

QTL involved in the modification of cyanidin compounds in black and red raspberry fruit

J. M. Bushakra · C. Krieger · D. Deng · M. J. Stephens ·
A. C. Allan · R. Storey · V. V. Symonds · D. Stevenson ·
T. McGhie · D. Chagné · E. J. Buck · S. E. Gardiner

Received: 8 August 2012 / Accepted: 15 November 2012 / Published online: 9 December 2012
© Springer-Verlag Berlin Heidelberg 2012

Abstract Fruit from *Rubus* species are highly valued for their flavor and nutritive qualities. Anthocyanin content contributes to these qualities, and although many studies have been conducted to identify and quantify the major anthocyanin compounds from various *Rubus* species, the genetic control of the accumulation of these complex traits in *Rubus* is not yet well understood. The identification of

the regions of the genome involved in the production of anthocyanins is an important first step in identifying the genes underlying their expression. In this study, ultra and high-performance liquid chromatography (UHPLC and HPLC) and two newly developed *Rubus* linkage maps were used to conduct QTL analyses to explore the presence of associations between concentrations of five anthocyanins in fruit and genotype. In total, 27 QTL were identified on the *Rubus* linkage maps, four of which are associated with molecular markers designed from transcription factors and three of which are associated with molecular markers designed from anthocyanin biosynthetic pathway candidate

Communicated by H. Nybom.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-012-2022-4) contains supplementary material, which is available to authorized users.

J. M. Bushakra · C. Krieger · T. McGhie · D. Chagné (✉) ·
E. J. Buck · S. E. Gardiner
The New Zealand Institute for Plant & Food
Research Limited (PFR), Batchelar Road,
Private Bag 11600, Palmerston North 4442, New Zealand
e-mail: David.Chagne@plantandfood.co.nz

C. Krieger
e-mail: Celia.Krieger@ensaia.inpl-nancy.fr

T. McGhie
e-mail: Tony.McGhie@plantandfood.co.nz

E. J. Buck
e-mail: Emily.Buck@plantandfood.co.nz

S. E. Gardiner
e-mail: Sue.Gardiner@plantandfood.co.nz

J. M. Bushakra · V. V. Symonds
Massey University, Institute of Molecular Biosciences,
Private Bag 11222, Palmerston North 4442, New Zealand
e-mail: V.V.Symonds@massey.ac.nz

Present Address:

J. M. Bushakra
United States Department of Agriculture, Agricultural Research
Service, National Clonal Germplasm Repository,
33447 Peoria Road, Corvallis, OR 97333, USA

Present Address:

C. Krieger
Ecole National Supérieure d'Agronomie et des Industries
Alimentaires, 2 Avenue de la Forêt de Haye BP 172,
54505 Vandoeuvre les Nancy Cedex, France

D. Deng · D. Stevenson
PFR, East Street, Private Bag 3230, Hamilton 3214,
New Zealand
e-mail: Dawei.Deng@plantandfood.co.nz

D. Stevenson
e-mail: David.Stevenson@plantandfood.co.nz

M. J. Stephens
PFR, Old Mill Road, RD3, Motueka 7198, New Zealand
e-mail: Jo.Stephens@plantandfood.co.nz

A. C. Allan · R. Storey
PFR, Mt Albert Research Centre, Private Bag 92169,
Auckland 1142, New Zealand
e-mail: Andrew.Allan@plantandfood.co.nz

R. Storey
e-mail: Roy.Storey@plantandfood.co.nz

genes. The results of this study suggest that, while QTL for anthocyanin accumulation have been identified on six of seven *Rubus* linkage groups (RLG), the QTL on RLG2 and RLG7 may be very important for genetic control of cyanidin modification in *Rubus*.

Introduction

One of the largest and most diverse of plant genera, *Rubus* contains between 600 and 800 species, including blackberry and red and black raspberry, is distributed worldwide (Thompson 1995) and has a long history of use for food and medicinal purposes (reviewed by Hummer 2010). Biologically active (bioactive) compounds found in red and black raspberry have been extracted for individual evaluation, in either animal models or *in vitro* assays, for their health-promoting characteristics including the reduction of cell proliferation in rat models (Chen et al. 2011), anti-inflammatory effect on ulcers (Montrose et al. 2011), and reduction of blood cholesterol levels (Ash et al. 2011). These reports, and others, have led to the marketing of bioactive compounds as ingredients in “functional foods”, that is products reported to have a specific health benefit (Espín et al. 2007; Finley et al. 2011).

Anthocyanins (ACY) are a class of water-soluble flavonoids that are derived from phenylalanine and stored in the plant vacuole. They provide plant tissues with a range of colors from orange/red to violet/blue, are widely distributed in flowering plants (Holton and Cornish 1995; Tanaka et al. 2008), and are among the bioactive compounds found in *Rubus* plants. Although the ACY profiles of cultivated red raspberry (*Rubus idaeus* L.), black raspberry (*R. occidentalis* L.), tetraploid blackberry and hybrid berry (*Rubus* sp.) fruit are the best characterized of the genus (Connor et al. 2005; Dossett et al. 2010; Jennings and Carmichael 1980; McGhie et al. 2002; Scalzo et al. 2008; Torre and Barritt 1977), the ACY biosynthetic pathway has yet to be examined in *Rubus*. However, the biosynthetic pathway of ACY production appears to be well conserved in *Arabidopsis thaliana*, maize, snapdragon, and petunia (reviewed by Holton and Cornish 1995), as well as in apple (Takos et al. 2006), and many of the genes encoding the enzymes involved are conserved in both sequence and function (reviewed by Grotewold 2006; Holton and Cornish 1995; Tanaka et al. 2008).

Several studies have been conducted to determine the predominant ACY compounds in *Rubus* species (Connor et al. 2005; Cooney et al. 2004; Dossett et al. 2010, 2011; Jennings and Carmichael 1980; Määttä-Riihinen et al. 2004; Ozgen et al. 2008; Seeram et al. 2001; Torre and Barritt 1977; Tulio et al. 2008). Black raspberry fruits are characterized by two xylose-containing pigments, cyanidin

3-*O*-sambubioside (C3Sb) and cyanidin 3-*O*-2^G-xylosylrutinoside (C3XR), and overall high ACY concentrations (200 mg/100 g fruit) (Ozgen et al. 2008; Tulio et al. 2008); whereas red raspberry fruits are characterized by the presence of cyanidin 3-*O*-sophoroside (C3S), cyanidin 3-*O*-2^G-glucosylrutinoside (C3GR), and overall lower ACY concentrations (60 mg/100 g fruit) (Jennings and Carmichael 1980; Torre and Barritt 1977). Other ACY compounds found in *Rubus* fruits are cyanidin 3-*O*-rutinoside (C3R), cyanidin 3-*O*-glucoside (C3G), and pelargonidin 3-*O*-rutinoside (P3R) (Dossett et al. 2011; Kassim et al. 2009; Torre and Barritt 1977). Each of the cyanidin-containing compounds is a variation on the molecule cyanidin 3-*O*-glucoside, resulting from the enzyme-catalyzed addition of different sugar moieties, either glucose, rhamnose or xylose (Grotewold 2006; Holton and Cornish 1995). Pelargonidin 3-*O*-rutinoside (P3R), and cyanidin 3-*O*-glucoside (C3G) have similar structures, with the basic pelargonidin and cyanidin three-ring molecule differing by the presence of a second hydroxyl on the cyanidin B-ring 3'-carbon, the addition of which is catalyzed by flavanone 3' hydroxylase (F3'H) early in the biosynthetic pathway (Grotewold 2006; Holton and Cornish 1995).

Four previous studies have been conducted to detect QTL in a population of *Rubus idaeus* ‘Latham’ × ‘Glen Moy’, both red-fruited cultivars. An assessment of fruit development from bud break to ripeness identified QTL involved in stages of fruit ripening and cane height associated with regions on linkage groups (LG) 2, 3, 5 and 6 (Graham et al. 2009). Two QTL involved in the variation of the accumulation of eight specific ACY compounds were mapped to regions of LG 1 and 4 (Kassim et al. 2009). Two candidate genes were strongly associated with the QTL mapped by Kassim et al. (2009), a putative *bHLH* transcription factor on LG1 and a putative *bZIP* transcription factor on LG4. QTL involved in the variation in fruit color during ripening and the accumulation of total ACY were mapped to LG 2, 3, 4, and 6 (McCallum et al. 2010). Candidate genes associated with the QTL for total ACY accumulation mapped by McCallum et al. (2010) were a putative *TIP* small molecule transporter on LG2; and *dihydroflavonol 4-reductase* (*DFR*), along with *flavonol synthase* (*FLS*) both on LG4. Each of these studies identified unique QTL, although the QTL on LG3 show some overlap, and the QTL associated with marker Rub-FruitE4 on LG4 are in common between the studies of McCallum et al. (2010), and Kassim et al. (2009), although for different traits. The most recent study assesses QTL associated with total ACY content along with other fruit quality traits (Dobson et al. 2012). This study identified two QTL associated with total ACY on LG2 and LG6 with the QTL on LG2 associated with the candidate *TIP* transcription factor that is also underlying the QTL mapped by

McCallum et al. (2010). The linkage group nomenclature used in these four studies follows Graham et al. (2004). These four studies provide insight into QTL of several fruit quality traits in a single population of red raspberry progeny over several years, as well as in different locations and growing conditions.

New genetic linkage maps constructed from a progeny of black raspberry (*Rubus occidentalis*) 96395S1 (deep purple fruit), and red raspberry (*R. idaeus*) ‘Latham’ (red fruit), have been developed recently (96395S1 × ‘Latham’) (Bushakra et al. 2012). These maps, deliberately enriched for markers derived from orthologous sequences and candidate genes coding for enzymes involved in the polyphenolic biosynthetic pathway, allow the association of candidate genes with QTL for variations in amounts of fruit ACY compounds in a new population and previously unstudied species of *Rubus*. Improving our understanding of the genetics controlling the production of and interactions between levels of ACY compounds will assist in the development of marker-assisted breeding (MAB) to improve the efficiency of breeding new *Rubus* cultivars with the most beneficial balance of compounds for human health. To this end, fruit ACY phenotype data were collected over three seasons and analyzed for marker associations using the ‘Latham’ (red raspberry) and 96395S1 (black raspberry) genetic linkage maps (Bushakra et al. 2012) to identify QTL.

Table 1 Derivation of markers mapped on *Rubus occidentalis* 96395S1 × *R. idaeus* ‘Latham’ that were designed from transcription factors and genes for enzymes involved in the polyphenolic biosynthetic pathway (Bushakra et al. 2012)

Marker name	Sequence source species	Accession	Putative gene function
EMFxaCAD1B ^a	Strawberry	AF320110	Cinnamyl alcohol dehydrogenase
Ri4CCL3	Raspberry	AF239685.1	4-Coumarate: coA ligase 3
Ri4Co1	Raspberry	AF239687	4-Coumarate: coA ligase 1
RiANR	Strawberry	DQ664193	Anthocyanidin reductase
RiANS	Raspberry	AY695818	Anthocyanidin synthase
RiCHI	Apple	CN946541	Chalcone isomerase
RiDFR	Strawberry	AY695813.1	Dihydroflavonol 4-reductase
RiF3H	Raspberry	EU078685	Flavanone-3-hydroxylase
RiF3'H	Apple	FJ919631	Flavonoid 3' hydroxylase
RiFLS	Strawberry	DQ087252.1	Flavonol synthase
RiGT	Apple	EB124403	UDP-glucosyltransferase
RiHCT/HQT	Apple	NP_199704	Hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase
RiPAL2	Raspberry	AF237955.1	Phenylalanine ammonia-lyase 2
RiUFGT	Strawberry	AY575056	UDP glucose:flavonoid-3-O-glucosyltransferase
RoC4H	Raspberry	FJ554629.1	Cinnamate-4-hydroxylase
RibHLH	Apple	ABB84474.1	b-Helix loop helix transcription factor
RiMYB_209	Raspberry	FJ527833	Myb-related transcription factor
RiMYB10	Raspberry	Unpublished	MYB10 transcription factor
RubFAC-003_MYB1	Strawberry	AF401220	MYB1 transcription factor
RubFruitE4 ^b	Raspberry	Unpublished	bZIP transcription factor

^a Sargent et al. (2007)

^b Graham et al. (2004)

Methods

Plant material

The diploid purple raspberry mapping population consists of single, non-replicated population of 500 full-sib F₁ individuals derived from an inter-specific cross between parents of contrasting fruit color, *Rubus occidentalis* selection 96395S1 and *R. idaeus* ‘Latham’ (Bushakra et al. 2012). The pedigree of each parent is described by Bushakra et al. (2012). The population is located at the Plant & Food Research (PFR) site near Motueka (41°058'S, 172°584'E). A subset of 155 individuals (Bushakra et al. 2012) was selected for ACY extraction from fruit and QTL analysis.

Marker nomenclature and linkage groups

Markers were designed previously, and Tables 1 and 2 provide information on the sources of the orthologous sequence used for primer design (Bushakra et al. 2012). *Rubus* linkage groups (RLG) are named to correspond with the diploid *Fragaria vesca* LG nomenclature (Bushakra et al. 2012).

Sample collection, extraction, preparation, and analysis

In each of the 2009, 2010 and 2011 harvest seasons, field-grown fruit was collected three times, once per week over

Table 2 Derivation of markers mapped on *Rubus occidentalis* 96395S1 × *R. idaeus* ‘Latham’ that were designed from genes from pathways other than the polyphenolic pathway and other sequences (Bushakra et al. 2012)

Marker name	Sequence source	Accession	Putative gene function
EMFxaACO1B ^a	Strawberry	AJ851828	1-Aminocyclopropane-1-carboxylate oxidase
RiABP	Strawberry	X91839	Auxin-binding protein
RiAq	Raspberry	GT119101	Aquaporin
RiCAH	Raspberry	GT119097	Cytoplasmic aconitate hydratase (Aco2)
RiF2KP	<i>A. thaliana</i>	AT1G07110	Fructose-6-phosphate 2-kinase
RiFRK	<i>A. thaliana</i>	AT1G06020	Fructokinase-3
RiFRUCT4	<i>A. thaliana</i>	NM_101096.2	b-Fructofuranosidase
RiG6PD5	<i>A. thaliana</i>	NM_121314	Glucose-6-phosphate dehydrogenase 2
RiHKL	<i>A. thaliana</i>	AT1G50460	Hexokinase
RiPPC1	tomato	AJ243416.1	Phosphoenolpyruvate carboxylase 1
RiSAMS	raspberry	GT119107	s-Adenosylmethionine synthetase
RiSDH2	<i>A. thaliana</i>	AY133848.1	Sorbitol dehydrogenase
RiSNF4	raspberry	GU056038	Protein kinase
RiSUC4	<i>A. thaliana</i>	AT1G09960	Sucrose transporter 4
RubARSFL_134_APX	strawberry	AF158652	Cytosolic ascorbate peroxidase (ApxSC)
RubARSFL_35_Cel2	strawberry	AF054615	Cellulase
SNF2	peach	BU044922	Protein kinase
RiRosCOS1549	Rosaceae	ROSC_FMLY_CSA1_1549 ^b	Similar to Tsi1-interacting protein
RiRosCOS2343	Rosaceae	ROSC_FMLY_CSA1_2343 ^b	Uncharacterized protein
RiRosCOS2381	Rosaceae	ROSC_FMLY_CSA1_2381 ^b	Uncharacterized protein
RiRosCOS2409	Rosaceae	ROSC_FMLY_CSA1_2409 ^b	Putative 1-hydroxy-2-methyl-2-(E)-butenyl 4- diphosphate synthase, GcpE family protein
RosCOS1378 ^c	Rosaceae	ROSC_FMLY_CSA1_1378	Putative UDP-glucose:glycoprotein glucosyltransferase
PBCESSRFxa12 ^d	Strawberry	DQ180317	SSR
Ri_GDSNP00321	Apple	GD_SNP00321	SNP
Ri_GDSNP02690	Apple	GD_SNP02690	SNP

^a Sargent et al. (2007)

^b http://cgpdb.ucdavis.edu/rosaceae_assembly/rosaceae_sequences_412832_Dec_2007.Clean.COS.CDS.assembly

^c Cabrera et al. (2009)

^d Keniry et al. (2006)

3 weeks to represent the typical harvest season for raspberry in New Zealand (mid-December through early-January). Individuals were arranged in rows and after planting were assigned a unique identification number (EA501–EA701). Plants were grown following standard New Zealand commercial raspberry practices.

Fruit samples were collected on ice and then stored at -20°C until processing. The three samples collected per individual were pooled for processing, either by whole fruit solvent extraction (2009), or by juice extraction (2010, 2011) methods. In 2009, the first 200 F_1 individuals that were producing fruit were selected for analysis. Of those 200 plants, 155 had sufficient fruit production accumulated over the three harvests for analysis. In 2010 and 2011, the fruit from all 500 F_1 progeny were netted to protect the fruit from bird predation, and samples were collected from

all individuals, with the subset of 155 F_1 progeny being used for the QTL analysis.

Sample extraction, 2009: solvent extraction method

Frozen whole fruit samples were shipped overnight on dry ice from PFR, Motueka, New Zealand to PFR, Palmerston North for processing. The three replicates of fruit from each individual were pooled. Twenty grams, including seeds, of the pooled fruit samples, was homogenized with 100 mL ethanol/water/formic acid (80:20:1, v/v/v) using an Omni GLH homogenizer (Omni International, Marietta, GA) and stored at 1°C for 48 h before centrifugation at 1000 g for 10 min with a Jouan[®] Hema-C centrifuge (Jouan, Saint Herblain, France). An aliquot (180 μL) of the supernatant was transferred to a 96-well plate for

analysis by reversed-phase ultra high-performance liquid chromatography (UHPLC).

UHPLC analysis of fruit anthocyanins, 2009

Analyses of ACY composition and concentration in individuals in the subset of 155 F₁ progeny of 96395S1 × ‘Latham’, were performed on a Dionex Ultimate[®] 3000 Rapid Separation LC system equipped with a SRD-3400 solvent rack with four degasser channels, HPR-3400RS binary pump, WPS-3000RS thermostated autosampler, TCC-3000RS thermostated column compartment, and DAD-3000RS Diode Array Detector (Sunnyvale, CA, USA). A volume of 1.0 µL supernatant was injected into a Zorbax[™] SB-C18 2.1 × 150 mm, 1.8 µm analytical column (Agilent Technologies, Santa Clara, CA, USA) and analyzed at a mobile flow rate of 0.5 mL/min with column oven temperature maintained at 45 °C. Solvents were (A) 0.5 % phosphoric acid in water, and (B) acetonitrile/water (80:20 v/v); a binary gradient elution was performed as follows: isocratic at 90 % A:10 % B (0–0.5 min); linear gradient to 85 % A:15 % B (0.5–3.5 min); linear gradient to 50 % A:50 % B (3.5–9.0 min); linear gradient to 100 % B (9.0–9.5 min), isocratic at 100 % B (9.5–10.5 min); linear gradient to 90 % A:10 % B (10.5–11.0 min). Total sample run time was 13.5 min and spectral data were collected for the entire run.

Compound quantification, 2009

Quantification of ACY extracted from the subset of 155 F₁ progeny was carried out at 520 nm as cyanidin 3-*O*-glucoside equivalents (Mullen et al. 2003). Chromatographic data were collected and manipulated using the Chromeleon[®] Chromatography Management System version 6.8 (Dionex Corporation, Sunnyvale, CA, USA). External calibration curves were constructed for cyanidin 3-*O*-glucoside using standards from Extrasynthese (Genay, France). Components for which standards were not available were quantified using the standard curve of a related compound. The results are presented as µg/mL.

Sample extraction 2010 and 2011: juice extraction method

The three replicates of fruit from each individual from the F₁ progeny and the parents (fruit from both parents collected in 2010, fruit from only ‘Latham’ collected in 2011) were pooled to give an average fruit weight per individual of 348 g. Juice was extracted using a manual juicer with fine-screen mesh to exclude pulp and seeds. The juicer was washed well and dried between samples. For each extraction, a 1.0-mL sample was collected from the total juice, transferred to a 1.5-mL Eppendorf[®] tube, and stored at –80°C

until analysis. Juice samples were shipped overnight on dry ice from PFR, Motueka to PFR, Ruakura for analysis.

HPLC analysis of fruit anthocyanins: 2010 and 2011

Analyses of ACY composition and concentration for fruit from the subset of 155 F₁ progeny of 96395S1 × ‘Latham’ were performed on a Shimadzu 20-series analytical HPLC with a column oven, auto-sampler, 4-channel vacuum solvent degas module and diode-array detector (Portland, OR, USA). A volume of 10 µL supernatant was injected into a 150 × 2 mm, Synergi Polar-RP, 4 µ particle size, 80 Å pore size analytical column, fitted with a Security-Guard 3 × 2 mm Polar-RP guard cartridge (Phenomenex, Auckland, New Zealand). Flow rate was 0.6 mL/min and column oven temperature 50 °C. Solvents were (A) methanol (neat) and (B) 2 % aqueous formic acid, and the initial mobile phase was 5 % (A):95 % (B). The time program for pump B concentration was set up as 92 % at 2.5 min, 82 % at 5 min, 72 % at 7 min, 58 % at 9.5 min, 45 % at 11 min, 30 % at 12 min, 20 % from 12.5–13.3 min, returning to 95 % at 13.8 min and holding at that concentration until the end of the run at 15 min.

Quantification of ACY extracted from the 500 F₁ progeny was carried out at 520 nm, in comparison with standard solutions of cyanidin 3-*O*-glucoside (Mullen et al. 2003). UV/visible spectra were recorded from 245–600 nm in 1.2 nm steps. Results were expressed as mg cyanidin 3-*O*-glucoside equivalent/100 mL of juice.

Confirmation of compound identity, 2010 and 2011

To assist identification of compounds observed during HPLC runs, some samples were rerun on the same HPLC instrument, with the addition of a Shimadzu LCMS 2020, single quadrupole mass spectrometer, fitted with an electrospray ionization (ESI) source [Shimadzu Scientific Instruments (Oceania) Pty Limited, Auckland, NZ], following manufacturer’s default parameter settings. MS scans were carried out in both positive and negative modes in the same run, using a mass range of 140–2,000 Da. Confirmation of identity was achieved both through comparison of expected and observed molecular ion masses and “neutral losses” of sugars during ionization to leave aglycones or partially glycosylated ions. For example, cyanidin 3-*O*-glucoside exhibited a positive molecular ion at 449 Da and an ion at 287 Da, corresponding to loss of the glucose residue (162 Da), to leave the cyanidin ion.

Statistical analyses

The GenStat 14th edition software package (VSN International 2010) was used to analyze the phenotypic data.

The mean concentration values for the 155 F₁ progeny for all 3 years were calculated for each compound and within-year, pair-wise compound correlations were conducted using the data. A probability distribution plot was applied to verify the normality of the trait distributions, and the means, ranges, standard deviations (SD), skewness and kurtosis were calculated for these distributions. Transgressive segregants (those progeny whose values exceed or are less than the parental values by at least two SD) were identified for the compounds analyzed in 2010, as the parental data were available only for that year.

As the methods of analysis resulted in the presentation of data with different units for 2009 and 2010/2011, each compound was evaluated as a percentage of the total ACY for each year, to assess the proportion of each compound per year. Restricted maximum likelihood (REML) analysis was conducted on the mean percent data to assess the trait variance contributed by year-to-year variation and by differences in genotype. The response variate was the trait's percent of total ACY, the fixed model was constant, the random model was Year + Genotype + Year.Genotype, and there were 462 U (3 years × 155 individuals – 3 individuals with no data for 2011), with 3 U excluded for missing data.

QTL map construction

Markers for QTL mapping were selected from the parental linkage maps (Bushakra et al. 2012) with the goal of developing linkage groups with one marker every 2–5 centimorgans (cM) and including as many gene-based markers as possible (Tables 1, 2). Marker selection was based on map location, least missing data, best segregation ratio, and when possible, markers in common with other *Rubus* linkage maps. The linkage maps were generated using MapChart[®] 2.1 (Voorrips 2002).

QTL mapping

QTL analyses were conducted for the concentration of each ACY compound and for total ACY for each year using MapQTL[®] 5 (van Ooijen 2004). To limit the genotypic classes to two, QTL detection was performed on the parental maps (Bushakra et al. 2012). Logarithm of odds (LOD) thresholds for each map was estimated with a 1,000-permutation test for each trait and year. An interval mapping (IM) analysis was then performed to locate preliminary QTL. QTL with LOD scores greater than the genome-wide threshold of 95 % were considered significant. Markers with significant association with a trait were selected for use as cofactors in the Multiple QTL Method (MQM) to uncover the presence of minor QTL that might be hidden by large-effect QTL. A step size of 1 cM was

used for both IM and MQM analyses. The square of the partial correlation coefficient (R^2) was used to calculate the phenotypic variance explained by a single QTL. The 95 and 99 % confidence intervals for QTL positions were obtained by taking the two positions left and right of the point estimate of the QTL for the LOD ± 1 interval, and the next two positions out, left and right, for the LOD ± 2 interval (van Ooijen 1992), respectively, for QTL detected with IM and with MQM.

Epistasis calculations

An evaluation of potential epistasis was conducted on the traits C3S, C3XR C3XR_C3R, and C3GR, as these traits had at least two QTL that were stable over 2009 and 2010. Simple fixed-effect models, whereby the particular ACY was modeled by the putative QTL (as fixed effects) and all possible interactions were fitted. Residual plots were checked for the presence of outliers and to ascertain the validity of the assumption of normality (Supplemental Fig. S1). All analyses and figures were undertaken in R 2.13.0 with the models using the asreml 3.0.1 (P. Alspach, PFR, pers. comm.).

Comparative gene analysis

The annotated genes from the syntenic region of *Fragaria vesca* LGII (1:4 Mb) were mined to identify candidates potentially involved in the expression of the large-effect QTL located at the top of 'Latham' RLG2. Gene models identified in *Fragaria* were compared with the *Malus × domestica* 'Golden Delicious' (Velasco et al. 2010) and *Fragaria vesca* 'Hawaii 4' (Shulaev et al. 2011) draft genome sequences. Over 400 gene models were represented. The best basic local alignment search tool (BLAST) matches for the gene models were then compared with The *Arabidopsis* Information Research (TAIR) database (<http://www.arabidopsis.org>) to validate the predicted gene function.

Results

Description of traits

Seven ACY compounds were detected in the F₁ progeny in all 3 years and in both parents in 2010. Five of the compounds, co-eluting compounds cyanidin 3-*O*-glucoside and cyanidin 3-*O*-sambubioside (C3G_C3Sb), cyanidin 3-*O*-sophoroside (C3S), co-eluting compounds cyanidin 3-*O*-2^G-xylosylrutinoside and cyanidin 3-*O*-rutinoside (C3XR_C3R), cyanidin 3-*O*-2^G-glucosylrutinoside (C3GR), and pelargonidin 3-*O*-rutinoside (P3R) were detected in all 3 years, identified based on UHPLC or HPLC retention times confirmed with

Table 3 Anthocyanin compounds extracted from purple raspberry fruit collected from 155 F₁ progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* ‘Latham’, analyzed with ultra high-

performance liquid chromatography (UHPLC) in 2009 and HPLC in 2010 and 2011, and used for quantitative trait locus (QTL) analysis

Chemical name	Common name	Abbreviation
Cyanidin 3- <i>O</i> -glucoside	Cyanidin 3- <i>O</i> -glucoside	C3G
Cyanidin 3- <i>O</i> -[2-(xylosyl)glucoside]	Cyanidin 3- <i>O</i> -sambubioside	C3Sb ^a
Cyanidin 3- <i>O</i> -[2-(glucosyl)glucoside]	Cyanidin 3- <i>O</i> -sophoroside	C3S
Cyanidin 3- <i>O</i> -[6-(rhamnosyl)glucoside]	Cyanidin 3- <i>O</i> -rutinoside	C3R ^b
Cyanidin 3- <i>O</i> -[2-(xylosyl)-6-(rhamnosyl)glucoside]	Cyanidin 3- <i>O</i> -2 ^G -xylosylrutinoside	C3XR ^b
Cyanidin 3- <i>O</i> -[2-(glucosyl)-6-(rhamnosyl)glucoside]	Cyanidin 3- <i>O</i> -2 ^G -glucosylrutinoside	C3GR
Pelargonidin 3- <i>O</i> -[6-(rhamnosyl)glucoside]	Pelargonidin 3- <i>O</i> -rutinoside	P3R

^a Cyanidin 3-*O*-sambubioside co-eluted with C3G and was detected in 2009 and 2011 only^b Cyanidin 3-*O*-rutinoside and cyanidin 3-*O*-2^G-xylosylrutinoside co-eluted as a single peak in 2010 and 2011

mass spectrometry data, and therefore were selected for QTL analysis (Table 3). In the samples from 2009 and 2011, compounds C3G and C3Sb co-eluted as a single peak (C3G_C3Sb), while in the samples from 2010, C3G only was detected. In the samples from 2009, compounds C3XR and C3R were detected as separate peaks, while in the samples from 2010 and 2011, the two compounds co-eluted as a single peak (C3XR_C3R). Therefore, the concentrations of C3G_C3Sb 2009 and 2011 and C3G 2010, as well as the sum of C3XR 2009 and C3R 2009 (C3XR + C3R) and co-eluting compounds C3XR_C3R 2010 and 2011 were used for within-year comparisons and the proportions of total ACY for each compound were used for between-year comparisons. As the co-eluting compounds are each considered as a single trait, five traits in total were considered for analysis. The two unknown ACY, present in trace amounts, were excluded from further analysis. Total ACY was calculated as the sum of all analyzed compounds for each year.

Relative proportions of total ACY content of each analyzed compound, calculated from the mean compound values of the subset of 155 F₁ progeny, were similar for all 3 years (Table 4). C3GR had the highest overall proportion of total analyzed ACY in the progeny (33 %, 2009; 39 %, 2010; 37 %, 2011), followed by co-eluting compounds C3XR_C3R (sum of compounds 39 % 2009; 33 % 2010; 34 % 2011). The mean compound concentrations in 2010 and 2011 were not significantly different with the exception of C3GR, which was higher in 2010 (2010 81.7 mg/100 mL; 2011 71.1 mg/100 mL; $p < 0.05$). As the mean differences between 2010 and 2011 were largely not significant, all between-year comparisons are performed using data from 2009 and 2010 only.

Weather data, compiled by MetWatch Online (<http://www.metwatch.co.nz>), for the typical New Zealand raspberry growing season (1 August through 8 January) for 2009, 2010 and 2011, are shown in Supplemental Table S1. Total rainfall (millimeters, mm), mean daily temperature (degrees Celsius, °C), and growing degree days (GDD,

number of hours with day-time temperatures above the base temperature, in this case the base temperature is 7 °C) were all higher in 2010 than in 2009 and 2011.

Descriptive statistics

The probability distribution plot p -values performed on the concentration data indicate that the phenotypic distributions of co-eluting compounds C3G_C3Sb (2009) (Supplemental Fig. S2a), C3S (all 3 years) (Supplemental Fig. S2b), C3R (2009), C3XR (2009), co-eluting compounds C3XR_C3R (2010 and 2011) (Supplemental Fig. S2c), and P3R (all 3 years) (Supplemental Fig. S2e) deviated

Table 4 Mean proportions of anthocyanins detected in the F₁ progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* ‘Latham’ in 2009, 2010, and 2011

	Mean (%)	SD (%)	SEM (%)
C3G_C3Sb 2009	10	3	0
C3G 2010	7	2	0
C3G_C3Sb 2011	8	2	0
C3S 2009	17	11	1
C3S 2010	19	10	1
C3S 2011	21	10	1
C3XR + C3R 2009	39	18	1
C3XR_C3R 2010	33	14	1
C3XR_C3R 2011	34	13	1
C3GR 2009	33	11	1
C3GR 2010	39	8	1
C3GR 2011	37	2	0
P3R 2009	1	2	0
P3R 2010	2	2	0
P3R 2011	0	1	0

SD standard deviation, SEM standard error of the mean, C3S cyanidin 3-*O*-sophoroside, C3XR cyanidin 3-*O*-2^G-xylosylrutinoside, C3R cyanidin 3-*O*-rutinoside, C3G cyanidin 3-*O*-glucoside, C3Sb cyanidin 3-*O*-sambubioside, C3GR cyanidin 3-*O*-2^G-glucosylrutinoside, P3R pelargonidin 3-*O*-rutinoside

significantly ($p \leq 0.01$) from normal (Table 5, Supplemental Fig. S2). The phenotypic distributions of ACY compound C3G (2010), co-eluting compounds C3G_C3Sb (2011) (Supplemental Fig. S2a), and C3GR (all three years) (Supplemental Fig. S2d) were approximately normal ($p \geq 0.05$; Table 5).

In 2010, fruit collected from the black raspberry female parent, 96395S1, had higher concentrations of compounds P3R, C3G, and C3XR_C3R than the red raspberry male parent, 'Latham', which had higher concentrations of compounds C3S and C3GR (Table 5, columns 96395S1 and 'Latham'). In 2010 and 2011, 'Latham' contained moderate amounts of C3S and C3GR, both red raspberry-specific compounds, as well as relatively low quantities of black raspberry-specific compounds C3XR_C3R and C3G, while in 2010, 96395S1 had trace amounts of red raspberry-specific C3S and C3GR along with concentrations of C3XR_C3R and C3G typical for black raspberry. Concentrations of C3GR were greatest of all individual compounds in all 3 years. In 2010 both parents had concentrations of C3GR less than all of the progeny, 68 of which had concentrations greater than two SD of 'Latham', which had the higher concentration of the two parents (Table 5). Transgressive segregants (TS) were identified for traits P3R 2010, with two TS (96395S1 9.5 mg/100 mL, 'Latham' 0.8 mg/100 mL; progeny range 0.3–18.3 mg/

100 mL); C3S 2010, with 10 TS (96395S1 1.3 mg/100 mL, 'Latham' 34.6 mg/100 mL; progeny range 4.1–99.7 mg/100 mL); and C3GR 2010 with 68 TS (96395S1 3.0 mg/100 mL, 'Latham' 28.8 mg/100 mL; progeny range 32.1–178.8 mg/100 mL).

REML variance component analysis conducted on the mean percentage of each compound for each year suggests that the amount of variation contributed by differences between years has much less influence on compound proportion of total ACY accumulation than does genotype (Table 6). This observation is supported by the standard error (s.e.) of each of the components. The s.e. for component "year" is nearly the same as the variance of the component itself, whereas the s.e. for component "genotype" is 6–7 times less than the variation contributed by genotype.

Pearson's correlation coefficients calculated for each pair of compounds are shown in Table 7. A number of compounds correlated significantly with each other, with the strongest positive correlation ($r > 0.70$) between C3S and C3GR for all 3 years; the strongest negative correlation ($r = -0.55$) was between C3XR_C3R and C3S in 2010. In general, the correlation coefficients agreed between years. The results of the correlation analysis suggest that C3G (and C3G_C3Sb) is positively correlated with all other compounds, except a slightly negative

Table 5 Mean phenotypic values and basic statistical parameters for the five traits analyzed in 155 F1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham' in 2009 and 2011, and the progeny and parents in 2010

Trait	F ₁ mean	F ₁ maximum	F ₁ minimum	Standard deviation	Skewness	Kurtosis	Probability distribution	96395S1 ^a	'Latham' ^a	Transgressive segregants ^b
C3G_C3Sb_09	7.3	21.8	2.6	3.4	1.8	4.6	$p \leq 0.01$	n.m.	n.m.	
C3S_09	12.6	40.4	0.7	9.1	0.7	-0.3	$p \geq 0.05$	n.m.	n.m.	
C3XR_09	10.0	27.0	2.0	5.4	0.9	0.3	$p \leq 0.01$	n.m.	n.m.	
C3R_09	19.7	61.5	4.6	11.7	1.2	1.2	$p \leq 0.01$	n.m.	n.m.	
C3GR_09	24.8	54.7	6.6	11.0	0.5	-0.5	$p \geq 0.05$	n.m.	n.m.	
P3R_09	0.8	10.2	0.1	1.1	5.4	40.9	$p \leq 0.01$	n.m.	n.m.	
C3G_10	15.1	27.0	4.6	4.3	0.3	-0.3	$p \geq 0.05$	46.0	8.4	0
C3S_10	40.9	99.7	4.1	25.7	0.4	-0.9	$p \leq 0.01$	1.3	34.6	10
C3XR_C3R_10 ^c	69.5	160.3	19.0	28.0	0.8	0.5	$p \leq 0.01$	239.6	11.5	0
C3GR_10	81.8	178.8	32.1	29.6	0.4	-0.3	$p \geq 0.05$	3.0	28.8	68
P3R_10	3.6	18.3	0.3	2.7	2.5	10.5	$p \leq 0.01$	9.5	0.8	2
C3G_C3Sb_11	14.9	32.0	1.3	4.5	0.5	1.3	$p \geq 0.05$	n.m.	5.6	
C3S_11	41.4	107.3	0.6	24.7	0.4	-0.8	$p \leq 0.01$	n.m.	21.8	
C3XR_C3R_11 ^c	65.2	139.2	9.5	29.7	0.5	-0.4	$p \leq 0.01$	n.m.	6.3	
C3GR_11	71.3	163.4	19.8	26.2	0.6	0.5	$p \geq 0.05$	n.m.	18.7	
P3R_11	0.8	8.1	0.0	1.4	2.7	8.1	$p \leq 0.01$	n.m.	1.1	

Progeny size = 155

n.m. Not measured; C3S cyanidin 3-*O*-sophoroside, C3XR cyanidin 3-*O*-2^G-xylosylrutinoside, C3R cyanidin 3-*O*-rutinoside, C3G cyanidin 3-*O*-glucoside, C3Sb cyanidin 3-*O*-sambubioside, C3GR cyanidin 3-*O*-2^G-glucosylrutinoside, P3R pelargonidin 3-*O*-rutinoside

^a Parents not analyzed in 2009. 'Latham' only analyzed in 2011

^b TS: Number of individuals that exceed the parent by at least two SD

^c Compounds C3XR_C3R co-eluted as a single peak in 2010 and 2011

Table 6 Variance components for five anthocyanin compounds analyzed for 3 years (2009, 2010, 2011) in 155 F₁ progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* ‘Latham’

	Component	s.e.
% C3GR All		
Year	7.74	7.97
Genotype	42.01	6.23
Residual	35.63	2.89
Year	1.39	1.41
% C3G_C3Sb All		
Genotype	3.28	0.49
Residual	2.82	0.23
Year	4.05	4.26
% C3S All		
Genotype	71.84	9.45
Residual	31.8	2.57
Year	9.78	10.31
% C3XR_C3R All		
Genotype	150.59	20.46
Residual	82.49	6.68
Year	0.48	0.48
% P3R All		
Genotype	0.81	0.14
Residual	1.22	0.10

Values are based on the mean percentage of the total anthocyanin accumulation for each year

s.e. standard error, C3S cyanidin 3-*O*-sophoroside, C3XR cyanidin 3-*O*-2^α-xylosylrutinoside, C3R cyanidin 3-*O*-rutinoside, C3G cyanidin 3-*O*-glucoside, C3Sb cyanidin 3-*O*-sambubioside, C3GR cyanidin 3-*O*-2^α-glucosylrutinoside, P3R pelargonidin 3-*O*-rutinoside

correlation with C3GR 2009. Compound C3S has a strong positive correlation with C3GR and is negatively correlated with all other compounds, except a slightly positive correlation with P3R 2011. The correlations of compound C3GR generally agree with those of C3S with the exception of a positive correlation with both C3G_C3Sb and C3XR_C3R in 2011. All remaining compounds show a mix of positive and negative correlations relative to the other compounds. As some of the data distributions were non-normal, a Spearman’s ranking analysis was conducted and showed similar results as the Pearson’s correlations (Supplemental Table S2).

Scatter plots for compound associations between C3S 2009 and C3G_C3Sb 2009, C3S 2009 and C3XR + C3R 2009, C3GR 2009 and C3XR + C3R 2009, C3S 2010 and C3G 2010, as well as C3S 2011 and C3G_C3Sb 2011 show two distinct groupings. The occurrence of subgroups was investigated and an association was detected between the genotype of the progeny at locus RubFruitE4 on ‘Latham’ RLG2, and the compound associations that segregate into two subgroups (Fig. 1; Supplemental Fig. S3). The two

subgroups consist of 61 heterozygous (“Im”; Set A) and 94 homozygous (“II”; Set B) individuals, where “Im” and “II” represent allele combinations of marker RubFruitE4. Individual EA509 (“II”) consistently groups with the phenotypes of Set A, while individuals EA611 and EA697 (both “Im”) consistently group with the phenotypes of Set B.

When the Pearson’s correlation coefficient analysis was conducted on data Set A (Table 8) and Set B (Table 9), most of the correlations became positive, some strongly so, with *r* values as high as 0.85 for the correlation between C3G_C3Sb and C3S for 2009, in contrast to the finding of no significant correlation when the data were considered as a whole. The negative correlations between C3S 2009 and C3R 2009, and between C3S 2009 and C3XR 2009, when the data were considered as a whole, changed to non-significant correlations when the data were considered as separate sets.

A comparison of average, minimum, and maximum values of the four compounds that formed two distinct groups based on genotype at RubFruitE4, showed that though the compound concentrations overlapped, the mean concentration of Set B (“II”) was higher for red raspberry compounds C3GR and C3S, and the mean concentration of Set A (“Im”) was higher for black raspberry compounds C3XR_C3R and C3G_C3Sb. A Student *t* test performed on the differences between the means of the two sets indicates that the differences are statistically significant (*p* < 0.001) for all compounds except C3S 2010 (Table 10).

Linkage maps used for QTL analysis

The ‘Latham’ linkage map used for QTL analysis comprised nine LG with 69 markers spanning 533 cM. Markers per cM ranged from one marker every 2.4 cM (RLG3b), to one marker every 14.8 cM (RLG4), with an average of one marker every 7.8 cM. The largest gap was on RLG4 spanning 42.4 cM, followed by a gap of 31.4 cM on RLG6. RLG3 and RLG7 were both in two parts (designated a and b).

The 96395S1 linkage map used for QTL analysis comprised four LG with 20 markers spanning 161.5 cM. Markers per cM ranged from one marker every 5.2 cM (RLG5), to one marker every 11.1 cM (RLG3), with an average of one marker every 8.1 cM. The largest gap was 24.0 cM found on RLG6, followed by a gap of 19.7 cM on RLG3.

Identification of QTL

Individual compounds

QTL were identified using the mean trait concentration values for each individual progeny from each year separately on the parental genetic maps, the genotype data set

Table 7 Pearson's correlation coefficient calculations and statistical significance (*p* value) for each pair of the six anthocyanin compounds analyzed for 2009, 2010 and 2011 in 155 F₁ progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham'

Trait	<i>C3GR</i> <i>09</i>	<i>C3G</i> <i>C3Sb 09</i>	<i>C3S</i> <i>09</i>	<i>C3XR+</i> <i>C3R 09</i>	<i>C3GR</i> <i>10</i>	<i>C3G 10</i>	<i>C3S 10</i>	<i>C3XR</i> <i>C3R 10</i>	<i>C3GR</i> <i>11</i>	<i>C3G</i> <i>C3Sb 11</i>	<i>C3S 11</i>	<i>C3XR</i> <i>C3R 11</i>
C3G_C3Sb 09	-0.05											
C3S 09	0.78	0.08 ^b										
C3XR + C3R 09	-0.30^b	0.63	-0.50^b									
P3R 09 ^a	-0.35	0.16	-0.36	0.37								
C3G 10					0.17							
C3S 10					0.77	0.24 ^b						
C3XR_C3R 10					-0.19	0.36	-0.55					
P3R 10					-0.18	0.08	-0.25	0.22				
C3G_C3Sb 11									0.57			
C3S 11									0.72	0.43^b		
C3XR_C3R 11									0.32	0.62	-0.14	
P3R 11									0.16	0.05	0.05	-0.10

N = 155

Bold font indicates significance at *p* < 0.001. Italic font indicates significance at *p* < 0.05

C3S cyanidin 3-*O*-sophoroside, *C3XR* cyanidin 3-*O*-2^G-xylosylrutinoside, *C3R* cyanidin 3-*O*-rutinoside, *C3G* cyanidin 3-*O*-glucoside, *C3Sb* cyanidin 3-*O*-sambubioside, *C3GR* cyanidin 3-*O*-2^G-glucosylrutinoside, *P3R* pelargonidin 3-*O*-rutinoside, *RLG* *Rubus* linkage group

^a Missing data replaced by mean value (0.849)

^b These correlations changed when the data were considered as Sets A and B relative to marker RubFruitE4 on RLG2

for each individual compound, and IM followed by MQM. Markers RubFruitE4 and RiF2KP were selected for use as co-factors for MQM analysis conducted on the 'Latham' parental map; marker Ri60S was selected for use as the co-factor used for MQM analysis conducted on the 96395S1 parental map. Co-location of traits with markers was detected, at LOD scores greater than the genome-wide 95 % threshold, on all *Rubus* Linkage Group (RLG) except RLG4 (Supplemental Fig. S4). The 2009 phenotype data taken alone showed that all five traits had significant LOD scores associated with markers on the 'Latham' genetic

Table 8 Pearson's correlation coefficients calculated for each pair of compounds that segregated into two distinct groups in the F₁ progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham'

Trait	C3S 2009	C3R 2009	C3XR 2009	C3S 2010
C3R 2009	0.025			
C3XR 2009	-0.073			
C3GR 2009		0.11	0.44	
C3G 2010				0.81

N = 155

The 61 individuals comprising Set A are heterozygous ("Im") at locus RubFruitE4 on 'Latham' *Rubus* linkage group (RLG) 2

Bold font indicates significance at *p* < 0.001

C3S cyanidin 3-*O*-sophoroside, *C3XR* cyanidin 3-*O*-2^G-xylosylrutinoside, *C3R* cyanidin 3-*O*-rutinoside, *C3G* cyanidin 3-*O*-glucoside, *C3GR* cyanidin 3-*O*-2^G-glucosylrutinoside

map, while four traits (C3S, C3XR, C3G_C3Sb, and C3GR) had significant LOD scores associated with markers on the 96395S1 genetic map. Similarly, the 2010 and 2011 phenotype data, when each was taken alone, showed that all five traits had significant LOD scores associated with markers on the 'Latham' genetic map. On the 96395S1 map, traits C3S, C3G, C3GR, and P3R for 2010 had significant LOD scores associated with markers, while in 2011 traits C3XR_C3R, C3G_C3Sb, and C3GR had significant LOD scores associated with markers.

In total, 49 QTL were detected (Table 11; Supplemental Fig. S4), using IM followed by MQM, on all 'Latham' linkage groups except RLG4, and on 96395S1 RLG 3, 5, and 6. The 49 QTL can be summarized as 28 significant associations (21 on the 'Latham' map, and seven on the 96395S1 map), as QTL that mapped to the same interval, or with overlapping confidence intervals, in all 3 years are considered the same. Of the 21 QTL mapped to 'Latham', four (19 %) were stable over all 3 years, while of the seven QTL mapped to 96395S1, two (29 %) were stable in all 3 years; these six stable QTL are indicated in bold font on Table 11.

QTL were identified on seven of the nine RLG, varying from one each on 'Latham' RLG1, RLG7a and 96395S1 RLG6, to 12 on 'Latham' RLG2 (Supplemental Fig. S4). For the compounds analyzed in 2010, red raspberry-specific compounds C3S and C3GR, 50 % (5 of 10) of the loci displayed gene effects in the same direction as predicted by

the phenotype of the parental lines, while 50 % of the loci showed opposite effects (Table 11). For traits C3XR_C3R and C3G_C3Sb, of which C3XR and C3Sb are black

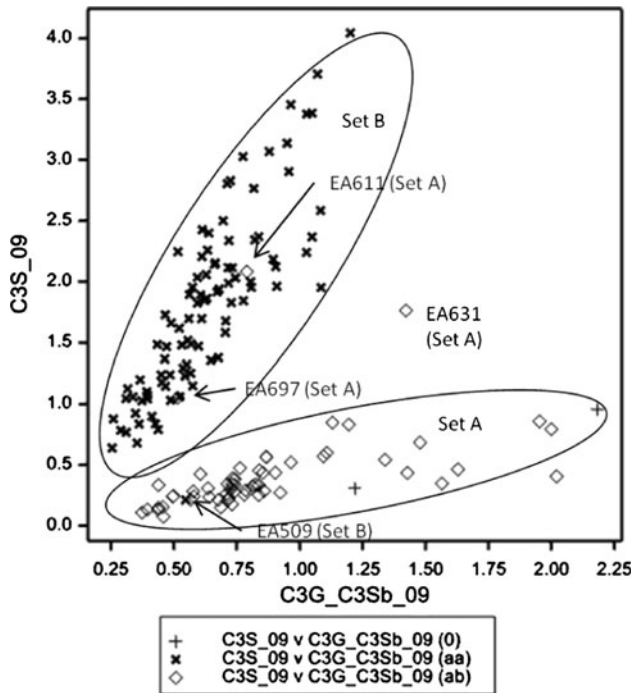


Fig. 1 Scatter plot for compounds cyanidin 3-*O*-sophoroside (C3S) and cyanidin 3-*O*-glucoside and cyanidin 3-*O*-sambubioside (C3G_C3Sb) in 2009. The subset of F₁ progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* ‘Latham’ segregates into two groups according to genotype at locus RubFruitE4 on ‘Latham’ *Rubus* linkage group (RLG) 2. Set A (open diamond symbols) consists of 61 heterozygous (“Im”) individuals; Set B (solid cross symbols) consists of 94 homozygous (“ll”) individuals, where “ll” and “Im” represent allele combinations. The correlation coefficient for the entire data set is $r = 0.08$; the correlation coefficient for Set A is $r = 0.57$, and the correlation coefficient for set B is $r = 0.85$. Individual EA631 is an outlier of Set A. Arrows indicate progeny that group opposite of their genotype

Table 10 Comparison of Set A and Set B for the four compounds that segregated into two distinct groups in the F₁ progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* ‘Latham’ during correlation analysis

Found in	Trait	Set A ave	Min	Max	Set B ave	Min	Max	‘Latham’	96395S1	<i>p</i> value ^a
RR	C3GR 2009	15.3	6.6	28.3	31.0	14.7	54.7			$p < 0.001$
	C3GR 2010	56.7	32.1	119.6	98.0	46.1	178.8	28.8	3.0	$p < 0.001$
BR	C3XR 2009	14.5	5.7	27.0	7.0	2.0	17.4			$p < 0.001$
BR _{RR} &BR ^b	C3XR_C3R 2010	88.0	0.0	160.3	55.0	19.0	130.5	11.5	239.6	$p < 0.001$
RR	C3S 2009	4.1	0.7	17.6	18.2	6.4	40.4			$p < 0.001$
	C3S 2010	14.5	4.1	26.2	58.0	17.7	99.7	34.6	1.3	$p < 0.001$
RR&BR ^b _BR	C3G_C3Sb 2009	8.9	3.7	21.8	6.3	2.6	12.0			$p < 0.001$
	C3G 2010	15.6	6.9	27.0	14.9	4.6	26.4	8.4	46.0	0.386

Ave average, Min minimum, Max maximum, RR red raspberry only compound, BR black raspberry only compound, C3S cyanidin 3-*O*-sophoroside, C3XR cyanidin 3-*O*-2^G-xylosylrutinoside, C3R cyanidin 3-*O*-rutinoside, C3G cyanidin 3-*O*-glucoside, C3Sb cyanidin 3-*O*-sambubioside, C3GR cyanidin 3-*O*-2^G-glucosylrutinoside; P3R pelargonidin 3-*O*-rutinoside

^a Probability associated with paired Student *t* test for the significance of the difference between the mean values for Set A and Set B

^b C3R and C3G are found in both species but in higher concentrations in BR

raspberry-specific compounds, predicted trait direction agreed with the parental phenotypes in 5 of 7 instances (71 %). Predicted trait direction for P3R was opposite of the parental phenotypes in both instances. One to six QTL have been identified for each trait, with the total phenotypic variation (R^2) explained by the QTL for each trait ranging from 0.07 (one QTL for P3R 2010 on 96395S1 RLG5) to 0.88 (four QTL for C3S 2010 on ‘Latham’ RLG 2, 3, 6 and 7). The largest cluster of QTL, representing all five traits over all 3 years, is located on ‘Latham’ RLG2 in association with marker RubFruitE4 and its LOD ± 2 confidence interval spans approximately 32 cM.

While the composition and overall proportions of the compounds were similar among years, the concentrations of the five ACY compounds analyzed varied widely among individuals in the progeny of interspecific cross 96395S1 × ‘Latham’ (Table 5). Table 12 summarizes the

Table 9 Pearson’s correlation coefficients calculated for each pair of compounds that segregated into two distinct groups in the F₁ progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* ‘Latham’

Trait	C3S 2009	C3R 2009	C3XR 2009	C3S 2010
C3R 2009	0.22			
C3XR 2009	0.19			
C3G_C3Sb 2009	0.85			
C3GR 2009		0.64	0.76	
C3G 2010				0.64

N = 155

The 94 individuals comprising Set B are homozygous (“ll”) at locus RubFruitE4 on ‘Latham’ *Rubus* linkage group (RLG) 2

Bold font indicates significance at $p < 0.001$

C3S cyanidin 3-*O*-sophoroside, C3XR cyanidin 3-*O*-2^G-xylosylrutinoside, C3R cyanidin 3-*O*-rutinoside, C3G cyanidin 3-*O*-glucoside, C3Sb cyanidin 3-*O*-sambubioside, C3GR cyanidin 3-*O*-2^G-glucosylrutinoside

Table 11 Mapping and significance information for five anthocyanin compounds mapped in *Rubus occidentalis* 96395S1 and *R. idaeus* 'Latham', % variation explained per locus, and the nearest mapped marker

Trait	Years	Parental map	Original LG ^a	RLG ^b	IM LOD max	cM of IM LOD max	MQM LOD max ^c	MQM LOD max	cM MQM LOD max	GW LOD 95% significance	% Explanation	Effect ^d	Nearest marker
C3S	2009	'Latham'	4	2	25.8	14.1				2.4	53.5	13.6	RubFruitE4 ^{f, g}
C3S	2010	'Latham'	4	2	33.4	15.1				2.5	64.3	41.3	RubFruitE4 ^f
C3S	2011	'Latham'	4	2	8.0	16.1				2.4	22.6	23.1	RubFruitE4 ^f
C3S	2009	'Latham'	2	3	3.8	70.4				2.4	10.6	-5.9	RiF2KP ^g
C3S	2010	'Latham'	2	3	3.3	70.4				2.5	9.4	-15.7	RiF2KP ^g
C3S	2009	'Latham'	3	6			3.5	30.9		2.4	6.5	-3.5	Ri4CoI ^f
C3S	2010	'Latham'	3	6			4.0	16.9		2.5	4.7	-6.5	RiRosCOS2381
C3S	2009	'Latham'	1	7	4.0	0.0				2.4	11.2	5.7	RibHLH ^{f, g}
C3S	2010	'Latham'	1	7	3.3	4.6				2.5	9.2	15.5	RiGT ^g
C3S	2009	96395S1	2	3	4.1	53.3				1.9	13.6	-5.7	Ri60S
C3S	2010	96395S1	2	3	3.2	63.9				1.9	9.5	-13.8	Ri60S
C3S	2011	96395S1			NS					1.9			
C3XR + C3R ^a	2009	'Latham'	4	2	26.5	15.1				2.5	55.3	-24.2	RubFruitE4 ^f
C3XR_C3R	2010	'Latham'	4	2	18.8	13.0				2.3	44.8	-36.8	RubFruitE4 ^f
C3XR_C3R	2011	'Latham'	4	2	5.6	16.1				2.3	17.5	-21.7	RubFruitE4 ^f
C3XR + C3R	2009	'Latham'	2	3			3.8	70.4		2.5	4.9	-3.3	RiF2KP
C3XR + C3R	2009	'Latham'	3	6			4.4	58.1		2.5	6.2	6.5	EMFxaACOIB
C3XR_C3R	2010	'Latham'	3	6			4.7	62.1		2.3	7.4	14.5	EMFxaACOIB
C3XR_C3R	2010	'Latham'	1	7	4.5	3.2				2.3	12.5	-20.0	RiFRK2 ^g
C3XR + C3R	2009	96395S1	3	6	2.0	24.1				1.9	5.8	7.5	Riunknp2 ^g
C3XR_C3R	2010	96395S1			NS					2.1			
C3XR_C3R	2011	96395S1	5	5	2.0	25.8				2.0	5.9	13.6	RiRosCOS536 ^g
C3G_C3Sb	2009	'Latham'	6	1			2.8	0.0		2.2	7.2	-1.7	Rub285a ^g
C3G_C3Sb	2009	'Latham'	4	2	5.4	15.1				2.2	15.4	12.9	RubFruitE4 ^f
C3G_C3Sb	2009	'Latham'	2	3	3.9	70.4				2.2	11.0	68.2	RiF2KP ^g
C3G	2010	'Latham'	2	3	3.2	70.4				2.4	9.0	68.0	RiF2KP ^g
C3G_C3Sb	2011	'Latham'	2	3	2.7	70.4				2.3	7.9	68.1	RiF2KP ^g
C3G_C3Sb	2009	'Latham'	3	6	3.8	10.0				2.2	12.5	7.8	RiRosCOS2381
C3G	2010	'Latham'	3	6	4.9	13.9				2.4	13.5	11.5	RiRosCOS2381 ^g
C3G_C3Sb	2009	'Latham'	1	7			2.5	20.7		2.2	6.1	1.6	ERubLR_SQ07-3_C07 ^g
C3G	2010	'Latham'	1	7			2.4	0.0		2.4	6.8	2.2	RibHLH ^{f, g}
C3G_C3Sb	2009	96395S1	2	3	4.2	66.4				1.9	12.0	-2.3	Rub12a ^g
C3G	2010	96395S1	2	3	3.5	66.4				2.0	10.0	-2.5	Rub12a ^g

Table 11 continued

Trait	Years	Parental map	Original LG ^a	RLG ^b	IM LOD max	cM of IM LOD max	MQM LOD max ^c	cM MQM LOD max	GW LOD 95% significance	% Explanation	Effect ^d	Nearest marker
C3G_C3Sb	2011	96395S1	2	3	2.4	54.3			1.9	8.3	-2.3	Ri60S
C3GR	2009	'Latham'	4	2	21.1	14.1			2.4	46.6	11.7	RubFruitE4^{f, g}
C3GR	2010	'Latham'	4	2	18.1	14.1			2.3	41.6	11.8	RubFruitE4^{f, g}
C3GR	2011	'Latham'	4	2	3.0	13.0			2.3	9.1	10.7	RubFruitE4^f
C3GR	2009	'Latham'	2	3	5.2	69.1			2.4	15.8	66.7	RiF2KP
C3GR	2010	'Latham'	2	3	3.3	70.4			2.3	9.3	68.1	RiF2KP ^g
C3GR	2010	'Latham'	5	5			5.1	49.6	2.3	8.2	14.0	RiF3'H ^g
C3GR	2010	'Latham'	3	6	2.7	75.3			2.3	9.7	16.2	RiG6PD5
C3GR	2011	'Latham'	1	7			2.4	3.7	2.3	6.3	-11.3	RiAcq ^g
C3GR	2009	96395S1	2	3	6.5	50.3			1.9	23.6	-7.9	Ri60S
C3GR	2010	96395S1	2	3	3.8	66.4			1.9	10.8	-19.2	Rub12a^g
C3GR	2011	96395S1	2	3	2.4	47.3			1.9	8.5	-10.7	Ri60S
P3R	2009	'Latham'	4	2	5.7	14.1			1.8	15.6	-0.8	RubFruitE4 ^f
P3R	2009	'Latham'	4	2			35.8	19.1	1.8	68.1	0.4	RiSNF4
P3R	2009	'Latham'	5	5	2.0	46.9			1.8	6.1	-0.5	RiF3'H
P3R	2010	'Latham'	5	5	3.6	53.6			2.3	10.7	-1.7	RiF3'H
P3R	2011	'Latham'			NS				2.3			
P3R	2009	96395S1			NS				1.5			
P3R	2010	96395S1	5	5	2.3	6.2			1.9	6.6	-1.4	RH_MEa0007a_ B01 ^g
P3R	2011	96395S1			NS				1.8			

C3S cyanidin 3-*O*-sophorose, *C3XR* cyanidin 3-*O*-2⁶-xylosylrutinoside, *C3R* cyanidin 3-*O*-rutinoside, *C3G* cyanidin 3-*O*-glucoside, *C3Sb* cyanidin 3-*O*-sambubioside, *C3GR* cyanidin 3-*O*-glucoside, *P3R* pelargonidin 3-*O*-rutinoside, *RLG Rubus* Linkage Group, *IM* interval mapping; *GW* genome-wide, *LOD* logarithm of odds, *NS* no significant LOD; *LG* linkage group, *cM* centimorgan

^a Linkage group (LG) numbering from Graham et al. (2004)

^b *Rubus* linkage group (RLG) numbering from Bushakra et al. (2012)

^c Multiple QTL Model (MQM) co-factors for traits mapped on the 'Latham' parental map were either RubFruitE4 (RLG2) or RiF2KP (RLG3). MQM co-factor for traits mapped on the 96395S1 parental map was Ri60S (RLG3)

^d Estimated additive effects calculated as the μ II - μ Im, where μ II and μ Im are the estimated mean of the distribution of the quantitative trait associated with the homozygous or heterozygous genotypes, respectively

Bold font indicates that the QTL was stable over all 3 years

^e Sum of C3XR and C3R compounds for 2009

^f Marker in common with Kassim et al. (2009)

^g Marker at peak of QTL

Table 12 Proportions of the total anthocyanin for each compound detected in the analyzed set of F₁ progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* ‘Latham’ in 2009 and 2010. Proportions of

anthocyanins from populations of black raspberry cultivars, red raspberry ‘Latham’ × ‘Glen Moy’ and purple raspberry from other studies are provided for comparison

Trait	Compound proportion of total anthocyanins 2009 ^a (%)	Trait	Compound proportion of total anthocyanins 2010 ^{a, b} (%)	‘Black raspberry cultivars’ ^d (%)	‘Latham’ × ‘Glen Moy’ ^e (%)	Purple raspberry populations ^f (%)
C3S	17	C3S	20	na	64	14%
C3XR + C3R ^c	39	C3XR_C3R	33	32–64	1.6	14–33
C3G_C3Sb	10	C3G	7	2–6 (C3Sb)	23 (C3G)	12 (C3G) 5 % (C3Sb)
C3GR	33	C3GR	39	na	23	26
P3R	1	P3R	2	1–5	1.3	na

na Information not available, C3S cyanidin 3-*O*-sophoroside, C3XR cyanidin 3-*O*-2^G-xylosylrutinoside, C3R cyanidin 3-*O*-rutinoside, C3G cyanidin 3-*O*-glucoside, C3Sb cyanidin 3-*O*-sambubioside, C3GR cyanidin 3-*O*-2^G-glucosylrutinoside, P3R pelargonidin 3-*O*-rutinoside

^a Calculated based on mean trait values of individual F₁ progeny (current study)

^b Differences in proportions between 2010 and 2011 were not significant therefore only 2010 data are used for this comparison

^c Sum of means for C3XR and C3R for 2009

^d Dossett et al. 2010

^e Kassim et al. 2009

^f Scalzo et al. 2008

proportions of individual ACY in the subset of 155 analyzed progeny of 96395S1 × ‘Latham’ compared with published results for red, black, and purple raspberry. Cyanidin 3-*O*-rutinoside (C3R) and cyanidin 3-*O*-2^G-xylosylrutinoside (C3XR) were the major ACY compounds identified in the black raspberry cultivars analyzed by others (Dossett et al. 2010; Tulio et al. 2008) comprising 39–90 % of the total ACY. The proportions of individual ACY in the subset of 96395S1 × ‘Latham’ analyzed progeny were more evenly distributed among the five identified compounds when compared with the other black raspberry cultivars (Dossett et al. 2010), but with a greater proportion of the co-eluting compounds cyanidin 3-*O*-glucoside and cyanidin 3-*O*-sambubioside (C3G_C3Sb) and smaller proportions of cyanidin 3-*O*-rutinoside (C3R), and co-eluting compounds cyanidin 3-*O*-2^G-xylosylrutinoside and cyanidin 3-*O*-rutinoside (C3XR_C3R). The subset of 96395S1 × ‘Latham’ analyzed progeny has smaller proportions of cyanidin 3-*O*-sophoroside (C3S) and cyanidin 3-*O*-glucoside (C3G), with greater proportions of cyanidin 3-*O*-rutinoside (C3R) and cyanidin 3-*O*-2^G-glucosylrutinoside (C3GR) when compared with ‘Latham’ × ‘Glen Moy’ (Kassim et al. 2009). The proportion of pelargonidin 3-*O*-rutinoside (P3R) was consistent among studies.

Ranking of individuals

Figure 2 shows the relative accumulation of percent C3S per genotype with individuals ranked from lowest to

highest percent for 2009 (data for other compounds are given in Supplemental Fig. S5). In general, rankings were consistent between 2009 and 2010 (those individuals with low accumulation in 2009 also had low accumulation in 2010). Percent accumulation in 2011 was the most variable and at the extremes appears to be opposite of the percent accumulation in 2009.

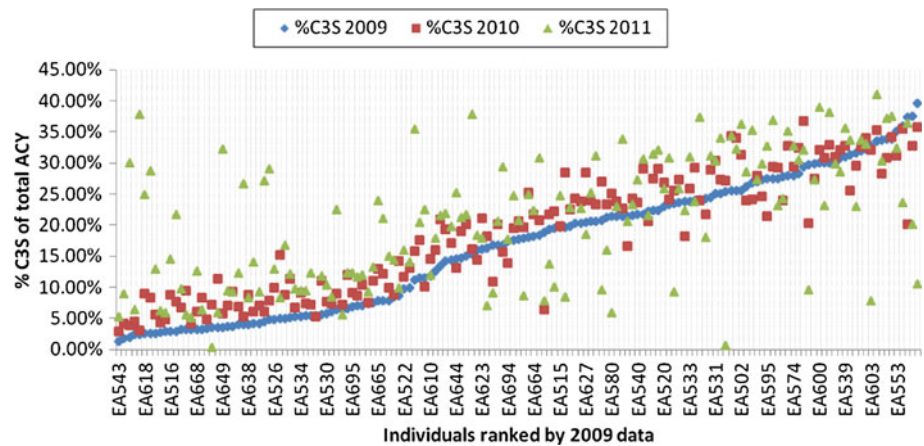
Total anthocyanin

QTL for total ACY were identified using the sum of the mean trait values for each individual progeny for each year separately on the ‘Latham’ parental genetic map, the genotype data set for total ACY, and IM. Two QTL were identified: one on ‘Latham’ RLG3 associated with marker RiF2KP in all 3 years (2009: LOD = 8.3, $R^2 = 0.22$; 2010: LOD = 5.1, $R^2 = 0.14$; 2011: LOD = 2.6, $R^2 = 0.08$) and one on ‘Latham’ RLG2 associated with marker RubFruitE4 (2010 only, LOD = 5.8, $R^2 = 0.16$) (Supplemental Fig. S6).

Candidate genes co-locating with QTL

On the ‘Latham’ parental map, two polyphenolic biosynthetic pathway gene-derived markers (RiF3’H and RiPAL) and four transcription factor (TF)-derived markers (RubFruitE4, RibHLH, RiMYB, RubFAC-003_MYB1) are associated with QTL on three different RLG (Table 11, Supplemental Fig. S4). No QTL on the 96395S1 parental

Fig. 2 Ranking of individuals based on the percentage of cyanidin 3-*O*-sophoroside (C3S) of the total anthocyanin (ACY) accumulated in 2009. This graph suggests that, generally, percent C3S accumulation in 2009 (*diamonds*) agreed with percent C3S accumulation in 2010 (*squares*). Percent C3S accumulation in 2011 was the most variable (*triangles*), with the low and high extremes trending to the opposite of 2009



map are associated with polyphenolic biosynthetic pathway gene- or TF-based markers.

Association of QTL with markers derived from polyphenolic biosynthetic pathway genes and transcription factors

Marker RiF3'H on RLG5 that was designed from *flavanone 3'-hydroxylase (F3'H)* sequence encoding the enzyme that catalyzes an early step in flavonoid metabolism, provides precursors for many classes of flavonoid compounds, and regulates the flavonoid pathway (Pelletier and Shirley 1996). On 'Latham' RLG5, QTL for P3R (2009, 2010) and C3GR 2010 co-located with the candidate gene-based marker RiF3'H. On 'Latham' RLG6, marker Ri4Co1, designed from the gene encoding 4-coumarate:coA ligase 1 which catalyzes the formation of 4-coumaroyl-CoA during the formation of flavonoids (Kuhn et al. 1984), is associated with, or is within the LOD \pm 1 confidence interval of, the QTL for C3S 2009. On 'Latham' RLG7a, marker RiPAL2, designed from the candidate gene *phenylalanine lyase (PAL)* that encodes the enzyme for biosynthesis of a wide range of phenylpropanoid products from phenylalanine (Liang et al. 1989), is within 2 cM of marker Ri_5O21, as well as within the LOD \pm 1 confidence interval of the QTL for co-eluting C3G_C3Sb 2009.

QTL for all traits analyzed on 'Latham' RLG2 co-segregated, with three recombinants (Fig. 1), with marker RubFruitE4, a putative *bZIP* TF (Kassim et al. 2009). On 'Latham' RLG7b, the marker RibHLH, designed from a TF similar to apple *MdbHLH33*, co-located with QTL for C3S (2009, 2010), co-eluting compounds C3XR_C3R 2010, C3G 2010, and C3GR 2011. Marker RiMYB, derived from a putative *MYB* TF, on 'Latham' RLG6 co-locates with, or is within the LOD \pm 1 confidence interval of, the QTL for C3XR + C3R 2009. Marker RiMYB is just outside the

confidence interval for QTL for total ACY and yield identified by McCallum et al. (2010).

Association of QTL with markers derived from additional gene sequences

On 'Latham' RLG3, QTL for traits C3GR (2009, 2010), C3G (2010) and C3G_C3Sb (2009, 2011), C3S (2009, 2010), and C3XR + C3R (2009) co-located with marker RiF2KP, designed from a regulatory enzyme-encoding gene (Draborg et al. 1999). On 96395S1 RLG3, marker Ri60S, which is similar to the 60S ribosomal protein L10 (Woodhead et al. 2010), co-located with three of the same four QTL (C3S, C3G, C3G_C3Sb, and C3GR) mapped on 'Latham' RLG3. Finally, on 'Latham' RLG6, marker RiRosCOS2381 co-located with QTL for C3G_C3Sb 2009, C3G 2010, and C3S 2010.

Comparative gene analysis of the top of RLG2

The genomic region surrounding the QTL associated with marker RubFruitE4 on RLG2 was examined for positional candidate genes, based on synteny with *Fragaria vesca* (alpine strawberry) (Bushakra et al. 2012) (Supplemental Fig. S7). The gene model prediction identified 409 gene models within the 4-Mb region of *Fragaria* LGII surrounding the homolog of *Rubus* marker RubFruitE4. Of these 409 gene models, 59 were considered as potential candidates for, or regulators of, genes involved in the phenylpropanoid biosynthetic pathway. Within this region, genes T51560 (probable *flavanol 3-O-glucosyltransferase* (EC 2.4.1.91), F2K13_200 [similarity]—*Arabidopsis thaliana*), and AAD17392.1 (putative *glucosyltransferase* [*Arabidopsis thaliana*]) are the most promising candidates, as their function is to catalyze the addition of sugar moieties to the cyanidin molecule.

Discussion

Individual anthocyanin concentrations in the F₁ progeny of 96395S1 × ‘Latham’

The interspecific progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* ‘Latham’ provides a genetic background in addition to the ‘Latham’ × ‘Glen Moy’ population (Graham et al. 2004) that can be used for analysis of the genetic and environmental controls of the ACY biosynthetic pathway and subsequent compound accumulation in the fruit of purple raspberry. The comparison of the 96395S1 × ‘Latham’ F₁ progeny subset with other black raspberry, red raspberry, and purple raspberry populations suggests that the concentrations of individual ACY differ widely between individuals, although the proportions and composition of the compounds appear to be consistent within a population. This indicates that the proportions and compositions of ACY are under strong genetic control, while ACY concentration is more subject to environmental influences.

Pearson’s correlation coefficient analyses

The positive correlation between C3G and the other compounds suggests that the enzyme-mediated addition of sugar moieties to this base molecule to produce other compounds does not deplete C3G (Fig. 3). The strong positive correlation between C3S and C3GR is interesting

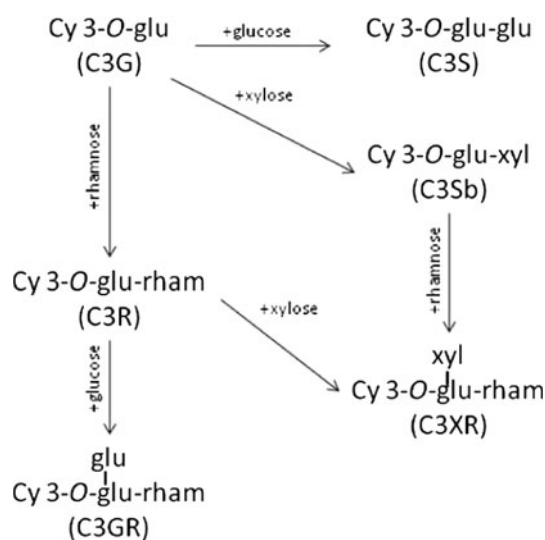


Fig. 3 Six cyanidin-based anthocyanin compounds found in red and black raspberry fruit. The addition of sugar moieties (above the arrows) modifies cyanidin 3-O-glucoside (C3G). C3S cyanidin 3-O-sophoroside, C3XR cyanidin 3-O-2G-xylosylrutinoside, C3R cyanidin 3-O-rutinoside, C3G cyanidin 3-O-glucoside, C3Sb cyanidin 3-O-sambubioside; C3GR cyanidin 3-O-2G-glucosylrutinoside

as they are both end-products resulting from the enzyme-mediated addition of glucose. This positive correlation suggests that glucose is prevalent in the fruit (Kafkas et al. 2008).

‘Latham’ *Rubus* linkage groups 2 and 7 and cyanidin modification

The comparisons conducted are based on markers in common between this and the other four *Rubus* QTL studies as presented in the introduction; however, we use the linkage group nomenclature which is based on the alignment of *Rubus* with *Fragaria vesca* (Bushakra et al. 2012). Dobson et al. (2012) conducted an analysis of the accumulation of total ACY in the ‘Latham’ × ‘Glen Moy’ progeny and identified two QTL on LG2 and LG6 (our RLG3 and RLG1, respectively). The analysis of individual compounds conducted in the same population by Kassim et al. (2009), identified two QTL on LG4 and LG1 (our RLG2 and RLG7, respectively). McCallum et al. (2010) present data based on Kassim et al.’s (2009) analysis of compounds that indicate that heterozygous loci at FruitE4 and at bHLH both independently influence the concentration of some of the compounds studied. Our study supports these QTL and validates that the QTL associated with markers RubFruitE4 and RubbHLH are involved in the sugar modification of cyanidin, rather than the production of total anthocyanins.

On RLG7, marker RubbHLH was designed from *Rubus* sequence similar to *MdbHLH33*, which encodes a component of the MYB/bHLH/WD protein complex involved in the regulation of the expression of ACY biosynthetic pathway genes and the subsequent accumulation of ACY in plant tissues (Allan et al. 2008; Baudry et al. 2004). On RLG2, apart from marker RubFruitE4 derived from a putative *bZIP* TF, which exhibited three recombinants with the phenotype (Fig. 1), no TF or biosynthetic gene-based marker co-segregated with the QTL. Comparative gene analysis of RLG2 with the homologous region of the *Fragaria* whole genome sequence (Shulaev et al. 2011) identified two genes that encode enzymes that may be responsible for the addition of sugars to the cyanidin molecule: a putative *flavonol 3-O-glucosyltransferase* (*F3OGT*) whose enzyme product is involved in the glycosylation of kaempferol (Tohge et al. 2005), and a putative *flavonol 7-O-glucosyltransferase* (*F7OGT*), the enzyme product of which catalyzes the transfer of a glucosyl group from UDP-glucose to the 7-hydroxy group of a quercetin molecule (Lim et al. 2004). These two genes, along with *Rubus bHLH*, would be candidates for gene expression analysis in tobacco (Espley et al. 2007), with *in planta* gene expression in *Rubus* progeny exhibiting the extremes of specific ACY compound composition.

Comparison of QTL identified in the progeny of 96395S1 × ‘Latham’ with those identified previously in the progeny of ‘Latham’ × ‘Glen Moy’

The QTL for total ACY accumulation in all 3 years that is located on ‘Latham’ RLG3 and associated with marker RiF2KP may be the same as the QTL identified by Dobson et al. (2012) (LG2 in their nomenclature). The QTL for total ACY accumulation in 2010 located on ‘Latham’ RLG2 and associated with marker RubFruitE4 were not identified by Dobson et al. (2012) and in the current study has a much lower LOD score and explains a much lower percentage of variation than the QTL for individual compounds. This finding could indicate that the large-effect QTL associated with marker RubFruitE4 on ‘Latham’ RLG2 may be associated with the modification of the base cyanidin molecule, rather than production of the cyanidin molecule, reinforcing F3OGT and F7OGT as candidates for enzyme-mediated sugar addition.

QTL for levels of eight ACY compounds [cyanidin 3-*O*-sophoroside (C3S), cyanidin 3-*O*-2^G-glucosylrutinoside (C3GR), cyanidin 3-*O*-glucoside (C3G), cyanidin 3-*O*-rutinoside (C3R), pelargonidin 3-*O*-rutinoside (P3R), pelargonidin 3-*O*-sophoroside (P3S), pelargonidin 3-*O*-glucoside (P3G) and pelargonidin 3-*O*-glucosylrutinoside (P3GR)], were previously mapped in association with marker RubFruitE4 and marker bHLH on the linkage map of ‘Latham’ × ‘Glen Moy’ (Kassim et al. 2009). QTL for five (C3S, C3GR, C3G, C3R, and P3R) of these eight compounds have been identified in this study and are also associated with the locus RubFruitE4. Neither C3Sb nor C3XR were reported as being detected in the progeny of ‘Latham’ × ‘Glen Moy’ (Graham et al. 2004). In the current study, the greatest amount of variation in expression of all six traits (14.9–64.3 %) is explained by locus RubFruitE4 on ‘Latham’ RLG2 (Table 10). This locus on the ‘Latham’ × ‘Glen Moy’ linkage map explains less (between 10 and 40 %) of the variation in expression of levels of the eight compounds (Kassim et al. 2009). The difference between these two studies in percent variation explained by this locus could result from the presence of black raspberry-specific xylose-containing compounds in the progeny of 96395S1 × ‘Latham’, as the ‘Latham’ × ‘Glen Moy’ progeny lack these compounds. The enzyme-mediated addition of another sugar moiety to the base cyanidin molecule is hence supported by this finding.

In the current study, markers RibHLH and RiFRK2 mapping within 3.2 cM of each other on ‘Latham’ RLG7b are associated with 6.8–11.2 % of the variation in expression of C3S 2009 and C3G 2010 in the set of 96395S1 × ‘Latham’ progeny analyzed, while locus bHLH on the ‘Latham’ × ‘Glen Moy’ linkage map explains 35–60 % of the variation in the expression of the

eight compounds detected in the progeny of ‘Latham’ × ‘Glen Moy’ by Kassim et al. (2009). Compound C3XR was not reported as being detected in the progeny of ‘Latham’ × ‘Glen Moy’ by Kassim et al. (2009). The difference between these two studies in percent variation explained by this locus could be the result of the more consistent proportions of compounds observed in the progeny of 96395S1 × ‘Latham’ compared with ‘Latham’ × ‘Glen Moy’. For example, excluding P3R, the range of percent compounds in 96395S1 × ‘Latham’ is 10–39 %, whereas for ‘Latham’ × ‘Glen Moy’ progeny the range is 1.6–64 % (Table 12).

On ‘Latham’ RLG6, the QTL for C3XR + C3R 2009, C3XR_C3R 2010, and C3GR 2010 are all associated with, or within the LOD ± 2 confidence interval of, marker RiMYB, which was designed from *Rubus* sequence homologous to TF *MdMYB10* (Lin-Wang et al. 2010) which, like *bHLH*, encodes a component of the MYB/bHLH/WD protein complex (Baudry et al. 2004). *MdMYB10* was isolated from red-fleshed apple and is responsible for controlling ACY biosynthesis in apple flesh (Espley et al. 2007). Additionally, *PavMYB10* from sweet cherry (*Prunus avium*) is located within the confidence interval for a QTL associated with fruit color on the ‘New York 54’ × ‘Emperor Francis’ consensus linkage map (Sooriyapathirana et al. 2010). RiMYB maps within 7 cM of candidate gene marker Ri4Co1 and within 5 cM of marker Rub2a1, which is shared between ‘Latham’ RLG6 and ‘Latham’ × ‘Glen Moy’ LG3. Marker Rub2a1 falls within the 95 % confidence interval of the QTL for total ACY identified by McCallum et al. (2010) and within 8 cM of Ri4CL1, which is derived from the same sequence as marker Ri4Co1 on ‘Latham’ RLG6.

QTL have been mapped to neither ‘Latham’ nor 96395S1 RLG4, nor its equivalent in the other *Rubus* studies. In summary, each linkage group, except RLG4, appears to carry one to several loci involved in the control of various aspects of fruit quality in *Rubus* spp. The markers that are associated with stable QTL (Table 10), for example RubFruitE4, RibHLH, RiRosCOS2381 and RiF3’H, would be good candidates for marker-assisted breeding (MAB) applications.

Conclusions

The accumulation of total ACY and the accumulation of individual compounds in the fruit of *Rubus* appear to be controlled by different regions of the genome, as QTL for total ACY are located on RLG2 and RLG3, while QTL that control accumulation of individual compounds are located on every linkage group except RLG4, with especially important QTL on RLG2 and RLG7. The cause of the

differences in the production of cyanidin 3-*O*-2^G-glucosylrutinoside (C3GR) by progeny, as demonstrated by the division of the analyzed progeny into two sets of individuals, indicates the need for closer examination of the genetic control of that trait. Additional comparative gene analysis of the region underlying the RLG2 QTL, as well as analysis of the forthcoming *Rubus idaeus* ‘Heritage’ genome sequence and gene annotation, will help to identify candidate genes that may encode the enzymes necessary to catalyze the reactions that produce individual ACY compounds.

Acknowledgments This work was completed as part of a Ph.D. thesis and supported in part by the PFR Excellence Program (JMB) and the “New Berries” FRST Programme (CO6XO807) (EJB). The authors would like to thank Andrew McLachlan (PFR) for help with the statistics; Peter Alspach (PFR) for help with the epistasis analysis; the Editor and three anonymous reviewers for helpful suggestions on the original manuscript.

References

- Allan AC, Hellens RP, Laing WA (2008) MYB transcription factors that colour our fruit. *Trends Plant Sci* 13:99–102
- Ash MM, Wolford KA, Carden TJ, Hwang KT, Carr TP (2011) Unrefined and refined black raspberry seed oils significantly lower triglycerides and moderately affect cholesterol metabolism in male Syrian hamsters. *J Med Food* 14:1032–1038
- Baudry A, Heim M, Dubreucq B, Caboche M, Weisshaar B, Lepiniec L (2004) TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *Plant J* 39:366–380
- Bushakra J, Stephens M, Atmadjaja A, Lewers K, Symonds V, Udall J, Chagné D, Buck E, Gardiner S (2012) Construction of black (*Rubus occidentalis*) and red (*R. idaeus*) raspberry linkage maps and their comparison to the genomes of strawberry, apple, and peach. *Theor Appl Genet* 125(2):311–327
- Cabrera A, Kozik A, Howad W, Arús P, Iezzoni A, van der Knaap E (2009) Development and bin mapping of a Rosaceae Conserved Ortholog Set (COS) of markers. *BMC Genom* 10:562
- Chen HS, Liu M, Shi LJ, Zhao JL, Zhang CP, Lin LQ, Liu Y, Zhang SJ, Jin JC, Wang L, Shen BZ, Liu JR (2011) Effects of raspberry phytochemical extract on cell proliferation, apoptosis, and serum proteomics in a rat model. *J Food Sci* 76:T192–T198
- Connor AM, Finn CE, McGhie TK, Alspach PA (2005) Genetic and environmental variation in anthocyanins and their relationship to antioxidant activity in blackberry and hybridberry cultivars. *J Am Soc Hortic Sci* 130:680–687
- Cooney JM, Jensen DJ, McGhie TK (2004) LC-MS identification of anthocyanins in boysenberry extract and anthocyanin metabolites in human urine following dosing. *J Sci Food Agric* 84:237–245
- Dobson P, Graham J, Stewart D, Brennan R, Hackett CA, McDougall GJ (2012) Over-seasons analysis of quantitative trait loci affecting phenolic content and antioxidant capacity in raspberry. *J Agric Food Chem* 60:5360–5366
- Dossett M, Lee J, Finn CE (2010) Variation in anthocyanins and total phenolics of black raspberry populations. *J Funct Foods* 2:292–297
- Dossett M, Lee J, Finn CE (2011) Characterization of a novel anthocyanin profile in wild black raspberry mutants: an opportunity for studying the genetic control of pigment and color. *J Funct Foods* 3:207–214
- Draborg H, Villadsen D, Hamborg Nielsen T (1999) Cloning, characterization and expression of a bifunctional fructose-6-phosphate, 2-kinase/fructose-2,6-bisphosphatase from potato. *Plant Mol Biol* 39:709–720
- Espín JC, García-Conesa MT, Tomás-Barberán FA (2007) Nutraceuticals: facts and fiction. *Phytochemistry* 68:2986–3008
- Espley RV, Hellens RP, Putterill J, Stevenson DE, Kuty-Amma S, Allan AC (2007) Red colouration in apple fruit is due to the activity of the MYB transcription factor, *MdMYB10*. *Plant J* 49:414–427
- Finley JW, Kong A-N, Hintze KJ, Jeffery EH, Ji LL, Lei XG (2011) Antioxidants in foods: State of the science important to the food industry. *J Agric Food Chem* 59(13):6837–6846
- Graham J, Smith K, MacKenzie K, Jorgenson L, Hackett C, Powell W (2004) The construction of a genetic linkage map of red raspberry (*Rubus idaeus* subsp. *idaeus*) based on AFLPs, genomic-SSR and EST-SSR markers. *Theor Appl Genet* 109:740–749
- Graham J, Hackett C, Smith K, Woodhead M, Hein I, McCallum S (2009) Mapping QTLs for developmental traits in raspberry from bud break to ripe fruit. *Theor Appl Genet* 118:1143–1155
- Grotewold E (2006) The genetics and biochemistry of floral pigments. *Ann Rev Plant Biol* 57:761–780
- Holton TA, Cornish EC (1995) Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell* 7:1071–1083
- Hummer KE (2010) *Rubus* pharmacology: antiquity to the present. *HortSci* 45:1587–1591
- Jennings DL, Carmichael E (1980) Anthocyanin variation in the genus *Rubus*. *New Phytol* 84:505–513
- Kafkas E, Özgen M, Özoğui Y, Türemitiş N (2008) Phytochemical and fatty acid profile of selected red raspberry cultivars: a comparative study. *J Food Quality* 31:67–78
- Kassim A, Poette J, Paterson A, Zait D, McCallum S, Woodhead M, Smith K, Hackett C, Graham J (2009) Environmental and seasonal influences on red raspberry anthocyanin antioxidant contents and identification of quantitative traits loci (QTL). *Mol Nutr Food Res* 53:625–634
- Keniry A, Hopkins CJ, Jewell E, Morrison B, Spangenberg GC, Edwards D, Batley J (2006) Identification and characterization of simple sequence repeat (SSR) markers from *Fragaria × ananassa* expressed sequences. *Mol Ecol Notes* 6:319–322
- Kuhn DN, Chappell J, Boudet A, Hahlbrock K (1984) Induction of phenylalanine ammonia-lyase and 4-coumarate: CoA ligase mRNAs in cultured plant cells by UV light or fungal elicitor. *Proc Nat Acad Sci* 81:1102–1106
- Liang XW, Dron M, Cramer CL, Dixon RA, Lamb CJ (1989) Differential regulation of phenylalanine ammonia-lyase genes during plant development and by environmental cues. *J Biol Chem* 264:14486–14492
- Lim E-K, Ashford DA, Hou B, Jackson RG, Bowles DJ (2004) *Arabidopsis* glycosyltransferases as biocatalysts in fermentation for regioselective synthesis of diverse quercetin glucosides. *Biotechnol Bioeng* 87:623–631
- Lin-Wang K, Bolitho K, Grafton K, Kortstee A, Karunairetnam S, McGhie T, Espley R, Hellens R, Allan A (2010) An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in Rosaceae. *BMC Plant Biol* 10:50
- Määttä-Riihinen KR, Kamal-Eldin A, Törrönen AR (2004) Identification and quantification of phenolic compounds in berries of *Fragaria* and *Rubus* species (Family Rosaceae). *J Agric Food Chem* 52:6178–6187
- McCallum S, Woodhead M, Hackett C, Kassim A, Paterson A, Graham J (2010) Genetic and environmental effects influencing fruit colour and QTL analysis in raspberry. *Theor Appl Genet* 121:611–627

- McGhie TK, Hall HK, Ainge GD, Mowat AD (2002) Breeding *Rubus* cultivars for high anthocyanin content and high antioxidant capacity. *Acta Hort (ISHS)* 585:495–500
- Montrose DC, Horelik NA, Madigan JP, Stoner GD, Wang L-S, Bruno RS, Park HJ, Giardina C, Rosenberg DW (2011) Anti-inflammatory effects of freeze-dried black raspberry powder in ulcerative colitis. *Carcinogenesis* 32:343–350
- Mullen W, Yokota T, Lean MEJ, Crozier A (2003) Analysis of ellagitannins and conjugates of ellagic acid and quercetin in raspberry fruit by LC-MSⁿ. *Phytochemistry* 64:617–624
- Ozgen M, Wyzgoski FJ, Tulio AZ, Gazula A, Miller AR, Scheerens JC, Reese RN, Wright SR (2008) Antioxidant capacity and phenolic antioxidants of midwestern black raspberries grown for direct markets are influenced by production site. *HortSci* 43:2039–2047
- Pelletier MK, Shirley BW (1996) Analysis of flavanone 3-hydroxylase in *Arabidopsis* seedlings (coordinate regulation with chalcone synthase and chalcone isomerase). *Plant Physiol* 111:339–345
- Sargent D, Rys A, Nier S, Simpson D, Tobutt K (2007) The development and mapping of functional markers in *Fragaria* and their transferability and potential for mapping in other genera. *Theor Appl Genet* 114:373–384
- Scalzo J, Currie A, Stephens J, McGhie T, Alspach P (2008) The anthocyanin composition of different *Vaccinium*, *Ribes*, and *Rubus* genotypes. *BioFactors* 34:13–21
- Seeram NP, Momin RA, Nair MG, Bourquin LD (2001) Cyclooxygenase inhibitory and antioxidant cyanidin glycosides in cherries and berries. *Phytomedicine* 8:362–369
- Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP, Burns P, Davis TM, Slovin JP, Bassil N, Hellens RP, Evans C, Harkins T, Kodira C, Desany B, Crasta OR, Jensen RV, Allan AC, Michael TP, Setubal JC, Celton J-M, Rees DJG, Williams KP, Holt SH, Rojas JJR, Chatterjee M, Liu B, Silva H, Meisel L, Adato A, Filichkin SA, Troggio M, Viola R, Ashman T-L, Wang H, Dharmawardhana P, Elser J, Raja R, Priest HD, Bryant DW, Fox SE, Givan SA, Wilhelm LJ, Naithani S, Christoffels A, Salama DY, Carter J, Girona EL, Zdepski A, Wang W, Kerstetter RA, Schwab W, Korban SS, Davik J, Monfort A, Denoyes-Rothan B, Arús P, Mittler R, Flinn B, Aharoni A, Bennetzen JL, Salzberg SL, Dickerman AW, Velasco R, Borodovsky M, Veilleux RE, Folta KM (2011) The genome of woodland strawberry (*Fragaria vesca*). *Nat Genet* 43:109–116
- Sooriyapathirana S, Khan A, Sebolt A, Wang D, Bushakra J, Lin-Wang K, Allan A, Gardiner S, Chagné D, Iezzoni A (2010) QTL analysis and candidate gene mapping for skin and flesh color in sweet cherry fruit (*Prunus avium* L.). *Tree Genet Genomes* 6:821–832
- Takos AM, Ubi BE, Robinson SP, Walker AR (2006) Condensed tannin biosynthesis genes are regulated separately from other flavonoid biosynthesis genes in apple fruit skin. *Plant Sci* 170:487–499
- Tanaka Y, Sasaki N, Ohmiya A (2008) Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J* 54:733–749
- Thompson MM (1995) Chromosome numbers of *Rubus* species at the National Clonal Germplasm Repository. *HortSci* 30:1447–1452
- Tohge T, Nishiyama Y, Hirai M, Yano M, Nakajima J, Awazuhara M, Inoue E, Takahashi H, Goodenowe D, Kitayama M, Noji M, Yamazaki M, Saito K (2005) Functional genomics by integrated analysis of metabolome and transcriptome of *Arabidopsis* plants over-expressing an MYB transcription factor. *Plant J* 42:218–235
- Torre LC, Barritt BH (1977) Quantitative evaluation of *Rubus* fruit anthocyanin pigments. *J Food Sci* 42:488–490
- Tulio AZ, Reese RN, Wyzgoski FJ, Rinaldi PL, Fu R, Scheerens JC, Miller AR (2008) Cyanidin 3-rutinoside and cyanidin 3-xylosylrutinoside as primary phenolic antioxidants in black raspberry. *J Agric Food Chem* 56:1880–1888
- van Ooijen JW (1992) Accuracy of mapping quantitative trait loci in autogamous species. *Theor Appl Genet* 84:803–811
- van Ooijen JW (2004) MapQTL[®] 5, Software for the mapping of quantitative trait loci in experimental populations. Kyazma, B.V., Wageningen, The Netherlands
- Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M, Coppola G, Costa F, Cova V, Dal Ri A, Goremykin V, Komjanc M, Longhi S, Magnago P, Malacarne G, Malnoy M, Micheletti D, Moretto M, Perazzolli M, Si-Ammour A, Vezzulli S, Zini E, Eldredge G, Fitzgerald LM, Gutin N, Lanchbury J, Macalma T, Mitchell JT, Reid J, Wardell B, Kodira C, Chen Z, Desany B, Niazi F, Palmer M, Koepke T, Jiwan D, Schaeffer S, Krishnan V, Wu C, Chu VT, King ST, Vick J, Tao Q, Mráz A, Stormo A, Stormo K, Bogden R, Ederle D, Stella A, Vecchiotti A, Kater MM, Masiero S, Lasserre P, Lespinasse Y, Allan AC, Bus V, Chagne D, Crowhurst RN, Gleave AP, Lavezzo E, Fawcett JA, Proost S, Rouze P, Sterck L, Toppo S, Lazzari B, Hellens RP, Durel C-E, Gutin A, Bumgarner RE, Gardiner SE, Skolnick M, Egholm M, Van de Peer Y, Salamini F, Viola R (2010) The genome of the domesticated apple (*Malus × domestica* Borkh.). *Nat Genet* 42:833–839
- Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered* 93:77–78
- VSN International L (2010) GenStat 12.2.0.3717 edn
- Woodhead M, Weir A, Smith K, McCallum S, MacKenzie K, Graham J (2010) Functional markers for red raspberry. *J Am Soc Hortic Sci* 135:418–427