning, and flesh bleeding. A set of 1,109 SNPs was mapped previously and 67 were linked with these complex traits. The prediction of the effects associated with these SNPs on downstream products from the 'peach v1.0' genome sequence was carried out. A total of 2,163 effects were detected, 282 effects (non-synonymous, synonymous or stop codon gained) were located in exonic regions (13.04 %) and 294 placed in intronic regions (13.59 %). An extended list of genes and proteins that could be related to these traits was developed. Two SNP markers that explain a high percentage of the observed phenotypic variance, UCD_SNP_1084 and UCD_SNP_46, are associated with zinc finger (C3HC4-type RING finger) family protein and AOX1A (alternative oxidase 1a) protein groups, respectively. In addition, phenotypic variation suggests that the observed polymorphism for SNP UCD_SNP_1084 [A/G] mutation could be a candidate quantitative trait nucleotide affecting quantitative trait loci for mealiness. The interaction and expression of affected proteins could explain the variation observed in each individual and facilitate understanding of gene regulatory networks for fruit quality traits in peach.

Keywords Chilling injury symptoms · Fruit quality · Proteins · Quantitative trait nucleotide · Single nucleotide polymorphisms · SNP effects

Introduction

Peach [*Prunus persica* (L) Batsch] is one of the genetically best-characterized horticultural crops and is the model species for *Prunus* crops due to its small genome size, self-fertilization, diploid genetics, and long history of genetic improvement (Arus et al. 2012). Several linkage maps, a physical map, a whole-genome sequence (Arus et al. 2012), and a large amount of transcriptomic/metabolomics/proteomic data are available to the scientific community in the Genome Database Rosaceae (GDR) (http://www.rosaceae.org/). More than 200 QTLs in this database have been described for peach associated with traits such as bloom date, flowering time, chilling requirements, flesh weight, glucose, heat requirement, etc. Major genes affecting fruit flesh color (Linkage Group 1 -LG1-), flesh color around the stone (LG3), flesh adhesion (LG4), non-acid fruit (LG5), fruit shape (LG6), and fruit skin color (LG6 and LG8), have been located on the *Prunus* reference TxE map (Dirlewanger et al. 2004). QTLs for chilling injury symptoms such as mealiness, graininess and leatheriness (Cantin et al. 2010) as well as several candidate genes for fruit quality have been reported previously (Ogundiwin et al. 2009).

More than 15,000 single nucleotide polymorphisms (SNPs) have been identified and are publicly accessible for

Electronic supplementary material The online version of this article (doi:10.1007/s11103-012-9989-8) contains supplementary material, which is available to authorized users.

P. J. Martínez-García (✉) · J. Fresnedo-Ramírez · D. E. Parfitt · T. M. Gradziel · C. H. Crisosto
Department of Plant Sciences, University of California, Davis, One Shields Ave, Davis, CA 95616, USA
e-mail: pjmartinezgarcia@ucdavis.edu

peach (Ahmad et al. 2011; Verde et al. 2012) and an additional set will be soon available (Arus et al. 2012). SNPs are the most common type of genetic variation (Wang et al. 1998) with one SNP every 100–300 bp in some crop species (Appleby et al. 2009). SNPs represent a single base change between two individuals at a defined location, and may be applied to diversity analysis, genetic trait mapping, association studies, and marker assisted selection (Duran et al. 2009). SNPs can be C⇔T and G⇔A transitions, A⇔C, G⇔T, A⇔T, and C⇔G transversions (Brookes 1999), and/or small insertions and deletions (Appleby et al. 2009). The exact nature of the allelic variants can be ascertained from sequence information. The location and type of sequence variation determines phenotypic variation at the gene, protein, metabolite, tissue, organ, and organism levels, determining organism development and response to the environment (Duran et al. 2009).

SNPs differ from other types of molecular markers because they can be used to relate altered gene function with specific DNA changes in the genome. They usually are bialellic, non-transferable across species, even among individuals of the same species as shown for peach (Verde et al. 2012), seem to be located in areas with low sequence conservation in animals (Castle 2011), and can be located in both coding and non-coding regions of the genome. This last characteristic permits interesting aspects of structure and genome dynamics to be interrogated. It is possible to infer changes in the reading frame of regions involved in the control of quantitative trait loci (QTL), and to identify quantitative trait nucleotide (QTN) variants causing variation in QTLs (Mackay et al. 2009). QTNs, depending on which genes and traits they are affecting, may be useful for studies of conservation and assessment of diversity, for the conduct of breeding programs, as well as to understand phenomena influencing natural selection on evolution of complex traits (van Tienderen et al. 2002).

From the total number of peach SNPs discovered, high throughput SNP assays permitted identification of around 7,500 peach SNP polymorphisms in several mapping populations (Martínez-García et al. 2012a; Verde et al. 2012). SNP analysis allowed the first study of nucleotide diversity in two candidate genes for fruit firmness and sugar content (Aranzana et al. 2012). The authors of these studies observed low levels of SNP polymorphism consistent with the low diversity levels reported for this species (Arulsekar et al. 1986; Font i Forcada et al. 2012).

Dense SNP linkage maps and QTLs analyses for peach fruit quality, chilling injury symptoms and aroma volatile compounds have been generated recently (Martínez-García et al. 2012a; Eduardo et al. 2012). SNPs may also be used to predict the protein/metabolism/phenotype effects that they could produce through modification of known genes or gene functions, permitting their use to target specific phenotypes of interest in peach breeding programs.

The presence of SNP genetic variants can produce changes in gene products, which can be linked to phenotypic differences. High throughput bioinformatic tools have been developed to predict post-transcriptional effects for large SNP populations. Two important web tools, PupaSuite and SNPeffect 3.0, have been created for analysis of SNP databases to predict the effects or expected phenotypes from specific sets of SNPs (Reumers et al. 2006). More recently, the software SnpEff has been developed (Cingolani et al. 2012). This software is a bioinformatics tool that annotates the variants (SNPs, insertions, deletions, and multiple nucleotide polymorphisms) and calculates the effects they produce on known genes present in the annotation of the reference genome sequence through an algorithm based on interval trees and implemented in the Java programming language.

The objective of this study was to obtain a list of candidate genes associated with fruit quality traits and chilling injury symptoms in peach using SNPs effect prediction. To accomplish this objective, mapped SNPs associated with QTLs for these traits (Martínez-García et al. 2012a) were used to predict the expected effects of SNP changes on annotated genes in the 'peach v1.0' genome sequence. We report a complete characterization and description of each effect obtained using the software SnpEff (Cingolani et al. 2012).

Materials and methods

Selection of SNPs

All information about the SNPs (discovery, names, genetic position), description of the two mapping populations (Pop-DF and Pop-DG), the fruit quality traits (yellow flesh and freestone-melting flesh) and chilling injury symptoms considered (mealiness, flesh bleeding and flesh browning) and description of QTLs (names, position, nearest markers, etc.) associated with these traits, have been described in previous published works from our group (Ahmad et al. 2011; Martínez-García et al. 2012a, b). A set of 67 SNPs markers significantly associated with QTLs for these important traits (Fig. 1), and an additional set of 1,042 SNPs from SNP mapping, a total set of 1,109 SNPs shared in both populations (Martínez-García et al. 2012a), were selected for SNP effect prediction on known genes using the functional information based on the annotation available for the Peach Reference Genome at: http://www.rosaceae.org/species/prunus_persica/genome_v1.0

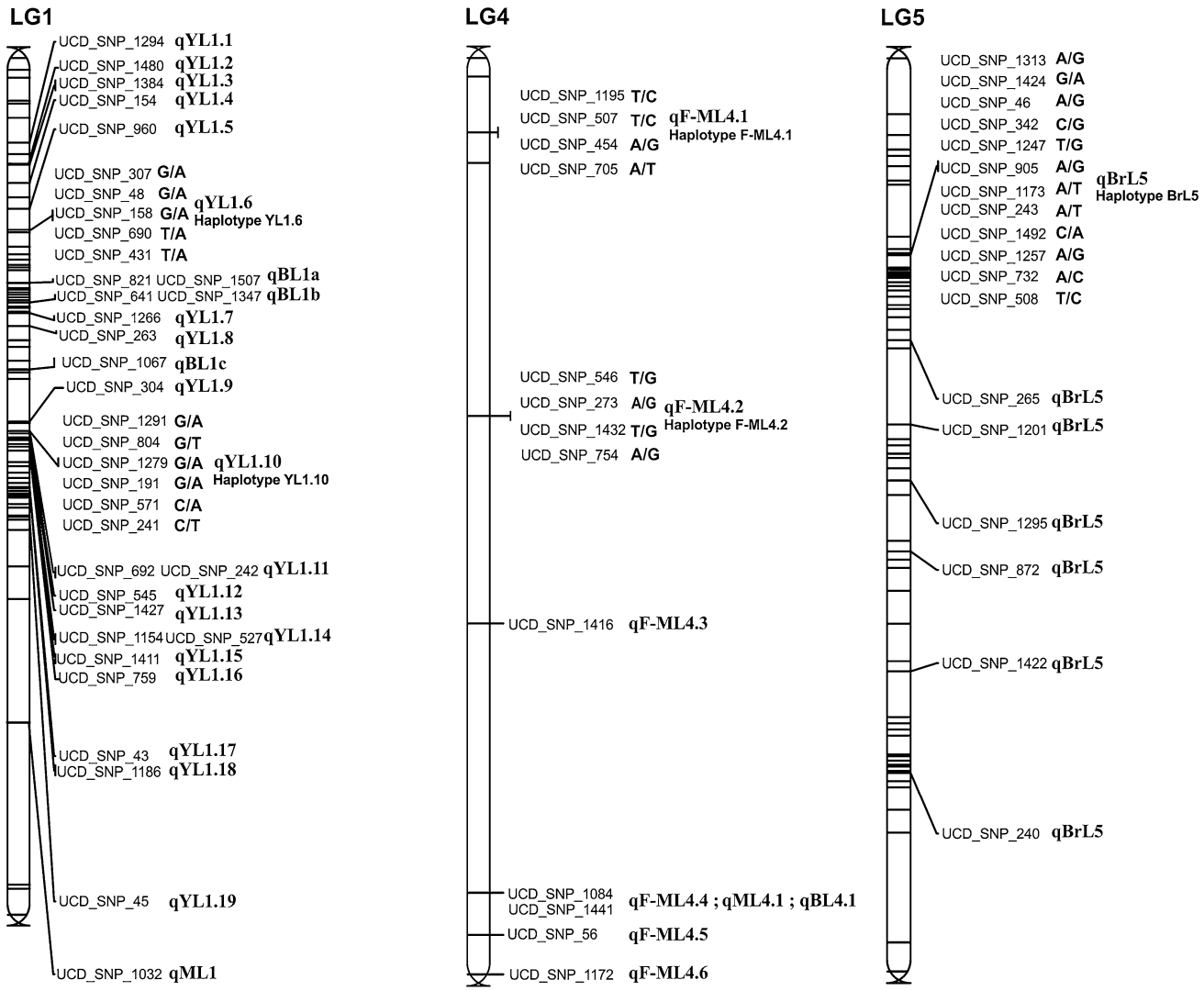

Fig. 1 SNPs markers and haplotypes significantly associated to QTLs for fruit quality and chilling injury symptoms in linkage group LG1, LG4 and LG5 previously described (Martínez-García et al. 2012b), qY = SNPs associated to QTL for yellow flesh, qF-M = SNPs associated to QTL for freestone-melting flesh, qB = SNPs associated to QTL for flesh bleeding, qBr = SNPs associated to QTL for flesh browning

Prediction of SNP effects

The initial 20.2 GB input file for this analysis was the combined pileup file created and compiled previously (Ahmad et al. 2011), which contains the alignments of 'Dr. Davis', 'Georgia Belle' and 'F8-1-42' against the 'peach v1.0' genome sequence. The software package VarScan 2.2.5 was used to check strand orientation for the SNPs (Koboldt et al. 2009). Subsequently, the output file was filtered and converted to a plain text file (.txt) to be analyzed using the software SnpEff ver. 3.0c. (Cingolani et al. 2012). The text file contained the scaffold locations of the SNPs, the reference nucleotide sequenced, the sequence reported as a change (the SNP), and the strand in which the SNP occurs (+ or −). Effect analyses were performed using the 'peach v1.0 genome' sequence and gene annotations with the default parameters of SnpEff.

A SnpEff predictor database file (17 MB) in binary format (.bin) was created to locate each SNP within annotated transcripts or intronic regions. Both HTML and text output files were generated from SnpEff. The output included the position of the SNP on the scaffold, the reference nucleotide, the changed nucleotide, if it was a transition or a transversion, transitions/tranversions ratio (Ts/Tv), warnings, gene ID, gene name, biotype, transcript ID, exon ID, exon rank effect, amino acid change (old aa/new aa), old codon/new codon, codon number [based on the coding sequence (CDS)], and CDS size. Warnings were provided if the predicted changes differed from those predicted using the 'peach v1.0' genome.

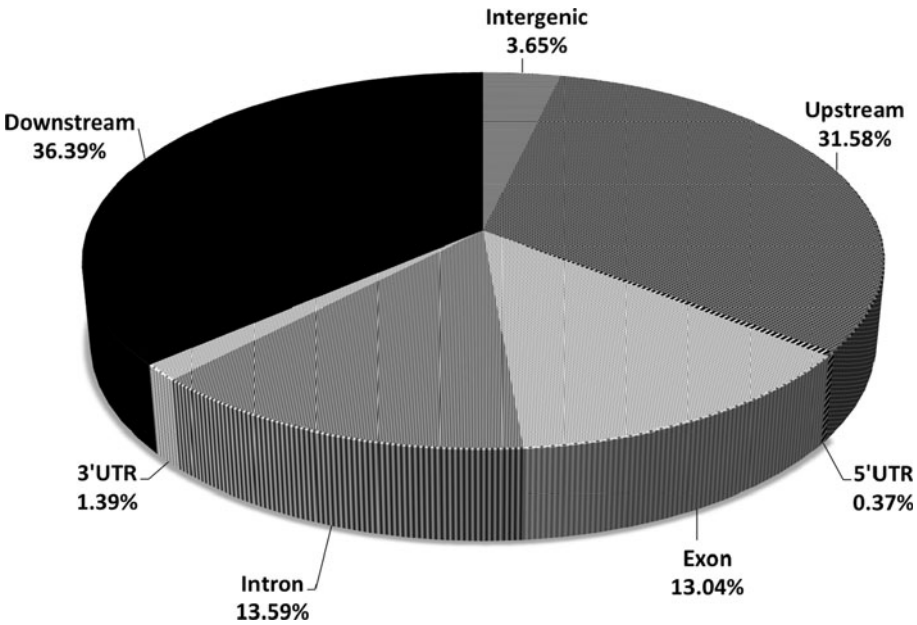

Fig. 2 Percentage of variants in each region of the 'peach v1.0' genome

Results

Prediction of SNP effects

Of 1,109 SNPs that were mapped and shared by the two mapping populations 2,163 alterations to gene products were predicted in a genome with an effective length of ~220 Mb in 13 scaffolds. This would represent one change in every 197,917 bases (supplemental file 1). A total of 2,163 individual effects, rather than 1,109, were identified from the SNP effect analysis because a change in one SNP can alter multiple transcripts at the same gene region, as reported in the electronic annotation developed for the peach genome. Scaffold_1 had the most variants (227 changes) while scaffold_53, scaffold _62, and scaffold _127 had only one variant each. No warnings were observed in this analysis.

Impact scores were used to estimate the types of changes that SNPs could effect. From the impact scores, 86.9 % of the effects were classified as modifiers and only one effect caused a high impact on genes (0.046 %). 56.3 % of the effects caused a silent change (synonymous) in the functional class of effects, 43.4 % caused a missense change (non-synonymous), and only one effect (0.3 %) was a nonsense change (one stop codon gained). Most SNP effects were predicted to occur downstream (36.39 %) and upstream (31.58 %) of genes. A total of two start codons and another stop codon were created as a result of SNP changes. Exons accounted for 13.04 % of the SNP effects while 13.59 % were located within introns. Intergenic regions had 79 (3.65 %) SNP effects (Fig. 2). The most frequent base change was from adenine to guanine (204 changes) and the least frequent was from guanine to cytosine (34 changes). There were 674 transitions (Ts) and 435 transversions (Tv), a Ts/Tv ratio of 1.55. The most frequent codon changes were TCA to TCG (seven changes), GAC to GAT (six changes), TTG to TTA, and AAG to AAA (five changes). The most frequent non-synonymous changes in amino acids were alanine to threonine and valine to alanine (five changes each). Amino acids with four changes each were aspartic acid to glutamic acid, valine to isoleucine, arginine to lysine, and lysine to asparagine. In our results, 122 SNPs were non-synonymous and located in exonic regions according to the available annotation (supplemental file 2). Thus, several genes and related proteins with potential changes have been reported in several different species as determined from gene homology.

SNP effects for homologous transcripts were obtained from 151 different organisms, from microbes such as *Acinetobacter* sp. to mammals such as *Homo sapiens,* based on the predicted biochemical and biological effects of the SNPs. Annotation of the peach v1.0 (http://www.rosaceae.org/species/prunus_persica/genome_v1.0) was generated by gene models that based homology prediction on information publically available from several organisms.[1] For this reason, our dataset also includes results reported from all classes of organisms. *Arabidopsis thaliana* L. provided 57 % of the homologous transcripts (Fig. 3). Many other homologous transcripts were found in plants of biological genetic and economic relevance

[1] Additional information can be consulted at http://services.appliedgenomics.org/projects/drupomics/.

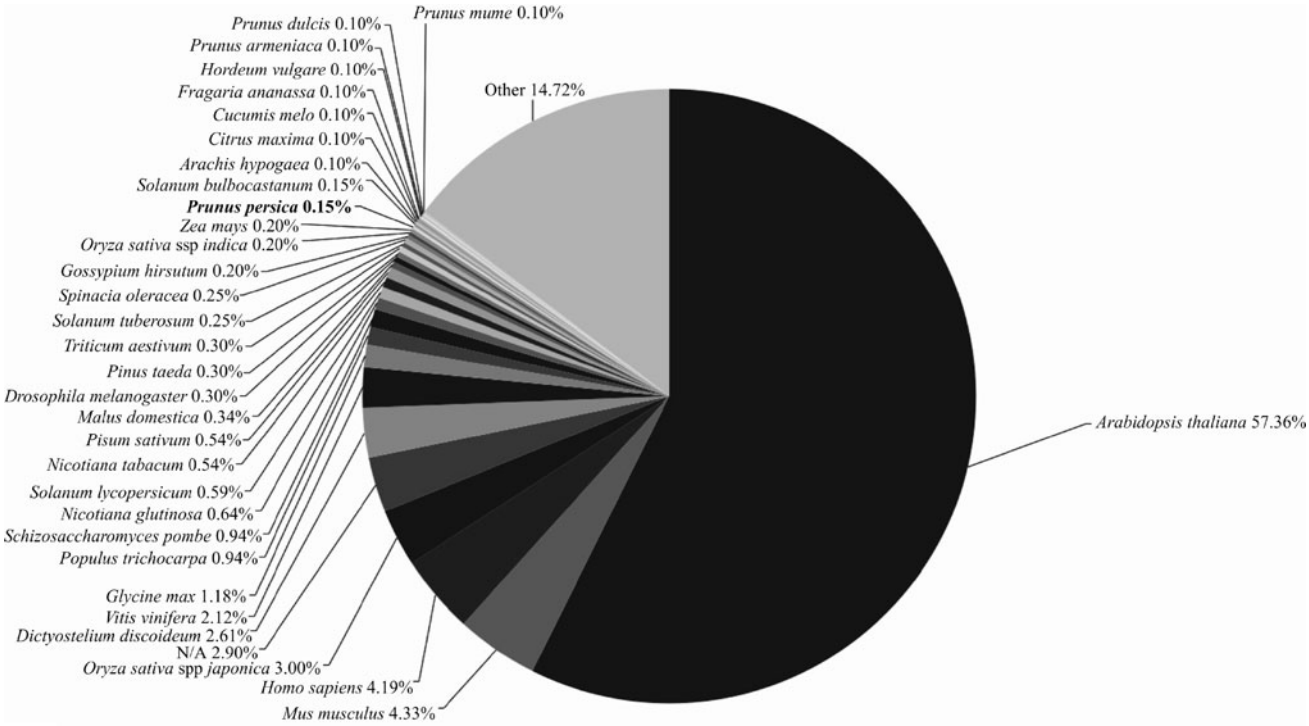

Fig. 3 Percentage of homologous transcripts found in each organism

including japonica rice (61), tomato (12), grape (45), *Populus* (22), soybean (24), and other *Prunus* species (9). Therefore, SNPs located in homologous transcript locations in peach could produce protein changes or cause changes in regulatory sequences potentially involved in traits of economic and breeding relevance (listed in supplemental file 2). Among the predicted effects were several examples of interest in peach including effects related to DEAD-box, ATP-dependent RNA helicase, and heat stress transcription factors from rice; agamous-like MADS-box proteins, ethylene-responsive transcription factors, and ATP binding from *Arabidopsis*; auxin-induced sucrose synthase and cytochrome proteins from soybean; heat shock proteins from petunia; cyanogenic beta-glucosidase from *Trifolium repens* L.; and auxin-binding protein ABP20 from peach (supplemental file 2).

Discussion

SNPs are providing information beyond statistical associations for linkage or QTL analyses. Our results allowed the identification of new putative candidate genes associated with fruit quality traits and chilling injury symptoms. The effects of 67 SNPs identified as significantly associated with fruit quality and chilling injury symptoms (Martínez-García et al. 2012a) are discussed in depth.

The Ts:Tv ratio for the peach sequences was 1.5:1.[2] This value was similar to the one obtained previously in studies of human SNPs from EST sequence trace databases and in a study comparing rodent and human sequences (Collins and Jukes 1994; Picoult-Newberg et al. 1999). In avian studies (chicken and mallard), the value was higher at 2.2:1 (Kraus et al. 2011; Sherry et al. 2001). Genomewide Ts/Tv ratios around values of 2.0 are common, while Ts/Tv ratios within exonic regions may be above 3.0. In plant species, such as maize, alfalfa, eikorn wheat (*Triticum monococcum* L.), barley and *Lotus*, the Ts/Tv ratios of Non-long Terminal Repeat (Non-LTR) retrotransposon sequences have been estimated as 3.9, 3.6, 1.9, 1.6, and 2.5, respectively (Vitte and Bennetzen 2006). Information about Ts:Tv ratios is scarce for *Prunus* species.

It has been suggested that the higher number of Ts events could be explained by a higher rate of C⇔T (transition) mutations due to the deamination of methyl-cytosines in CpG dinucleotides (Cooper and Krawczak 1989; Vignal et al. 2002); although methylation may also be

[2] Pablo Cingolani, developer of SnpEff, notes in the output of SnpEff 3.0 that: "this Ti/Tv ratio is a raw ratio. Some people prefer to use a ratio of rates, not observed events. In that case, you need to multiply by 2.0 (since there are twice as many possible transitions than transversions E[Ts/Tv] ratio is twice the ratio of events)". In our case, the value given in the output is Ti/Tv = 1.5494, multiplied by 2.0, the ratio of rates = 3.0988.

important, since C⇔T and G⇔A changes can occur after methylation (Coulondre et al. 1978) in both plants and animals (SanMiguel et al. 1998). In our study the most common nucleotide substitution was the transition A⇔G, which is the result of depurination, and is more common than deamination, which contributes a higher frequency of Ts events. The presence of transition mutations in maize ranges from 60 to 100 % of all the present mutations (Morton et al. 2006).

The estimation of the Ts:Tv ratio provides insight into processes of molecular evolution (Kimura 1980). DNA sequence evolution is more commonly influenced by transition mutations in comparison to transversion mutations (Wakeley 1996), the former being more frequent. Therefore, the Ts:Tv ratio provides information about the extent of transition bias, which varies depending on the organism, sequences/genes in an organism, or among species (Strandberg and Salter 2004). The Ts:Tv ratio helps to describe the rates of change and divergence; hence, it is a theoretical estimator of mutation rates and evolutionary divergence, which is not directly related to observed rates of change at the phenotypic level, since other factors (e.g. natural selection pressure) are important factors driving observed evolutionary change. For peach, the sequences analyzed come from a domesticated species under strong directed artificial selection pressure. Additionally, the Ts:Tv $>$ 1.5:1 ratio estimated on LTR retrotransposon sequences may suggest epigenetic silencing (Gruenbaum et al. 1981).

Proteins associated with candidate SNPs related to yellow flesh and freestone-melting flesh

The SNP marker UCD_SNP_759, associated with yellow flesh, showed an effect on RNA-binding protein 25. This proteins has been associated with unrelated biological processes in humans such as RNA splicing, apoptotic processes, mRNA processing, and regulation of alternative nuclear mRNA splicing via spliceosomes; however, no association with biological processes in plants has been observed until now. Although both SNP markers UCD_SNP_242 and UCD_SNP_692 presented same genetic map position (Fig. 1), only UCD_SNP_692 presented an effect associated with a sterile alpha motif (SAM) domain-containing protein. The sterile alpha motif (SAM) domain-containing protein (D7MM55 accession number in UniProtKB/TrEMBL) was also observed in *Arabidopsis thaliana* and *Arabidopsis lyrata*, associated with several functions related to growth patterns, developmental stages, and plant structure inferred from expression patterns (Hu et al. 2011). Other markers with putative significant effects were UCD_SNP_304, UCD_SNP_527, and UCD_SNP_1154, which were associated with other annotated functions such as defense response ATP-dependent helicase activity, ribonuclease H activity in Arabidopsis, and defense response by callose deposition in cell walls. In the haplotype block observed for yellow flesh (Fig. 1), several additional effects were obtained; two SNPs UCD_SNP_431 and UCD_SNP_690 were associated with two different sugar transporters (Sugar transport protein 13 and Sucrose transporter 2) and the SNP UCD_SNP_ 191 was associated to the family of transcription factor bZIP (Table 1).

Effects associated with freestone-melting flesh included the SNP marker UCD_SNP_ 1084, associated with a predicted effect on two proteins; a zinc RING finger family protein and pentatricopeptide repeat-containing protein. Three other markers presented different effects on several uncharacterized and unknown proteins, proteins involved in translation (mitochondrial ribosomal protein S10), and RNA-dependent DNA replication (LINE-1 reverse transcriptase homolog). Several F-box proteins were affected for some of the SNPs associated with the haplotypes blocks for this trait (Table 2). The possible role and function of the zinc RING finger family protein and pentatricopeptide repeat-containing protein are discussed in the following section.

Proteins associated with candidate SNPs related to mealiness and flesh bleeding

The predicted effect of SNP marker UCD_SNP_1032, significantly associated with the QTL for mealiness in the linkage group 1 (qML1) (Fig. 1), was related to a light-inducible protein (CPRF2), an uncharacterized aarF domain-containing protein kinase 1, and a probable calcium-binding protein (CML43).

CPRF2 protein belongs to the bZIP family; its function is binding to the G-box-like motif (5′-ACGTGGC-3′) of the chalcone synthase (CHS) gene promoter. G-box and G-box-like motifs are defined in promoters of certain plant genes that are regulated by light-induction or hormone control. A member of the bZIP family, bZIP protein BZ2 was found to be up-regulated, using the ChillPeach database, to identify differentially expressed genes in cold-treated compared to control mesocarp tissues of peach (Ogundiwin et al. 2008).

The SNP markers UCD_SNP_1084 and UCD_SNP_1441 are associated with QTLs for freestone-melting flesh (F-M4.4), mealiness (qML4), and flesh bleeding (qBL4) (Fig. 1). Three predicted effects of the SNP UCD_SNP_1084 were found; the locations of the variants were in upstream regions and in intron regions. The SNP location showed that the zinc finger (C3HC4-type RING finger) family protein was modified by this particular SNP. In the Ogundiwin et al. (2008) study of cold-responsive genes in peach, a putative zinc

Table 1 The predicted effects for each SNP from the haplotype block associated with Yellow flesh

SNP (peach genome position)	Effect	Homologous gene	Reference organism	Description
UCD_SNP_48 (scaffold_1:11822333)	UPSTREAM: 4440 bases	GI1L2_ARATH	*Arabidopsis thaliana*	Probable gibberellin receptor GID1L2
UCD_SNP_48 (scaffold_1:11822333)	SYNONYMOUS_CODING	VAC8_CRYNE	*Cryptococcus neoformans*	Vacuolar protein 8
UCD_SNP_158 (scaffold_1:11825623)	SYNONYMOUS_CODING	VAC8_CRYNE	*Cryptococcus neoformans*	Vacuolar protein 8
UCD_SNP_158 (scaffold_1:11825623)	UPSTREAM: 4548 bases	Y1677_ARATH	*Arabidopsis thaliana*	Probable LRR receptor-like serine/threonine-protein kinase At1g67720
UCD_SNP_191 (scaffold_1:26648503)	UPSTREAM: 1488 bases	AT1G14410.1	*Arabidopsis thaliana*	WHY1 (WHIRLY 1); DNA binding/telomeric DNA binding
UCD_SNP_191 (scaffold_1:26648503)	UPSTREAM: 2706 bases	AT1G27000.1	*Arabidopsis thaliana*	bZIP family transcription factor
UCD_SNP_191 (scaffold_1:26648503)	SYNONYMOUS_CODING	AT5G20220.2	*Arabidopsis thaliana*	Zinc knuckle (CCHC-type) family protein
UCD_SNP_241 (scaffold_1:26867949)	DOWNSTREAM: 1527 bases	AT2G02860.1	*Arabidopsis thaliana*	SUT2 (SUCROSE TRANSPORTER 2)
UCD_SNP_241 (scaffold_1:26867949)	UPSTREAM: 2544 bases	ATB13_ARATH	*Arabidopsis thaliana*	Homeobox-leucine zipper protein ATHB-13
UCD_SNP_307 (scaffold_1:11173257)	DOWNSTREAM: 3701 bases	BSDC1_BOVIN	*Bos taurus* (Bovine)	BSD domain-containing protein 1
UCD_SNP_307 (scaffold_1:11173257)	UPSTREAM: 897 bases	BI1L_ARATH	*Arabidopsis thaliana*	BI1-like protein
UCD_SNP_307 (scaffold_1:11173257)	DOWNSTREAM: 4214 bases	AT2G01050.1	*Arabidopsis thaliana*	Nucleic acid binding/zinc ion binding
UCD_SNP_431 (scaffold_1:11912526)	UPSTREAM: 3426 bases	STP13_ARATH	*Arabidopsis thaliana*	Sugar transport protein 13
UCD_SNP_431 (scaffold_1:11912526)	UPSTREAM: 3491 bases	STP13_ARATH	*Arabidopsis thaliana*	Sugar transport protein 13
UCD_SNP_431 (scaffold_1:11912526)	UPSTREAM: 621 bases	STP13_ARATH	*Arabidopsis thaliana*	Sugar transport protein 13
UCD_SNP_571 (scaffold_1:26742400)	DOWNSTREAM: 1448 bases	D1IF42_VITVI	*Vitis vinifera* (Grape)	Whole genome shotgun sequence of line PN40024, scaffold_26.assembly12x
UCD_SNP_571 (scaffold_1:26742400)	SYNONYMOUS_CODING	IQD31_ARATH	*Arabidopsis thaliana*	Protein IQ-DOMAIN 31
UCD_SNP_690 (scaffold_1:11880352)	INTRON	ACSL6_RAT	*Rattus norvegicus* (Rat)	Long-chain-fatty-acid–CoA ligase 6
UCD_SNP_804 (scaffold_1:26260043)	UPSTREAM: 3696 bases	N/A	N/A	N/A
UCD_SNP_1291 (scaffold_1:26155955)	UPSTREAM: 447 bases	Y1719_ARATH	*Arabidopsis thaliana*	Probable inactive receptor kinase At1g27190
UCD_SNP_1291 (scaffold_1:26155955)	UPSTREAM: 4055 bases	AT1G69980.1	*Arabidopsis thaliana*	Unknown protein
UCD_SNP_1291 (scaffold_1:26155955)	UPSTREAM: 4055 bases	AT1G69980.1	*Arabidopsis thaliana*	Unknown protein
UCD_SNP_1291 (scaffold_1:26155955)	DOWNSTREAM: 2994 bases	MYBJ_DICDI	*Dictyostelium discoideum* (Slime mold)	Myb-like protein J

Table 1 continued

SNP (peach genome position)	Effect	Homologous gene	Reference organism	Description
UCD_SNP_1279 (scaffold_1:26351546)	UTR_3_PRIME: 32 bases from CDS	GSTU6_ORYSJ	*Oryza sativa* L. ssp. *japonica* (Rice)	Probable glutathione S-transferase GSTU6

The location on the peach genome, the homologous gene, the reference organism and description of the protein associated with known protein coding regions are shown in the table

Table 2 The predicted effects for each SNP from the haplotype block associated with Freestone-Melting flesh

SNP (Peach genome position)	Effect	Homologous gene	Reference organism	Description
UCD_SNP_273 (scaffold_4:12199704)	DOWNSTREAM: 727 bases	AT4G17750.1	*Arabidopsis thaliana*	HSF1 (HEAT SHOCK FACTOR 1); DNA binding/protein binding/transcription factor
UCD_SNP_454 (scaffold_4:10881939)	INTRON	CTL2_MOUSE	*Mus musculus* (Mouse)	Choline transporter-like protein 2
UCD_SNP_507 (scaffold_4:10718725)	DOWNSTREAM: 4264 bases	CHIT1_HUMAN	*Homo sapiens* (Human)	Chitotriosidase-1
UCD_SNP_507 (scaffold_4:10718725)	INTRON	SPA3_ARATH	*Arabidopsis thaliana*	Protein SPA1-RELATED 3
UCD_SNP_546 (scaffold_4:12072804)	DOWNSTREAM: 3691 bases	AT5G54540.1	*Arabidopsis thaliana*	
UCD_SNP_546 (scaffold_4:12072804)	INTRON	MCES1_ARATH	*Arabidopsis thaliana*	mRNA cap guanine-N7 methyltransferase 1
UCD_SNP_546 (scaffold_4:12072804)	DOWNSTREAM: 4008 bases	BGL41_ARATH	*Arabidopsis thaliana*	Putative beta-glucosidase 41
UCD_SNP_705 (scaffold_4:10940959)	INTRON	NOG2_MOUSE	*Mus musculus* (Mouse)	Nucleolar GTP-binding protein 2
UCD_SNP_705 (scaffold_4:10940959)	UPSTREAM: 3387 bases	LRC40_XENLA	*Xenopus laevis* (African clawed frog)	Leucine-rich repeat-containing protein 40
UCD_SNP_754 (scaffold_4:13807204)	DOWNSTREAM: 4242 bases	FB138_ARATH	*Arabidopsis thaliana*	Putative F-box protein At3g10430
UCD_SNP_754 (scaffold_4:13807204)	UPSTREAM: 4064 bases	FB223_ARATH	*Arabidopsis thaliana*	Putative F-box protein At4g09190
UCD_SNP_754 (scaffold_4:13807204)	DOWNSTREAM: 1291 bases	AT3G19410.1	*Arabidopsis thaliana*	F-box family protein
UCD_SNP_1195 (scaffold_4:10565069)	SYNONYMOUS_CODING	TT12_ARATH	*Arabidopsis thaliana*	Protein TRANSPARENT TESTA 12
UCD_SNP_1195 (scaffold_4:10565069)	UPSTREAM: 4905 bases	N/A	N/A	N/A
UCD_SNP_1195 (scaffold_4:10565069)	UPSTREAM: 2576 bases	Y1235_ORYSJ	*Oryza sativa* L. ssp. *japonica* (Rice)	B3 domain-containing protein Os01g0723500
UCD_SNP_1432 (scaffold_4:13746645)	UPSTREAM: 951 bases	AT5G35735.1	*Arabidopsis thaliana*	Auxin-responsive family protein
UCD_SNP_1432 (scaffold_4:13746645)	UPSTREAM: 1912 bases	AT1G11925.1	*Arabidopsis thaliana*	Encodes a Stigma-specific Stig1 family protein
UCD_SNP_1432 (scaffold_4:13746645)	UPSTREAM: 4754 bases	PROSC_BOVIN	*Bos taurus* (Bovine)	Proline synthetase co-transcribed bacterial homolog protein

The location on the peach genome, the homologous gene, the reference organism and description of the protein associated with known protein coding regions are shown in the table

RING finger ankyrin protein was shown to be up-regulated in cold treated mesocarp tissue.

The previous candidate genes associated with mealiness were previously bin-mapped on Linkage Group 4 (MADS-RIN, ACO3, TAT, Unk17, and Unk28) and on LG6 (UnK15 and Unk24) of the fruit quality gene map of *Prunus* (Ogundiwin et al. 2009). In the present study, the presence of these candidate genes was not detected. It can be a result of the marker technology used in each study. Since both studies used the same mapping population (Pop-DG). Differences in the results of an assessment of genetic variation depending on the type of molecular marker used (SSR and AFLP in their case) have been reported by Ramakrishnan et al. (2004). Additionally, we used a greater number of markers, which provide high resolution and accuracy to the study.

Classification by SnpEff suggests that the UCD_SNP_1084 variant is located in the intronic region of the gene, causing a modifier effect in the function of the zinc RING finger family protein. For the pentatricopeptide repeat-containing protein, the change is upstream of the gene, also producing a modifier effect.

Rice (*Oryza sativa*) zinc finger (C3HC4-type RING finger) family proteins have been related to growth, development, and stress response (Ma et al. 2009). Those results suggested that zinc finger genes respond to abiotic stresses, and may be mediated by abscisic acid (ABA)-dependent pathways. ABA is often referred to as the typical 'stress hormone' due to its important role in regulating plant response to several stresses (Zhang et al. 2006). In addition, a previous study of gene expression in virus infected grapevine leaves showed that the zinc finger family of proteins was induced during viral infection (Espinoza et al. 2007). In apple (*Malus domestica* Borkh.), Li et al. (2011) have determined that stress conditions, by themselves, are sufficient to induce or control expression of regulators of target proteins as RING finger proteins (containing E3 ligases).

The relationship between cold acclimation and chilling injury and identified cold-responsive genes expressed in peach has been studied previously (Tittarelli et al. 2009). These authors found 164 contigs with preferential expression in cold stored peach fruits at 4 °C. The C3HC4-type RING finger family was associated with cellular protein metabolism, but with a statistically non-significant cold-induction response similar to the cold response in Arabidopsis (down-regulated) for these genes. An additional study during bud dormancy in peach by suppression subtractive hybridization (SSH) showed that a zing-finger transcription factor was induced by cold (Leida et al. 2010). These authors observed that the expression of this transcription factor was about five times lower at the end of the dormancy than in the first collected samples.

An experiment in collaboration between research teams at UC Davis and IBMCP CSIC-UPV to identify differentially expressed genes during chilling response, using samples of mesocarp from fruits of two progeny peach trees of Pop-DG, was carried out (unpublished data C. Crisosto personal communication). The researchers found that zinc finger (C3HC4-type RING finger) family proteins involved with cell wall redox homeostasis showed a positive correlation with mealiness at several stages of cold storage and ripening, supporting the results from the present study.

The results from SnpEff also associated UCD_SNP_1084 with a pentatricopeptide repeat-containing protein. An analysis at transcriptomic, proteomic, and metabolomic levels in grapevine berry development and postharvest drying (withering) identified a pentatricopeptide repeat-containing protein as a putative withering-specific biomarker (Zamboni et al. 2010).

The SNP marker UCD_SNP_1084 revealed a Guanine (G) to Adenine (A) transition. The cultivar "Dr. Davis" (female cultivar used to obtain Pop-DG) showed a genotype (GG) and "Georgia Belle" (male cultivar of Pop-DG) presented a genotype (GA) for this locus. The frequency of both genotypes (GG and GA) in the entire population was 42 and 58 %, respectively. 'Dr. Davis' (GG) is resistant to mealiness while 'Georgia Belle' (GA) has a high susceptibility to mealiness. The progenies with the GG genotype showed resistance to mealiness and progenies with the GA genotype presented a medium–low susceptibility to mealiness (Fig. 4). Since 'Georgia Belle' is more susceptible to mealiness than the heterozygous progeny, additional loci may be conditioning the high level of susceptibility in 'Georgia Belle', while other less dominant resistance genes are present in the progeny.

Results from prior studies of this group and those presented here suggest that a zinc finger (C3HC4-type RING finger) family protein is associated with a gene mapped on LG4, which is positionally related to the loci associated to freestone-melting flesh, mealiness and flesh bleeding. In addition, results from a recently published study, using principal components analysis of gene expression in four different fruit with different sensibility to chilling injury, suggested that RING finger-like protein (ID: AT5G19430) has a role in cold response adaptation (Dagar et al. 2012). The gene marked by UCD_SNP_1084 is a potential candidate gene for developing resistance to chilling injury symptoms. The SNP marker UCD_SNP_1084 could be an important biomarker (a QTN) for fruit quality and chilling injury symptoms such as mealiness and flesh bleeding in peach.

The QTN for mealiness (causing an observable difference in expression) must be validated with a larger number of samples (e.g. with different cultivars or breeding lines)

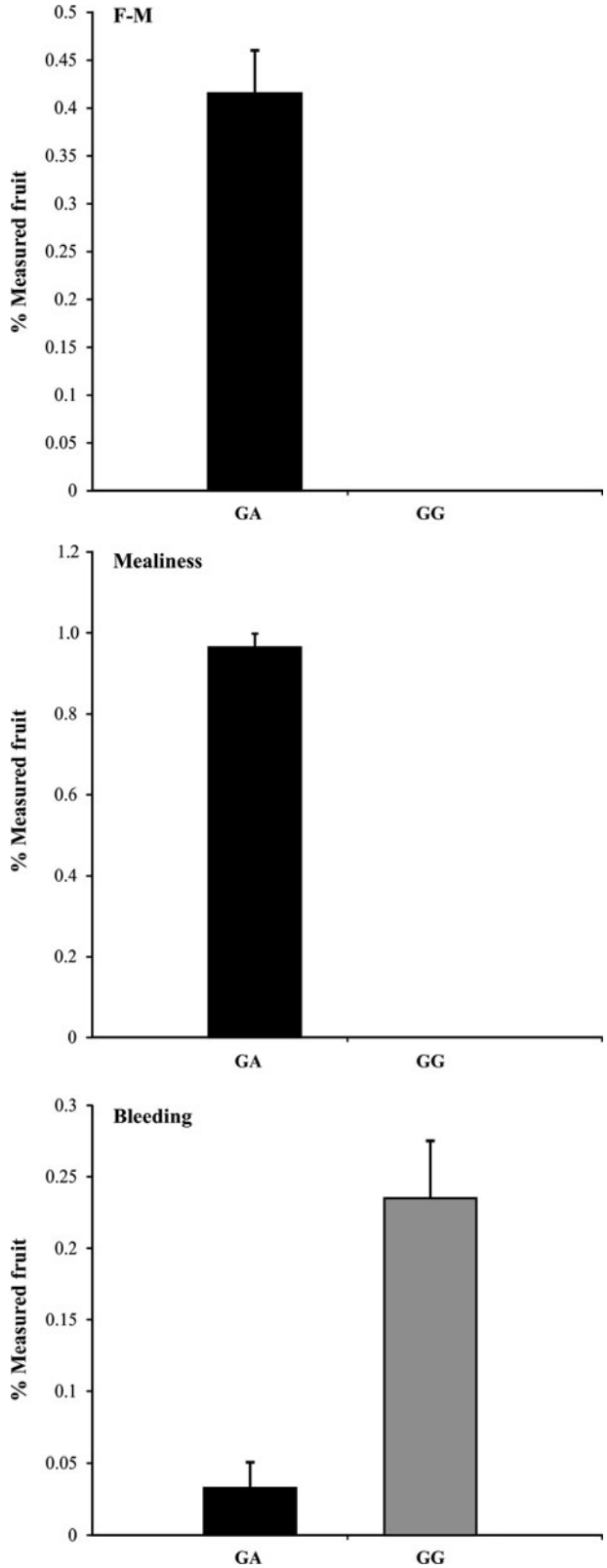

Fig. 4 Freestone-melting flesh (F-M), mealiness and bleeding symptoms distribution associated with the genotype obtained by SNP marker UCD_SNP_1084. Georgia belle (freestone-melting flesh cultivar) 'GA' (susceptible) and Dr Davis (clingstone, non-melting flesh cultivar) presented a 'GG' (resistant)

to follow the changes in phenotype (particular symptom-protein changes) associated with differences in genotype. This QTN may facilitate the characterization metabolic pathways and improve protocols for phenotyping by identification of biochemical changes affecting structure or availability of substrates, with subsequent application to peach breeding programs. A good example of this was presented by Fournier-Level et al. (2009) in grape. Trait identification, validation, studies of inheritance, stability, and mechanisms through which the trait is affected by QTNs, may improve genetic gain for these characters in breeding programs (Weller and Ron 2011).

Validated QTNs can serve as screening markers for peach germplasm, as the assumption of linkage disequilibrium (LD) between the marker and the QTL is no longer needed for selection, which makes selection at the genome level more efficient. This is important since LD relationships change over time and differs among populations (even among clusters of material in structured breeding programs), which affects precision of selection. Therefore, application of these QTNs eliminates the need for Genome-Wide Association Studies (GWAS) and permits direct application of Marker-assisted selection. If a marker (such as a SNP or microsatellite) is in LD, the allelic frequencies will not accurately reflect the allelic frequencies of the QTN in the breeding population. Thus, when QTN information is applied to breeding programs, in which just a portion of the population is genotyped, the QTN (statistical) effects can be treated as fixed, which reduces the bias in the analyses. More research is needed to determine the potential of this particular QTN for selection in peach breeding programs and to determine how close the SNP is to the functional gene component.

Only one predicted effect was observed for the SNP marker UCD_SNP_1441. This SNP was associated with one putative uncharacterized protein found in grape (*Vitis vinifera*), but there is no information about its function.

The four SNPs (UCD_SNP_821, UCD_SNP_641, UCD_SNP_1347, and UCD_SNP_1067), associated with QTL for flesh bleeding (qBL1) (Fig. 1), were associated with effects in proteins such as putative protein phosphatase 2C 13, dihydrolipoyl dehydrogenase 1 (mitochondrial), TMV resistance protein N, probable protein Pop3, and some unknown proteins. These proteins are related to plant defense responses, cell redox homeostasis, and protein dephosphorylation. In addition to the QTL on G4, several candidate genes for flesh bleeding were mapped on G1, previously (Ogundiwin et al. 2009). One of these genes PME3 (pectin methylesterase) can be related to the probable protein Pop3. Probable protein Pop3 (locus XP_002280639) in *Vitis vinifera* was predicted by automated computational analysis and derived from a genomic sequence (NW_003724019), annotated using the gene

prediction method: GNOMON, supported by mRNA and EST evidence), and showed a similarity of 78.9 % with putative protein Pop3 (HS1). It also has a 100 % similarity with a plant invertase/pectin methylesterase inhibitor domain-containing protein (locus NP_683571.1) in *Arabidopsis thaliana* (Salanoubat et al. 2000).

Flesh bleeding could be related to phenolic metabolism (C. Crisosto personal communication), however no relationship has previously been observed between the proteins described here and response to fruit quality or storage cold stress.

Proteins associated with candidate SNPs related to flesh browning

Flesh browning SNPs were associated with different functional proteins for enzyme inhibitor activity, DNA binding, RNA binding (UCD_SNP_265), metal ion binding, structural constituent of ribosome (UCD_SNP_1295), transcription factor activity (UCD_SNP_1422), ATP binding, glutamate-cysteine ligase activity, and structural molecule activity (UCD_SNP_240). The marker UCD_SNP_46, a marker that belongs to the haplotype

Table 3 The predicted effects for each SNP from the haplotype block associated with flesh browning

SNP (Peach genome position)	Effect	Homologous gene	Reference organism	Description
UCD_SNP_46 (scaffold_5:2042442)	SYNONYMOUS_CODING	AOX1A_ARATH	*Arabidopsis thaliana*	Alternative oxidase 1a, mitochondrial
UCD_SNP_243 (scaffold_5:2196116)	SYNONYMOUS_CODING	D1IVR9_VITVI	*Vitis vinifera*	Whole genome shotgun sequence of line PN40024, scaffold_33.assembly12x
UCD_SNP_342 (scaffold_5:2043034)	DOWNSTREAM: 330 bases	AOX1A_ARATH	*Arabidopsis thaliana*	Alternative oxidase 1a, mitochondrial
UCD_SNP_508 (scaffold_5:2245360)	UPSTREAM: 4594 bases	M310_ARATH	*Arabidopsis thaliana*	Uncharacterized mitochondrial protein AtMg00310
UCD_SNP_508 (scaffold_5:2245360)	UPSTREAM: 244 bases	Q9FW98_ORYSJ	*Oryza sativa* L. ssp. *japonica*	Putative non-LTR retroelement reverse transcriptase
UCD_SNP_732 (scaffold_5:2236495)	INTRON	AT4G19140.1	*Arabidopsis thaliana*	Unknown protein
UCD_SNP_732 (scaffold_5:2236495)	DOWNSTREAM: 2585 bases	M310_ARATH	*Arabidopsis thaliana*	Uncharacterized mitochondrial protein AtMg00310
UCD_SNP_905 (scaffold_5:2106547)	UPSTREAM: 1173 bases	FH14_ARATH	*Arabidopsis thaliana*	Formin-like protein 14
UCD_SNP_1173 (scaffold_5:2185638)	DOWNSTREAM: 3868 bases	Y1691_LEGPH	*Legionella pneumophila* ssp. *pneumophila*	Uncharacterized transporter lpg1691
UCD_SNP_1173 (scaffold_5:2185638)	DOWNSTREAM: 1470 bases	Y381_RICFE	*Rickettsia felis* (*Rickettsia azadi*)	Putative ankyrin repeat protein RF_0381
UCD_SNP_1247 (scaffold_5:2094400)	DOWNSTREAM: 1588 bases	FH14_ARATH	*Arabidopsis thaliana*	Formin-like protein 14
UCD_SNP_1247 (scaffold_5:2094400)	SYNONYMOUS_CODING	DUR31_SCHPO	*Schizosaccharomyces pombe*	Probable urea active transporter 1
UCD_SNP_1257 (scaffold_5:2208207)	INTERGENIC			
UCD_SNP_1313 (scaffold_5:2005159)	INTRON	RECQ1_HUMAN	*Homo sapiens*	ATP-dependent DNA helicase Q1
UCD_SNP_1424 (scaffold_5:2024428)	INTRON	RECQ1_RAT	*Rattus norvegicus* (Rat)	ATP-dependent DNA helicase Q1
UCD_SNP_1424 (scaffold_5:2024428)	UPSTREAM: 4187 bases	AT2G13770.1	*Arabidopsis thaliana*	Unknown protein
UCD_SNP_1492 (scaffold_5:2198771)	NON_SYNONYMOUS_CODING	D1IVR9_VITVI	*Vitis vinifera*	Whole genome shotgun sequence of line PN40024, scaffold_33.assembly12x

The location on the peach genome (scaffold 5), the homologous gene, the reference organism and description of the protein associated with known protein coding regions are shown in the table

found for flesh browning (Table 3), was associated with AOX1A [Alternative Oxidase 1a (mitochondrial)]. This protein has been associated with TCA cycle stability, suppression of reactive oxygen species, and with a role in maintaining energy homeostasis (Crichton et al. 2005).

The location of UCD_SNP_46 (variant) was identified by SnpEff as the exon of AOX1A and is a synonomous coding change in the protein, resulting in a low functional impact. The function of AOX1A was related "by similarity" to other protein family members (from UniProtKB/Swiss-Prot database) that catalyze cyanide-resistant oxygen consumption and may increase respiration when the cytochrome respiratory pathway is restricted or in response to low temperatures. In potato tubers, cold stress stimulates the AOX activity and the enzyme can serve a protective function against cell damage in response to reactive oxygen species production (Pinheiro et al. 2004). In addition AOX1A was mapped to the central segment of chromosome VIII in potato for which the gene PPO (polyphenol oxidase) was previously described (Werij et al. 2007).

The candidate gene RHB1 was up-regulated in cold treated mesocarp tissue (Ogundiwin et al. 2008) and mapped to linkage group 2. It is co-located with one minor browning QTL reported by Ogundiwin et al. (2009). However, RHB1was not associated with flesh browning in this study. The browning potential of peaches depends on the total amount of phenolic compounds present in the fruit and the level of activity of polyphenol oxidase (PPO). Normally, the phenols and polyphenol oxidase are located in separate compartments in the cell (Kader and Chordas 1984). Flesh browning damage expression is related to cell membrane leakage due to damage or senescence, which leads to changes in membrane permeability and merging of phenolic compounds (substrate) and the PPO (catalyzer) in the cell, producing a browning reaction. The interaction between both proteins, AOX1A and PPO, must be studied in greater depth to clarify the antioxidant attribute related to this symptom in peach.

No SNP associations were observed to leucoanthocyanidin dioxygenase (PpLDOX) or to an enzyme in the anthocyanin biosynthesis pathway that is co-located with a major QTL controlling cold-induced flesh browning in Pop-DG (Ogundiwin et al. 2008). However, the SNP marker UCD_SNP_1295 was associated with cytochrome P450 72A1 (secologanin synthase), conditioning for an oxidoreductase activity (acting on paired donors with incorporation or reduction of molecular oxygen). Both leucoanthocyanidin dioxygenase (locus ppa007738 m.g; scaffold_5: 9817673–9819380) and cytochrome P450 72A1 could affect the same function. These results suggested that UCD_SNP_1295 is a possible candidate SNP marker to use in marker-assisted breeding.

Conclusions

The combination of classical QTL analysis, genomic annotation resources in peach, and bioinformatics tools allowed us to obtain an extended list of new candidate genes that can contribute to the understanding of important quality traits in peach. In a woody fruit species, such as other *Prunus* species, breeding strategies have been limited by long reproductive cycles characterized by long juvenile periods, complex reproductive biology, and a high degree of heterozygosity. The information obtained through the study of variants in the peach genome can be used to determine effects of the observed alleles on transcripts. This information may aid selection of specific variants for genotyping studies and permit the discovery of new candidate genes for fruit quality as well as other biologically significant loci. Candidate genes, functional genes, or modifiers, can be very precisely identified with tools such as SnpEff, and then marked and transferred via either conventional or alternative breeding strategies into new genetic backgrounds for cultivar development.

The results presented in this paper are part of a larger project that began with the sequencing of specific peach cultivars, followed by development of SNP markers in breeding populations segregating for important quality traits, construction of high-density SNP maps, analysis of QTLs, and determination of genome regions controlling fruit quality characteristics. These results may also be translated to candidate gene discovery in other stone fruits such as plum or apricot.

Acknowledgments We gratefully acknowledge the support of the National Research Initiative of USDA's National Institute of Food and Agriculture (NIFA) grant # 2008-35300-04432 and US-Israel Binational Agriculture Research and Development Fund (BARD) Grant no. US-4027-07 for provided financial support to this project, as well as UC Davis, UC Agricultural Experiment Station, USDA-CREES (Hatch Experiment Station funding) and CONACYT-UCMEXUS, which provides a PhD fellowship to J. Fresnedo-Ramírez. We would especially like to thank Pablo Cingolani, developer of SnpEff at McGill University, and Dr. María Luisa Badenes for very useful comments on a previous draft. We thank the International Peach Genome Initiative (IPGI) led by Drs Bryon Sosinski, Ignazio Verde and Daniel Rokhsar for making publicly available the draft of assembled and annotated peach genome (peach v1.0), gene annotation and homology files through JGI, GDR and IGA.

References

Ahmad R, Parfitt DE, Fass J, Ogundiwin E, Dhingra A, Gradziel TM, Lin D, Joshi NA, Martínez-García PJ, Crisosto CH (2011) Whole genome sequencing of peach (*Prunus persica* L.) for SNP identification and selection. BMC Genomics 12:569. doi: 10.1186/1471-2164-12-569

Appleby N, Edwards D, Batley J (2009) New technologies for ultra-high throughput genotyping in plants. Methods Mol Biol 513:19–39. doi:10.1007/978-1-59745-427-8_2

Aranzana MJ, Illa E, Howad W (2012) A first insight into peach [*Prunus persica* (L.) Batsch] SNP variability. Tree Genet Genomes. doi:10.1007/s11295-012-0523-6

Arulsekar S, Parfitt DE, Kester DE (1986) Comparison of isozyme variability in peach and almond cultivars. J Hered 77:272–274

Arus P, Verde I, Sosinski B, Zhebentyayeva T, Abbott AG (2012) The peach genome. Tree Genet Genomes 8(3):531–547. doi:10.1007/S11295-012-0493-8

Brookes AJ (1999) The essence of SNPs. Gene 234(2):177–186

Cantin CM, Crisosto CH, Ogundiwin EA, Gradziel T, Torrents J, Moreno MA, Gogorcena Y (2010) Chilling injury susceptibility in an intraspecific peach [*Prunus persica* (L.) Batsch] progeny. Postharvest Biol Technol 58(2):79–87. doi:10.1016/J.Postharvbio.2010.06.002

Castle JC (2011) SNPs occur in regions with less genomic sequence conservation. PLoS ONE 6(6). doi:10.1371/journal.pone.0020660

Cingolani P, Platts A, le Wang L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w (1118); iso-2; iso-3. Fly 6(2):80–92. doi:10.4161/fly.19695

Collins DW, Jukes TH (1994) Rates of transition and transversion in coding sequences since the human-rodent divergence. Genomics 20(3):386–396

Cooper DN, Krawczak M (1989) Cytosine methylation and the fate of Cpg dinucleotides in vertebrate genomes. Hum Genet 83(2):181–188

Coulondre C, Miller JH, Farabaugh PJ, Gilbert W (1978) Molecular-basis of base substitution hotspots in *Escherichia coli*. Nature 274(5673):775–780

Crichton PG, Affourtit C, Albury MS, Carre JE, Moore AL (2005) Constitutive activity of Sauromatum guttatum alternative oxidase in *Schizosaccharomyces pombe* implicates residues in addition to conserved cysteines in alpha-keto acid activation. FEBS Lett 579(2):331–336. doi:10.1016/j.febslet.2004.10.107

Dagar A, Pons Puig C, Marti Ibanez C, Ziliotto F, Bonghi C, Crisosto CH, Friedman H, Lurie S, Granell A (2012) Comparative transcript profiling of a peach and its nectarine mutant at harvest reveals differences in gene expression related to storability. Tree Genet Genomes. doi:10.1007/s11295-012-0549-9

Dirlewanger E, Graziano E, Joobeur T, Garriga-Caldere F, Cosson P, Howad W, Arús P (2004) Comparative mapping and marker-assisted selection in Rosaceae fruit crops. P Natl Acad Sci USA 101(26):9891–9896. doi:10.1073/pnas.0307937101

Duran C, Appleby N, Edwards D, Batley J (2009) Molecular genetic markers: discovery, applications, data storage and visualisation. Curr Bioinform 4(1):16–27

Eduardo I, Chietera G, Pirona R, Pacheco I, Troggio M, Banchi E, Bassi D, Rossini L, Vecchietti A, Pozzi C (2012) Genetic dissection of aroma volatile compounds from the essential oil of peach fruit: QTL analysis and identification of candidate genes using dense SNP maps. Tree Genet Genomes. doi:10.1007/s11295-012-0546-z

Espinoza C, Vega A, Medina C, Schlauch K, Cramer G, Arce-Johnson P (2007) Gene expression associated with compatible viral diseases in grapevine cultivars. Funct Integr Genomics 7(2):95–110. doi:10.1007/S10142-006-0031-6

Font i Forcada C, Oraguzie N, Igartua E, Moreno M, Gogorcena Y (2012) Population structure and marker–trait associations for pomological traits in peach and nectarine cultivars. Tree Genet Genomes 1–19. doi:10.1007/s11295-012-0553-0

Fournier-Level A, Le Cunff L, Gomez C, Doligez A, Ageorges A, Roux C, Bertrand Y, Souquet JM, Cheynier V, This P (2009) 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 183(3):1127–1139. doi:10.1534/Genetics.109.103929

Gruenbaum Y, Navehmany T, Cedar H, Razin A (1981) Sequence specificity of methylation in higher-plant DNA. Nature 292(5826):860–862

Hu TT, Pattyn P, Bakker EG, Cao J, Cheng JF, Clark RM, Fahlgren N, Fawcett JA, Grimwood J, Gundlach H, Haberer G, Hollister JD, Ossowski S, Ottilar RP, Salamov AA, Schneeberger K, Spannagl M, Wang X, Yang L, Nasrallah ME, Bergelson J, Carrington JC, Gaut BS, Schmutz J, Mayer KFX, de Peer YV, Grigoriev IV, Nordborg M, Weigel D, Guo YL (2011) The Arabidopsis lyrata genome sequence and the basis of rapid genome size change. Nat Genet 43(5):476–481. doi:10.1038/Ng.807

Kader AA, Chordas A (1984) Evaluation the browning potential of peaches [cultivar screening]. Calif Agric 38(3–4):14–15

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide-sequences. J Mol Evol 16(2):111–120

Koboldt DC, Chen K, Wylie T, Larson DE, McLellan MD, Mardis ER, Weinstock GM, Wilson RK, Ding L (2009) VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics 25(17):2283–2285. doi:10.1093/Bioinformatics/Btp373

Kraus RHS, Kerstens HHD, Van Hooft P, Crooijmans RPMA, Van der Poel JJ, Elmberg J, Vignal A, Huang YH, Li N, Prins HHT, Groenen MAM (2011) Genome wide SNP discovery, analysis and evaluation in mallard (*Anas platyrhynchos*). BMC Genomics 12. doi:10.1186/1471-2164-12-150

Leida C, Terol J, Martc G, Agustc M, LlcLcer G, Badenes ML, Rcos G (2010) Identification of genes associated with bud dormancy release in *Prunus persica* by suppression subtractive hybridization. Tree Physiol 30(5):655–666

Li YZ, Wu BJ, Yu YL, Yang GD, Wu CG, Zheng CC (2011) Genome-wide analysis of the RING finger gene family in apple. Mol Genet Genomics 286(1):81–94. doi:10.1007/S00438-011-0625-0

Ma K, Mao JH, Li XH, Zhang QF, Lian XM (2009) Sequence and expression analysis of the C3HC4-type RING finger gene family in rice. Gene 444(1–2):33–45. doi:10.1016/J.Gene.2009.05.018

Mackay TFC, Stone EA, Ayroles JF (2009) The genetics of quantitative traits: challenges and prospects. Nat Rev Genet 10(8):565–577. doi:10.1038/Nrg2612

Martínez-García PJ, Parfitt D, Ogundiwin E, Fass J, Chan H, Ahmad R, Lurie S, Dandekar A, Gradziel T, Crisosto C (2012a) High density SNP mapping and QTL analysis for fruit quality characteristics in peach (*Prunus persica* L.). Tree Genet Genomes 1–18. doi:10.1007/s11295-012-0522-7

Martínez-García PJ, Cameron P, Parfitt D, Ogundiwin E, Fresnedo-Ramírez J, Dandekar A, Gradziel T, Crisosto C (2012b) Influence of year and genetic factors on chilling injury susceptibility in peach (*Prunus persica* (L.) Batsch). Euphytica 185(2):267–280

Morton BR, Bi IV, McMullen MD, Gaut BS (2006) Variation in mutation dynamics across the maize genome as a function of regional and flanking base composition. Genetics 172(1):569–577. doi:10.1534/Genetics.105.049916

Ogundiwin EA, Marti C, Forment J, Pons C, Granell A, Gradziel TM, Peace CP, Crisosto CH (2008) Development of ChillPeach genomic tools and identification of cold-responsive genes in peach fruit. Plant Mol Biol 68(4–5):379–397. doi:10.1007/S11103-008-9378-5

Ogundiwin EA, Peace CP, Gradziel TM, Parfitt DE, Bliss FA, Crisosto CH (2009) A fruit quality gene map of Prunus. BMC Genomics 10. doi:10.1186/1471-2164-10-587

Picoult-Newberg L, Ideker TE, Pohl MG, Taylor SL, Donaldson MA, Nickerson DA, Boyce-Jacino M (1999) Mining SNPs from EST databases. Genome Res 9(2):167–174

Pinheiro C, Passarinho JA, Ricardo CP (2004) Effect of drought and rewatering on the metabolism of Lupinus albus organs. J Plant Physiol 161(11):1203–1210. doi:10.1016/J.Jplph.2004.01.016

Ramakrishnan AP, Meyer SE, Waters J, Stevens MR, Coleman CE, Fairbanks DJ (2004) Correlation between molecular markers and adaptively significant genetic variation in *Bromus tectorum* (Poaceae) an inbreeding annual grass. Am J Bot 91(6):797–803

Reumers J, Maurer-Stroh S, Schymkowitz J, Rousseau F (2006) SNPeffect v.20: a new step in investigating the molecular phenotypic effects of human non-synonymous SNPs. Bioinformatics 22(17):2183–2185. doi:10.1093/Bioinformatics/Btl348

Salanoubat M, Lemcke K, Rieger M, Ansorge W, Unseld M, Fartmann B, Valle G, Blöcker H, Perez-Alonso M, Obermaier B, Delseny M, Boutry M, Grivell LA, Mache R, Puigdomènech P, De Simone V, Choisne N, Artiguenave F, Robert C, Brottier P, Wincker P, Cattolico L, Weissenbach J, Saurin W, Quétier F, Schäfer M, Müller-Auer S, Gabel C, Fuchs M, Benes V, Wurmbach E, Drzonek H, Erfle H, Jordan N, Bangert S, Wiedelmann R, Kranz H, Voss H, Holland R, Brandt P, Nyakatura G, Vezzi A, D'Angelo M, Pallavicini A, Toppo S, Simionati B, Conrad A, Hornischer K, Kauer G, Löhnert TH, Nordsiek G, Reichelt J, Scharfe M, Schön O, Bargues M, Terol J, Climent J, Navarro P, Collado C, Perez-Perez A, Ottenwälder B, Duchemin D, Cooke R, Laudie M, Berger-Llauro C, Purnelle B, Masuy D, de Haan M, Maarse AC, Alcaraz JP, Cottet A, Casacuberta E, Monfort A, Argiriou A, Flores M, Liguori R, Vitale D, Mannhaupt G, Haase D, Schoof H, Rudd S, Zaccaria P, Mewes HW, Mayer KF, Kaul S, Town CD, Koo HL, Tallon LJ, Jenkins J, Rooney T, Rizzo M, Walts A, Utterback T, Fujii CY, Shea TP, Creasy TH, Haas B, Maiti R, Wu D, Peterson J, Van Aken S, Pai G, Militscher J, Sellers P, Gill JE, Feldblyum TV, Preuss D, Lin X, Nierman WC, Salzberg SL, White O, Venter JC, Fraser CM, Kaneko T, Nakamura Y, Sato S, Kato T, Asamizu E, Sasamoto S, Kimura T, Idesawa K, Kawashima K, Kishida Y, Kiyokawa C, Kohara M, Matsumoto M, Matsuno A, Muraki A, Nakayama S, Nakazaki N, Shinpo S, Takeuchi C, Wada T, Watanabe A, Yamada M, Yasuda M, Tabata S (2000) Sequence and analysis of chromosome 3 of the plant *Arabidopsis thaliana*. Nature 408(6814):820–823

SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL (1998) The paleontology of intergene retrotransposons of maize. Nat Genet 20(1):43–45

Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K (2001) dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 29(1):308–311

Strandberg AKK, Salter LA (2004) A comparison of methods for estimating the transition: transversion ratio from DNA sequences. Mol Phylogenet Evol 32(2):495–503. doi:10.1016/J.Ympev.2004.01.013

Tittarelli A, Santiago M, Morales A, Meisel LA, Silva H (2009) Isolation and functional characterization of cold-regulated promoters, by digitally identifying peach fruit cold-induced genes from a large EST dataset. BMC Plant Biol 9. doi:10.1186/1471-2229-9-121

van Tienderen PH, de Haan AA, van der Linden CG, Vosman B (2002) Biodiversity assessment using markers for ecologically important traits. Trends Ecol Evol 17(12):577–582

Verde I, Bassil N, Scalabrin S, Gilmore B, Lawley CT, Gasic K, Micheletti D, Rosyara UR, Cattonaro F, Vendramin E, Main D, Aramini V, Blas AL, Mockler TC, Bryant DW, Wilhelm L, Troggio M, Sosinski B, Aranzana MJ, Arus P, Iezzoni A, Morgante M, Peace C (2012) Development and evaluation of a 9 K SNP array for peach by internationally coordinated SNP detection and validation in breeding germplasm. PLoS ONE 7(4):e35668. doi:10.1371/journal.pone.0035668

Vignal A, Milan D, SanCristobal M, Eggen A (2002) A review on SNP and other types of molecular markers and their use in animal genetics. Genet Sel Evol 34(3):275–305. doi:10.1051/Gse:2002009

Vitte C, Bennetzen JL (2006) Analysis of retrotransposon structural diversity uncovers properties and propensities in angiosperm genome evolution. Proc Natl Acad Sci USA 103(47):17638–17643. doi:10.1073/Pnas.0605618103

Wakeley J (1996) The excess of transitions among nucleotide substitutions: new methods of estimating transition bias underscore its significance. Trends Ecol Evol 11(4):158–163. doi:10.1016/0169-5347(96)10009-4

Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris MS, Shen NP, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ, Lipshutz R, Chee M, Lander ES (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280(5366):1077–1082

Weller JI, Ron M (2011) Invited review: quantitative trait nucleotide determination in the era of genomic selection. J Dairy Sci 94(3):1082–1090. doi:10.3168/jds.2010-3793

Werij JS, Kloosterman B, Celis-Gamboa C, de Vos CHR, America T, Visser RGF, Bachem CWB (2007) Unravelling enzymatic discoloration in potato through a combined approach of candidate genes, QTL, and expression analysis. Theor Appl Genet 115(2):245–252. doi:10.1007/S00122-007-0560-Y

Zamboni A, Di Carli M, Guzzo F, Stocchero M, Zenoni S, Ferrarini A, Tononi P, Toffali K, Desiderio A, Lilley KS, Pe ME, Benvenuto E, Delledonne M, Pezzotti M (2010) Identification of putative stage-specific grapevine berry biomarkers and omics data integration into networks. Plant Physiol 154(3):1439–1459. doi:10.1104/Pp.110.160275

Zhang J, Jia W, Yang J, Ismail AM (2006) Role of ABA in integrating plant responses to drought and salt stresses. Field Crop Res 97(1):111–119. doi:10.1016/j.fcr.2005.08.018