CrossMark

Identification of major stable QTLs for flower color in roses

A. Henz · T. Debener · M. Linde

Abstract Flower color is one of the most important traits of ornamental roses. Anthocyanins are the major secondary metabolites responsible for the red and pink colors found among rose cultivars. Color varies depending on the combination of particular anthocyanins, their co-factors and their concentrations. Several genetic investigations have indicated that variation in flower color is dependent on monogenic factors and quantitative trait loci (QTL). Here, we analyze quantitative variation of total anthocyanins in diploid rose progeny. We demonstrate that the environment produces relatively small effects; the main causes of variation in anthocyanin content are the genetic differences between individuals. Two major QTLs were detected in all six tested environments. Four additional QTLs were found only in a subset of the environments. Some of the QTLs either co-segregate or are located close to the map positions of known structural genes of the anthocyanin biosynthesis pathway or transcriptional regulators of anthocyanin biosynthesis. This information might be used to characterize tetraploid parental genotypes for their potential to pass on higher anthocyanin contents to their progeny.

Introduction

The pigmentation of angiosperm flowers has important functions in attracting pollinators in many zoophilic plant species (Schistl and Johnson 2003). Flower color is one of the most important traits in ornamental horticulture, as it determines to a large extent the attractiveness of floricultural crops.

Due to the high economic importance of ornamental roses among floricultural crops, many studies have been conducted on the chemical composition of rose flowers (Jay et al. 2003; Forkmann 2003). The color of rose flowers is dominated by flavonoids and, to a smaller extent, by carotenoids (De Vries et al. 1974). Among the flavonoids, anthocyanin glucosides, mainly cyanidin and pelargonidin glucosides, are the prominent colorigenic anthocyanidins (Biolley et al. 1994; Mikanagi et al. 1995; Jay et al. 2003). Flavonoid biosynthesis has been well studied and is highly conserved among plants (Grotewold 2006; Tanaka et al. 2008). Therefore, many of the structural and

A. Henz · T. Debener (✉) · M. Linde
Institut für Pflanzengenetik, Leibniz Universität
Hannover, Herrnhäuser Str. 2, 30419 Hannover, Germany
e-mail: debener@genetik.uni-hannover.de

A. Henz
Division of Pharmacognosy, Department of Medicinal
Chemistry, BMC - Biomedical Center, Uppsala
University, Box 574, 751 23 Uppsala, Sweden

regulatory genes have been isolated and characterized in roses. However, the functional characterization of these genes has been hampered by a lack of efficient transformation systems (Forkmann 2003; Debener and Hibrand-Saint Oyant 2009). Therefore, no direct link between variants of these genes and variation in flower color has been demonstrated in roses.

Ornamental roses are mostly tetraploid complex multispecies hybrids, to which at least ten different rose species have contributed genetic material (Gudin 2000). As the composition of colorigenic substances is mainly genetically determined, a vast diversity of colors is found among commercial rose varieties (Jay et al. 2003). These vary not only in the composition of colorigenic substances but also in their concentration. In addition to genetic factors, the intensity and composition of flower color are also dependent on environmental factors and the developmental stage of the flower. Common observations are that red and pink colors in flowers that are based on anthocyanins fade with the onset of senescence and that color intensity is influenced by soil parameters such as pH (Schmitzer et al. 2010; Schmitzer and Stampar 2010).

A number of genetic studies in roses indicated that individual, mostly codominant loci are involved in the expression of red, pink and yellow colors in flowers (Marshall et al. 1983, De Vries et al. 1980; Byrne 2009). The construction of various molecular marker maps in diploid and tetraploid roses revealed the chromosomal position of some of the major loci involved in flower color expression (Debener and Mattiesch 1999; Byrne 2009; Spiller et al. 2011). However, the observed variation in flower color in roses indicates that quantitative variation in flower color intensity is frequently observed in rose populations. Surprisingly, no detailed study has been conducted thus far on the loci involved in quantitative variation in color intensity that considers different environmental conditions. However, knowledge of the number of loci involved in the expression of color intensity is important for the development of sophisticated breeding strategies.

The aim of the current study was therefore to analyze genetic factors influencing total anthocyanin content in young rose flowers of diploid progeny, to map these factors to a genetic linkage map of roses and to assess the influence of the environment on QTLs for flower color.

Materials and methods

Plant material

The F_1 population 97/7 used in this study originates from a *R. multiflora* background and is comprised of 270 individuals. This population was derived from 95/13-39 × 82/78-1 cross, as described by Linde et al. (2006). Three clonally propagated copies of the plant population have been grown in the following environments: in a greenhouse [G] in soil under semi-controlled conditions (no extra lighting, temperature regulated through roof ventilation) and at two different locations under field conditions (an experimental field plot at the Faculty of Natural Sciences [F] in Herren-hausen, Hannover, and an experimental field plot in Ruthe [R], south east of Hannover, Germany). These three clonal populations were sampled during the years 2009, 2010 and 2011. This resulted in measurements for the following five environments: R-2009, F-2010, F-2011, G-2011 and R-2011. DNA extractions were performed from leaf samples from the greenhouse population as described in Linde et al. (2006).

Phenotypic analysis of anthocyanin extracts

Petal samples were collected in 2009–2011, from the beginning of April until the middle of June. To avoid bleaching of the anthocyanins, sampling was conducted in the morning at 8–12 a.m. To define the anthocyanin content, three to five recently (at the day of sampling) opened buds at flower development stage 1–2 were selected from each genotype and were kept on ice until weighing. Flower developmental stages were determined according to Picone et al. (2004). Total anthocyanins were extracted from 50 mg of petals in 1 ml of MeOH/HCl (99:1) at room temperature for 4 h or overnight at 4 °C. For each environment, at least three repetitions per genotype were carried out on different days. The total anthocyanin content was measured with a UV mc2 SAFAS photometer at 525 nm in 1-ml disposable cuvettes. If necessary, the extracts were diluted to keep the measured OD_{525} at 0.2–2. For diluted samples, the OD was then calculated by multiplying the measured OD with the dilution factor.

Development of candidate gene markers

To establish markers for genes within the anthocyanin biosynthesis pathway, we used sequence data from *Rosa hybrida* and *Rosa rugosa* deposited at NCBI (http://www.ncbi.nlm.nih.gov/). The primers were designed using Primer3 (http://frodo.wi.mit.edu/primer3/) and tested with the parental genotypes and a few progeny plants of the mapping population 97/7. The presence of sequence polymorphisms was determined via single-strand conformation polymorphism (SSCP) gel analysis (Orita et al. 1989). The SSCP analysis and the fluorescence labeling of PCR products according to Schuelke (2000) using M13-tailing followed a modified procedure as described by Spiller et al. (2010). These markers are listed in supplementary Table 1.

Biostatistics

Estimation of genotype–environment interactions

Statistics were calculated either with MS Excel 2007 or with the statistics software [R] (R Development Core Team 2015). The data were tested for normal distribution using a nonparametric Kolmogorov–Smirnov test ($\alpha = 0.05$) (WinSTAT MS Excel 2007).

To estimate the influence of the environment on the anthocyanin content, we used the data obtained in 2011 from the environments F, G and R. From these three environments, we could obtain measurements for the largest number of identical individuals in the three plots.

The genotype–environment interactions were calculated according to the formula $P = G + E + GE$ (Clausen et al. 1940). Forty-four genotypes were chosen that had a coefficient of variance of ≤ 0.25 in their anthocyanin content.

Linkage mapping and QTL analysis

The genetic linkage map of the 97/7 population (Linde et al. 2006), with updated data published by Spiller et al. (2011), was used for the QTL analysis. DNA samples which had to be replaced were extracted with the Qiagen DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions using 70 mg of dried leaf tissue. The molecular markers developed in this study were integrated into the Spiller consensus map using JoinMap4 (van Ooijen 2006) and the Kosambi mapping function with the same settings used by Spiller et al. (2010). The independence LOD option was used, and mapping was performed with a recombination frequency threshold of 0.4 and a LOD > 1. The overall grouping LOD was ≥7. Marker segregation data were analyzed in CP (*cross-pollinated*) mode. Independent maps for the parental plants were calculated using the test cross strategy, as described by Stam (1993). Charts were generated using Map-Chart version 2.2 (http://www.biometris.wur.nl/uk/Software/MapChart/). For a more convenient visualization of the map structure, several markers were removed. Only 1–2 markers were displayed per cM. Interval mapping (Lander and Botstein 1989) was performed using the computer software package MapQTL6 (van Ooijen 2004). Only maps from the first round of calculations in JoinMap4 were used for QTL mapping. QTL detection was performed using the mean values of the anthocyanin content of all five environments separately.

Settings for the nonparametric Kruskal–Wallis rank-sum test for single marker influence and for the significance threshold of LOD scores estimated by permutation tests were used as described in Spiller et al. (2011). QTLs were considered significant if they reached at least a chromosome-wide level of significance in interval mapping. A distinction was made between *major QTLs* (at a genome-wide level) and *minor QTLs* (at a chromosome-wide level).

Furthermore, MQM mapping (Jansen and Stam 1994) using MapQTL6 was performed to identify additional putative QTLs, as described in Linde et al. (2006). Those markers were chosen as cofactors, which showed the highest significance with the Kruskal–Wallis rank-sum test in the region around a putative QTL.

Results

Large, genotype-dependent variation in anthocyanin content in diploid progeny

The cross between genotypes 95/13-39 (petal color: pink) and 82/78-1 (Sp3) (petal color: red) resulted in the F$_1$ progeny 97/7, which displayed a large variation in petal color. The spectrum of colors ranged from violet over deep red to light rose (supplementary

Figure 1). The anthocyanin content was analyzed in samples from three consecutive years, 2009–2011, and three different locations. Five independent datasets (referred to as environments) were used (Table 1). Altogether, 904 genotype × environment combinations have been analyzed in the five different environments.

In all five environments, a large variation in the total anthocyanin content was observed. In particular, large variation was observed in the maximum absorption at OD_{525}; genotypes in the greenhouse environment showed the highest OD_{525} values (Table 1). The population in the greenhouse displayed the largest variation in anthocyanin content. Individuals in the field plot in R-2009 had the smallest range. The phenotypic distribution and the variance in the data were similar in all environments. Correlation coefficients between the environments varied from 0.89 to 0.95. Scatter plots of a subset of comparisons based on a set of 80 genotypes that were scored in all five environments are shown in Fig. 1, also displaying the high correlations between the OD_{525} values measured in the different environments. The distribution of anthocyanin contents was right-skewed for all five environments; distinct classes could not be identified. Slightly different variation coefficients for the measured anthocyanin contents were obtained for the different environments; these ranged from 0.14 (R2011) to 0.16 (F2011), and the medial coefficient of variance over all test plots was 0.15.

Low genotype–environment interaction for the anthocyanin content

A comparison of OD values of clones from the same genotypes in three different environments in 2011 showed that there were no significant differences in the variation in anthocyanin content between the environments; the Kruskal–Wallis rank-sum test was used for this comparison ($P = 0.146$).

Some genotypes differed considerably in their anthocyanin content between environments; other genotypes were very stable. The genotype–environment interaction ($P = G + E + GE$) was strongest in genotypes in the greenhouse (-1.440 to $+7.027$ OD). Genotypes grown on the Faculty field plot had the lowest genotype–environment interaction (-2.41 to $+1.537$ OD). However, the general differences between the environments were quite low. This is reflected in the calculation of environmental variance (E). The greenhouse and the field plot in Ruthe showed slightly positive environmental variances (higher anthocyanin content) of $+0.178$ and $+0.151$, respectively. The Faculty plot had a negative variance of -0.263.

Map construction: mapping of candidate genes

The parental plants 95/13-39 and Sp3 and their segregating progeny 97/7 were screened for polymorphisms using SSCP markers. A total of 11 SSCP primer pairs for nine candidate genes for anthocyanin biosynthesis were developed. These generated 23 polymorphic fragments. In total, 17 of the fragments could be mapped in the two parental maps and are shown in supplementary Table 1. Only four markers deviated in the expected segregation using a Chi-square test at $P = 0.05$. The SSCP markers were integrated into the map for population 97/7 from Linde et al. (2006), which was modified by Spiller et al. (2011). A separate map was constructed for each parent, with an overall grouping LOD ≥ 7. Each map

Table 1 Variation in anthocyanin content for the five environments (test area Ruthe = R, Faculty plot = F, greenhouse = G). Values represent the OD_{525}

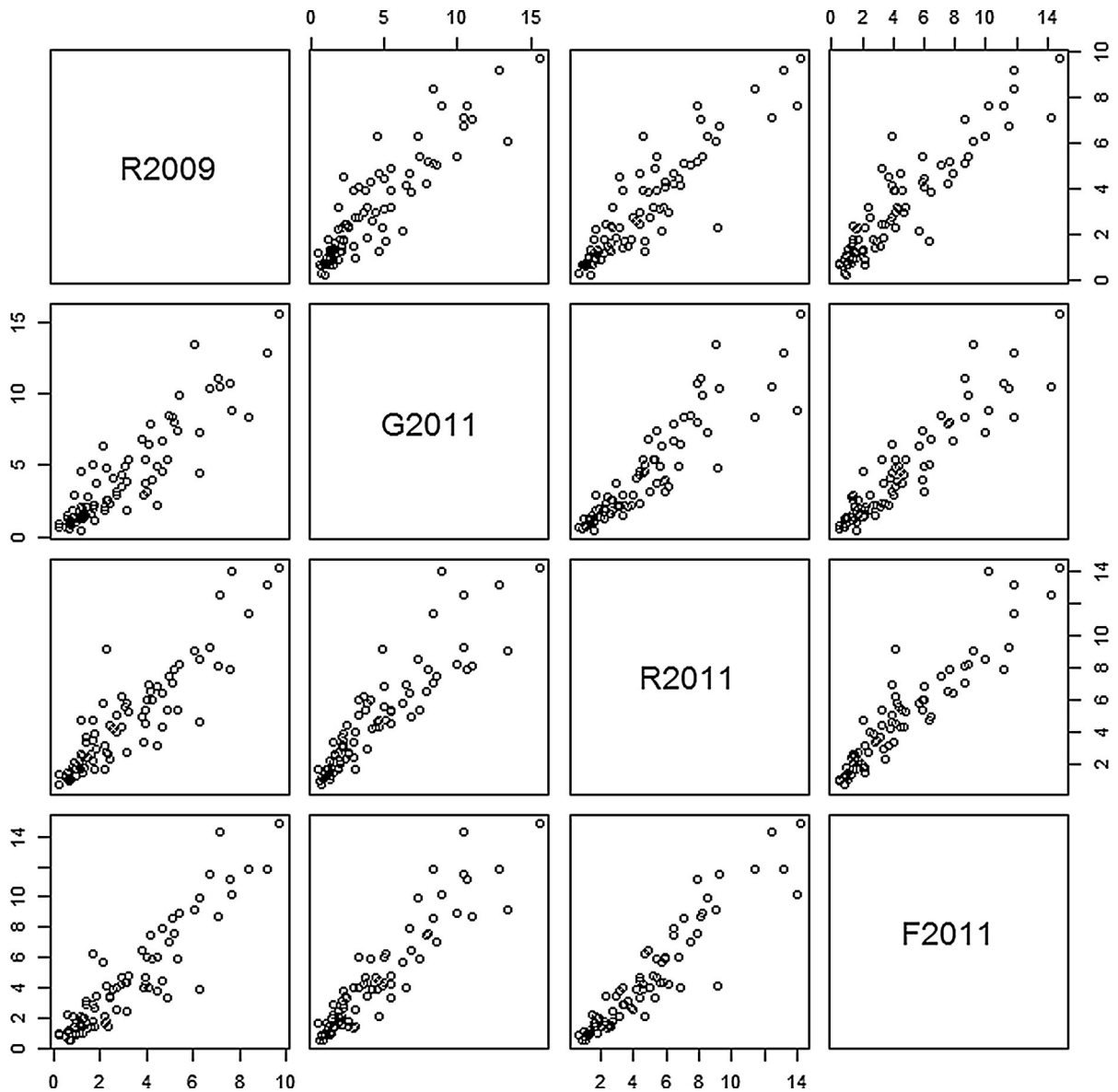| Environment	Number of analyzed genotypes	Min–Max	Mean	Median
R-2009	156	0.235–10.602	3.13	2.423
F-2010	230	0.790–17.0	5.0	3.81
G-2011	210	0.496–23.462	4.483	3.331
F-2011	175	0.483–17.336	4.275	3.386
R-2011	133	0.678–15.19	4.838	4.19

Fig. 1 Selection of scatterplots for OD_{525} values measured for a selection of 80 genotypes scored in all environments. Values on the *x*- and *y*-axis indicate the measured OD_{525} in the respective environment. Four of the five environments were selected due to space constraints. Displayed are from *left to right* and from *top to bottom*: R2009, G2011, R2011, F2011

consisted of seven LGs corresponding to the seven rose chromosomes. The maps for the parental lines 95/13-39 and Sp3 cover total lengths of 689.8 and 738.3 cM and contain 301 and 277 markers, respectively. The mean size per LG is 98.5 and 105.5 cM for 95/13-39 and Sp3, respectively. The maps for 95/13-39 and Sp3 contain means of 43 and 40 markers and average distances of 2.3 and 2.7 cM, respectively.

QTL analysis

The Kolmogorov–Smirnov test ($\alpha = 0.05$) revealed that none of the datasets were normally distributed. The data can be used for QTL analysis; however, the skewed distribution has to be taken into account when the results are interpreted (van Ooijen 1992). A permutation test with 1000 permutations ($\alpha = 0.05$)

was used to determine the significance threshold for QTL interval mapping that resulted in genome-wide LOD scores of 4.1 and 3.9 and chromosome-wide LOD scores of 3.0 and 2.9 for 95/13-39 and Sp3, respectively. In general, QTLs in the same region with overlapping confidence intervals were considered to be the same locus. Kruskal–Wallis tests with Bonferroni-adjusted significance thresholds were used to identify single loci that affect anthocyanin content. Additional QTLs on the same LG were only treated as a separate QTL if the main peaks of the LOD curves were at least 20 cM apart. A total of two major QTLs and four minor QTLs were detected. The detected QTLs could be subdivided into the following two groups: (A) major QTLs with genome-wide LOD scores ≥4.1 in the parental genotype 95/13-39 or ≥3.9 in the parental genotype Sp3, where QTL effects were detected in all five environments, or (B) QTLs that were observed in only some environments (Figs. 2, 3). In Table 2, the locations, LOD scores and 2-LOD confidence intervals are presented.

The two regions with environmentally stable QTLs are located on LG2 and LG6 (Fig. 2). In the center of LG2, the calculated LOD curves exceed the thresholds in the same regions in both parents and are stable over the five analyzed environments. Two highly significant markers (Kruskal–Wallis), AAT1-Intron2 and RMS065, are located in this region. Only the calculated LOD curve for the environment R-11 and the parent 95/13-39 did not exceed the chromosome-wide LOD score (minor QTL). Through MQM mapping, its significance was raised to a LOD of 7.6. The peak of the LOD curve for environment F-10 and the parent Sp3 is slightly shifted; however, the 2 LOD confidence intervals overlap with all other QTL positions. The calculated LOD curves on LG6 indicate a very high influence on anthocyanin content. All curves exceed the genome-wide LOD score continuously over the entire LG. Very high LOD scores, up to 25, are also reached. In particular for the parent 95/13-39, highly significant markers are distributed over the entire LG. The calculated LOD curves exceed the threshold in the same regions close to the highly significant microsatellite marker RMS108-2 on both parental maps. These LOD curves are stable over the five analyzed environments with few exceptions. The calculated QTLs for the parent 95/13-39 are very stable in terms of their positions. The peak of the LOD curve in environment R-09 in the parent Sp3 is shifted,

and the 2-LOD confidence intervals are not overlapping. At the same position, a second QTL was found for the dataset F-11. The marker bHLH is located within the 1-LOD confidence interval. All curves had similar shapes and pointed to two regions on LG6. The QTL for dataset R-09 reaches a LOD score of 16.19 in the same region (∼45 cM) as the other environments, which is almost as high as the LOD score at 62.8 cM (16.55). However, the two peaks were not more than 20 cM apart. LOD score curves for the environments F-10, G-11 and R-11 showed an additional peak in the region of the bHLH marker; however, peaks were not more than 20 cM away from the highest peak. Only in the dataset F-10 were the two peaks more than 20 cM apart. Therefore, this region was considered to be a second QTL. The marker for the transcription factor bHLH showed a significant influence only for R-11. On LG6, the MYB1 marker was mapped (Fig. 2); however, this marker did not have any significant influence on the trait.

On LG2 of both parents, more major QTLs were detected at the top of the LG. Out of five possible environments, four (for 95/13-39) and three (for Sp3) exceeded the genome-wide LOD threshold in the same region. The highest LOD scores were reached almost exclusively at 0.0 cM (cf. Table 2). Confidence intervals of 1 LOD were located in regions without any mapped markers. For the evaluation, this has to be taken into account. On LG6, an additional minor QTL was found for R-09 through MQM mapping. An LOD score of 2.98 only slightly exceeds the chromosome-wide threshold of 2.9; however, this minor QTL spans over half of the LG.

Several minor QTLs are located on LG3 and LG4 (Table 2; Fig. 3). At the lower end of LG3, several QTLs were detected. In each parent, the markers analyzed in three out of five environments exceeded the significant LOD threshold; for the most part, this was the genome-wide threshold. Located in this region are the markers Myb10 and R2R3-Myb. The Kruskal–Wallis test showed that marker Myb10 had a significant influence in the environments F-10, G-11 and F-11. However, the K values did not exceed the threshold after Bonferroni correction. The marker R2R3-Myb did not have any influence. Furthermore, three QTL were detected on LG3 of the parent 95/13-39 through MQM mapping. The positions of the QTLs MQM_R-09 and MQM_F-11 are similar; however, these markers only exceeded the chromosome-wide
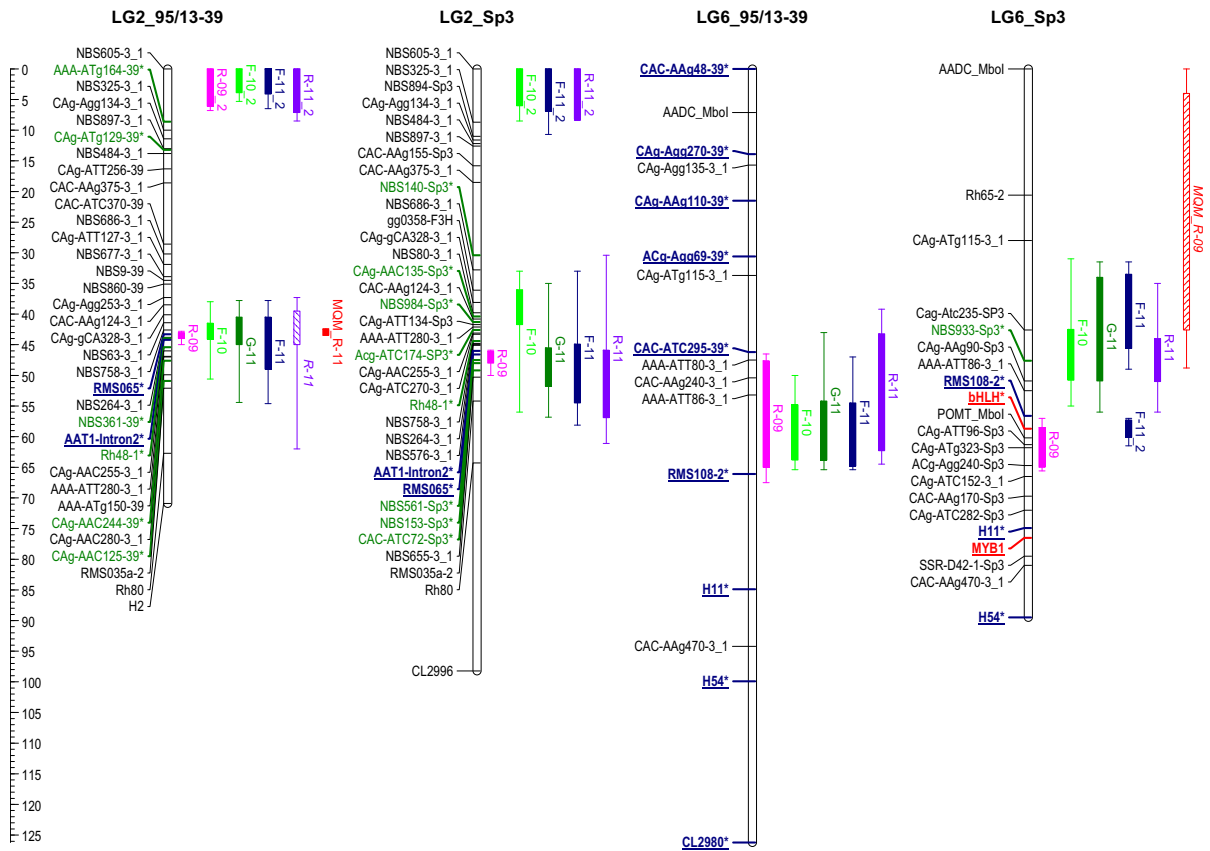
Fig. 2 Linkage maps of LG2 and LG6 of the parental lines 95/13-39 and Sp3 and a summary of the two major QTL regions. The positions of the QTLs are represented by *bars*, with 2-LOD confidence intervals on the *right* of the LG. *Unfilled bars* represent minor QTLs that only exceed the chromosome-wide LOD threshold. The SSCP markers mapped in this study are indicated by *red*, *bold* and *underlined* locus names. Loci that show highly significant *K* values after Bonferroni correction in all environments are indicated by *blue*, *bold* and *underlined* locus names with *asterisks*. Loci with partly significant *K* values are indicated by *green* locus names and with *asterisks*. (Color figure online)

threshold. For the environment F-10, the marker Rh50 showed a significant influence. On LG4, additional major and minor QTLs were detected. These QTLs are partly located in similar regions. For the parent 95/13-39, major QTLs for the environments R-09 and R-11 mapped to the same region. An additional QTL for R-09 is located at the end of the LG4. Another minor QTL was detected through MQM mapping at the top of LG4 in the genotype 95/13-39 for the environment R-09; the marker FLS-1 was within a 1 LOD confidence interval and approximately 60 cM in the environment G-11. The marker GT7-2, developed in this study, is located close to the last QTL. For the parent Sp3, major QTLs were also detected on the top of LG4. These QTLs were close to the marker FLS-2 for the environment R-09 and F-10. One minor QTL for R-11 was in the same region. Additional QTLs were found through MQM mapping. One major QTL for R-11 and one minor QTL for F-10 and G-11 were close to the developed markers GT7-1 and GT7-3 at the center of LG4.

Discussion

The red and pink hues that characterize the flowers of many wild and cultivated roses are due to the accumulation of anthocyanins and flavonoids in rose petals. As the components of the flavonoid biosynthetic pathways are well conserved among higher plant species, flavonoids represent the best-understood group of secondary plant products (Forkmann 2003;
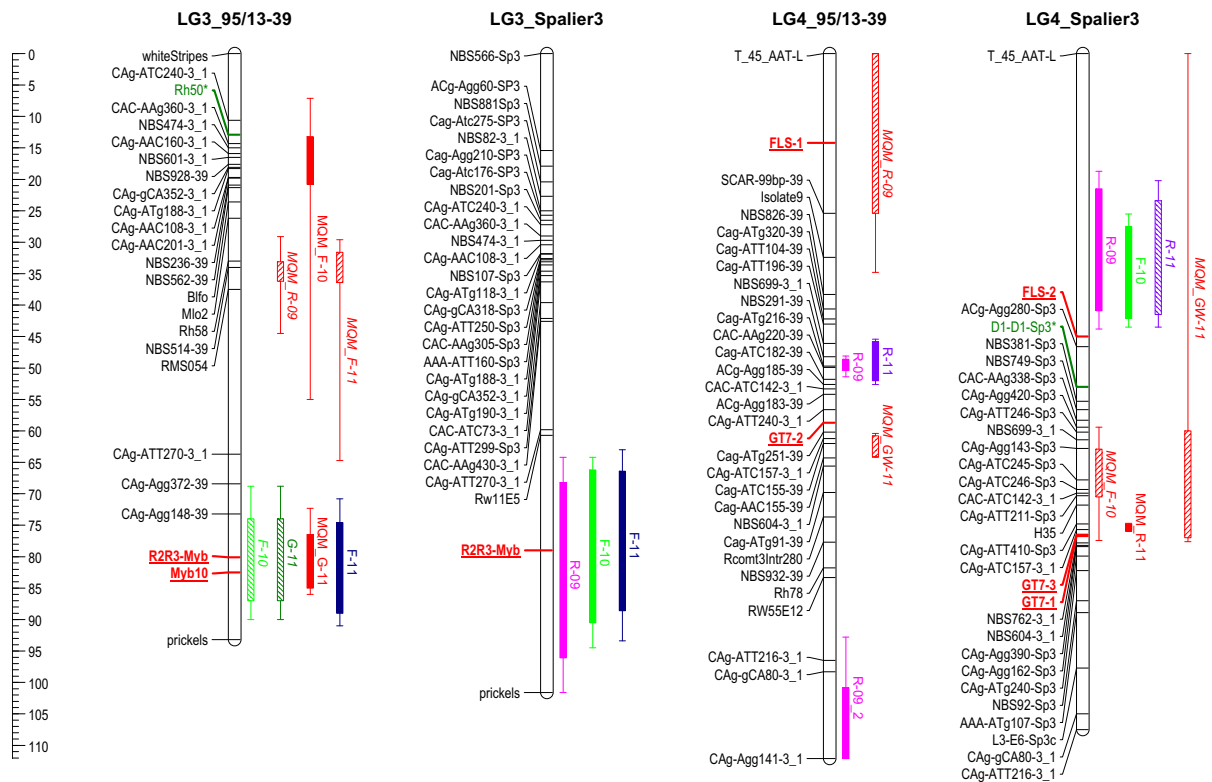
Fig. 3 Linkage maps of LG3 and LG4 of the parental plants 95/13-39 and Sp3 and a summary of several minor loci. The positions of QTLs are represented by *bars* with 2-LOD confidence intervals on the *right* of the LG. *Unfilled bars* represent minor QTLs that only exceed the chromosome-wide LOD threshold. The SSCP markers mapped in this study are indicated by *red*, *bold* and *underlined* locus names. Loci with partly significant *K* values after Bonferroni correction are indicated in *green* and with *asterisks*. (Color figure online)

Xu et al. 2015). For roses, the composition of anthocyanins has been analyzed for a number of genotypes (Marshall et al. 1983; Biolley and Jay 1993; Jay et al. 2003). Several genes involved in the flavonoid biosynthesis pathway have been isolated and studied so far (Forkmann 2003; Ogata et al. 2005). In addition, the inheritance of the color components of rose flowers was analyzed in a number of segregating populations; this revealed single genes as factors explaining the segregation patterns observed among the progeny although quantitative variation was found as well (De Vries et al. 1980; Marshall et al. 1983, Debener 1999). Despite the high degree of conservation in flavonoid biosynthesis, differences in the regulation of parts of this pathway have been identified between species (Grotewold 2006; Tanaka et al. 2008). Therefore, little information is available for most ornamental crops such as roses.

No study thus far describes the loci underlying the quantitative inheritance of anthocyanidin accumulation in roses. Transgenic approaches for the manipulation of rose flower color have been reported (Katsumoto et al. 2007). For political reasons, these techniques are not an option for commercial rose production in Europe and other regions of the world. Furthermore, attempts to manipulate flower color by genetic engineering into commercially acceptable phenotypes have proven to be more difficult than expected (Tanaka et al. 2008). Therefore, knowledge concerning the factors influencing the quantity of anthocyanins in rose petals would be a valuable resource for breeding new rose varieties.

Here, we report the detection of major QTLs that are stable across several environments as well as a number of minor QTLs influencing anthocyanin content in roses. We chose a diploid rose population for the following two reasons: Segregation patterns are much easier to interpret in diploids, and we could build our analysis on previously mapped markers in diploid populations.

Table 2 Summary of QTLs for anthocyanin content detected in the population 97/7 (results for minor QTLs are written in italics)

	LG	Environment	95/13-39			Sp3		
			QTL peak position (cM)	LOD score	2-LOD confidence interval (cM)	QTL peak position (cM)	LOD score	2-LOD confidence interval (cM)
Environmental stable	2	R-09	43.3	8.43	42.8–45.0	46.6	8.35	45.9–50.0
		F-10	43.3	7.1	38.0–50.6	39.8	7.48	33.0–56.0
		G-11	43.3	5.24	37.8–54.4	46.6	5.31	35.0–56.8
		F-11	43.3	6.24	37.8–54.6	46.6	6.26	33.0–58.1
		R-11	*43.3*	*3.99*	*37.3–62.0*	49.2	4.35	30.4–61.1
		MQM_R-11	43.3	7.6	42.4–43.5			
	6	R-09	60.2	15.0	46.5–67.5	62.8	16.55	57.0–65.6
		F-10	59.2	23.4	50.0–65.4	47.6	25.09	31.0–55.0
		G-11	59.2	16.63	43.0–65.4	44.5	16.5	31.5–56.0
		F-11	60.2	20.74	47.0–65.4	38.0	22.53	34.5–49.0
		F-11_2				58.7	19.94	57.0–61.5
		R-11	47.5	14.49	39.2–64.5	47.6	13.8	35.0–56.0
Environmental dependent	2	R-09_2	3.0	4.76	0.0–6.8			
		F-10_2	0.0	11.07	0.0–5.3	0.0	12.43	0.0–8.5
		F-11_2	0.0	6.44	0.0–6.5	0.0	6.75	0.0–10.7
		R-11_2	0.0	6.75	0.0–8.5	0.0	5.59	0.0–8.3
	3	R-09				79.0	4.35	62.4–101.6
		F-10	*82.5*	*3.11*	*68.8–90.0*	72.0	4.93	64.2–94.5
		G-11	*82.5*	*3.21*	*68.8–90.0*			
		F-11	84.5	5.43	70.8–91.0	79.0	6.48	63.0–93.4
		MQM_G-11	84.5	4.35	72.3–86.0			
		MQM_R09	*3.7*	*34.0*	*29.1–44.5*			
		MQM_F-10	4.5	15.9	7.1–55.0			
		MQM_F-11	*3.7*	*34.0*	*29.6–64.7*			
	4	R-09	49.4	6.2	48.1–51.4	31.0	4.35	18.7–43.8
		R-09_2	112.1	4.2	92.8–112.1			
		F-10				35.0	4.69	25.5–43.5
		R-11	50.0	4.2	45.4–52.6	*34.0*	*2.83*	*20.2–43.5*
		MQM_R-09	*23.1*	*3.5*	*0.0–34.8*			
		MQM_F-10				67.8	*3.31*	*59.4–77.4*
		MQM_G-11	*62.1*	*3.74*	*60.4–64.2*	67.8	*3.22*	*0.0–77.6*
		MQM_R-11				75.7	4.07	74.8–76.0
	6	MQM_R-09				*31.0*	*2.98*	*0.0–48.8*

The type of phenotypic variation that we observed indicates a quantitative inheritance of anthocyanin content in rose petals. This is in contrast with previous findings, in which we mapped a major locus for pink flower color as a single dominant gene in another diploid population (population 94/1) with a similar genetic background resulting from hybridizations of garden roses with *R. multiflora* (Debener 1999; Debener and Mattiesch 1999; Spiller et al. 2011). The position of this major locus corresponds to one of the major QTLs on LG2, which was linked to the microsatellite marker RMS065 in this study. In the previous study, it was shown that both parents of population 94/1 carry one allele that confers pink flower color; this resulted in a 3:1

segregation of pink versus white (Debener 1999). If color intensities were measured by anthocyanin extraction, segregation in a 1:2:1 ratio would be detected with homozygotes displaying higher anthocyanin contents. Our observation in the present study that one of the major QTLs is located at the same position indicates that allelic variation in the same gene in combination with other factors might be responsible for the observed variation in color. The monogenic inheritance in the previous study indicates that both parents carried one defective allele of the gene responsible for monogenic segregation.

The marker RMS065, linked to a major QTL on LG2, was located on a genomic contig of 7.2 kb derived from a partial genome sequence of a *R. multiflora* hybrid (Byrne, Debener, Klein unpublished). At a distance of 3.5 kb to the SSR motif, an open reading frame could be identified that mapped to the *Prunus* genome at a distance of 7.1 kb away from a *Prunus* orthologue of a F3H gene that converts naringenin to dihydrokaempferol, an early step in anthocyanidin biosynthesis. Therefore, F3H is a candidate gene for this QTL on LG2.

This locus is one of two QTLs present in both parents and in all environments tested. These QTLs would be interesting target loci for use as selection tools that confer more intensive anthocyanin concentrations in rose petals. The observed quantitative variation in anthocyanin content in the current study may therefore be due to special combinations in alleles in the parents that does not lead to homozygous recessive progeny for key genes in anthocyanin biosynthesis, as described for populations segregating for single genes such as 94/1 (Spiller et al. 2011).

It is known that anthocyanin concentrations in other plant organs such as apple fruits are much more dependent on environmental factors such as light (Ban et al. 2007; Espley et al. 2007; Xie et al. 2011). It is astonishing that, in our experiments, very little variation between the environments could be detected. Although not recorded in detail, it is obvious that the greenhouse and field environments differed strongly in terms of temperature profiles and irradiation intensities. This indicates that the genes underlying the two major QTLs are central regulators of anthocyanin concentration and are not much affected by environmental conditions/influences.

The intensity and hue of anthocyanin coloration of petals are dependent on a number of additional factors, such as petal pH, co-pigments and types of glycosylation (Grotewold 2006; Tanaka et al. 2008). Furthermore, the stability of coloration over the lifetime of flowers from anthesis to senescence varies dramatically between genotypes and was not considered in the present study as petals were sampled at a single, early stage of flower development. The analysis of single genotypes by HPLC and petal extracts treated with hydrochloric acids for the removal of glycosyl residues shows that mainly cyanidin residues are responsible for the pink hues observed in our populations (data not shown).

Apart from F3H, a candidate gene detected by indirect evidence based on the *Prunus* genome sequence, a number of other candidate genes co-localize with QTLs in our study. The second major QTL present in all environments is located in the center of LG6 in a region where a marker derived from a bHLH gene was mapped. Markers generated from two genes for an R2R3 and a Myb10 transcription factor co-segregate with the QTL on the lower part of linkage group 3. Markers for FLS and GST genes map to the area of a QTL on linkage group 4. bHLH, R2R3 and Myb transcription factors are important regulators of anthocyanin biosynthesis. In a number of studies concerning the quantitative variation in anthocyanin concentrations in fruits and flower petals, members of these gene families were shown to influence anthocyanin concentrations in either a qualitative or a quantitative manner (Fournier-Level et al. 2009; Ban et al. 2014; Bushara et al. 2013; Yuan et al. 2013). Therefore, it is likely that allelic variants of orthologues of these transcription factors are responsible for some of the observed variation in anthocyanin content in our population.

Recently we found that PCR-based markers of the same GST, Myb, FLS and bHLH factors were associated with variation in anthocyanins in a set of 96 independent rose varieties analyzed in an association study (Schulz, unpublished). This further supports the role of these genes in the expression of total anthocyanins in roses.

Only some of the known flavonoid biosynthesis genes could be used as segregating markers in our population (supplementary Table 1). For others such as chalcone isomerase or F3H, no useful markers could be generated for PCR-based assays. Extending the marker analysis by GBS- or by NGS-based amplicon sequencing of additional sequences from

genes of the anthocyanin biosynthesis pathway in the segregating population could improve the analysis of candidate genes.

With the data presented here, we show that major loci influencing the total anthocyanin content in rose petals are largely independent of the environment. Therefore, these loci serve as interesting molecular markers for marker-assisted selection. As anthocyanins are easy to select for in conventional breeding programs, marker-assisted selection in segregating progeny might not be of primary importance. However, the dose of markers of superior alleles might provide useful information for the selection of appropriate parents; such selection will affect the fraction of tetraploid progeny expressing the trait, for which predictions based on the phenotypes of the parents alone are not conclusive.

References

Ban Y, Honda C, Hatsuyama Y, Igarashi M, Bessho H, Moriguchi T (2007) Isolation and functional analysis of a MYB transcription factor gene that is a key regulator for the development of red coloration in apple skin. Plant Cell Physiol 48:958–970

Ban Y, Mitani N, Hayashi T, Sato A, Azuma A, Kono A, Kobayashi S (2014) Exploring quantitative trait loci for anthocyanin content in interspecific hybrid grape (*Vitis labruscana* × *Vitis vinifera*). Euphytica 198:101–114

Biolley JP, Jay M (1993) Anthocyanins in modern roses-chemical and colorimetric features in relation to the color range. J Exp Bot 44:1725–1734

Biolley JP, Jay M, Viricel MR (1994) Flavonoid diversity and metabolism in 10 *Rosa* × *hybrida* cultivars. Phytochemistry 35:413–419

Bushara JM, Krieger C, Deng D, Stephens MJ, Allan AC, Storey R, Symonds VV, Stevenson D, McGhie T, Chagne D et al (2013) QTL involved in the modification of cyanidin compounds in black and red raspberry fruit. Theor Appl Genet 126:847–865

Byrne DH (2009) Rose structural genomics. In: Folta KM, Gardiner SE (eds) Genetics and genomics of rosacea, Plant genetics and genomics: crops and models 6. Springer Science and Business Media, New York, pp 353–379

Clausen J, Keck DD, Hiesey WM (1940) Experimental studies on the nature of species I. Effect of varied environments on Western North American plants. Carnegie Institute of Washington Publication No. 520, 452 pp

De Vries DP, Van Keulen HA, Debruyn JW (1974) Breeding research on rose pigments 1. Occurrence of flavonoids and carotenoids in rose petals. Euphytica 23:447–457

De Vries DP, Garretsen F, Dubois LAM, Van Keulen HA (1980) Breeding research on rose pigments II. Combining ability analyses of variance of four flavonoids in F1 populations. Euphytica 29:115–120

Debener T (1999) Genetic analysis of horticulturally important morphological and physiological characters in diploid roses. Gartenbauwissenschaft 64:14–20

Debener T, Hibrand-Saint Oyant L (2009) Genetic engineering and tissue culture of roses. In: Folta KM, Gardiner SE (eds) Genetics and genomics of rosacea, Plant genetics and genomics: crops and models 6. Springer Science and Business Media, New York, pp 393–409

Debener T, Mattiesch L (1999) Construction of a genetic linkage map for roses using RAPD and AFLP markers. Theor Appl Genet 99:891–899

Espley RV, Hellens RP, Putterill J, Stevenson DE, Kutty-Amma S, Allan AC (2007) Red coloration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. Plant J 49:414–427

Forkmann G (2003) Flavonoid molecular biology. In: Roberts AV, Debener T, Gudin S (eds) Encyclopedia of rose science. Elsevier, Amsterdam, pp 256–263

Fournier-Level A, Le Cunff L, Gomez C, Doligez A, Ageorges A, Roux C, Bertrand Y, Souquet J, 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 183:1127–1139

Grotewold E (2006) The genetics and biochemistry of floral pigments. Annu Rev Plant Biol 57:761–780

Gudin S (2000) Rose: genetics and breeding. Plant Breed 17:159–189

Jansen RC, Stam P (1994) High-resolution of quantitative traits into multiple loci via interval mapping. Genetics 136:1447–1455

Jay M, Biolley JP, Fiasson JL, Fiasson K, Gonnet JF, Grossi C, Raymond O, Viricel MR (2003) Anthocyanins and other flavonoid pigments. In: Roberts AV, Debener T, Gudin S (eds) Encyclopedia of rose science. Elsevier, Amsterdam, pp 248–256

Katsumoto Y, Fukuchi-Mizutani M, Fukui Y, Brugliera F, Holton TA, Karan M, Nakamura N, Yonekura-Sakakibara K, Togami J, Pigeaire A et al (2007) Engineering of the rose flavonoid biosynthetic pathway successfully generated blue-hued flowers accumulating delphinidin. Plant Cell Physiol 48:1589–1600

Lander ES, Botstein D (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199

Linde M, Hattendorf A, Kaufmann H, Debener T (2006) Powdery mildew resistance in roses: QTL mapping in different environments using selective genotyping. Theor Appl Genet 113:1081–1092

Marshall HH, Campbell CG, Collicutt LM (1983) Breeding for anthocyanin colours in *Rosa*. Euphytica 32:205–216

Mikanagi Y, Yokoi M, Ueda Y, Saito N (1995) Flower flavonol and anthocyanin distribution in subgenus *Rosa*. Biochem Syst Ecol 23:183–200

Ogata J, Kanno Y, Itoh Y, Tsugawa H, Suzuki M (2005) Plant biochemistry: anthocyanin biosynthesis in roses. Nature 435:757–758

Orita M, Suzuki Y, Sekiya T, Hayashi K (1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain-reaction. Genomics 5:874–879

Picone JM, Clery RA, Watanabe N, MacTavish HS, Turnbull CGN (2004) Rhythmic emission of floral volatiles from *Rosa damascena semperflorens* cv. 'Quatre Saisons'. Planta 219:468–478

R Development Core Team (2015) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Schistl FP, Johnson SD (2003) Pollinator-mediated evolution of floral signals. Trends Ecol Evol 28:307–315

Schmitzer V, Stampar F (2010) Changes in anthocyanin and selected phenolics in "DORcrisett" rose flowers due to substrate pH and foliar application of sucrose. Acta Hortic 870:89–93

Schmitzer V, Veberic R, Osterc G, Stampar F (2010) Color and phenolic content changes during flower development in groundcover rose. J Am Soc Hortic Sci 135:195–202

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol 18:233–234

Spiller M, Berger RG, Debener T (2010) Genetic dissection of scent metabolic profiles in diploid rose populations. Theor Appl Genet 120:1461–1471

Spiller M, Linde M, Hibrand-Saint Oyant L, Tsai C-J, Byrne DH, Smulders MJM et al (2011) Towards a unified genetic map for diploid roses. Theor Appl Genet 122:489–500

Stam P (1993) Construction of integrated genetic-linkage maps by means of a new computer package—Joinmap. Plant J 3:739–744

Tanaka Y, Sasaki N, Ohmiya A (2008) Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. Plant J 54:733–749

van Ooijen JW (1992) Accuracy of mapping quantitative trait loci in autogamous species. Theor Appl Genet 84:803–811

van Ooijen JW (2004) MapQTL® 6, Software for the mapping of quantitative trait loci in experimental populations. Kyazma B.V., Wageningen

van Ooijen JW (2006) JoinMap 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma B.V., Wageningen

Xie RJ, Zheng L, He SL, Zheng YQ, Yi SL, Deng L (2011) Anthocyanin biosynthesis in fruit tree crops: genes and their regulation. Afr J Biotechnol 10:19890–19897

Xu W, Dubos C, Pepiniec L (2015) Transcriptional control of flavonoid biosynthesis by Myb–bLHL–WDR complexes. Trends Plant Sci 20:176–185

Yuan YW, Sagawa JM, Young RC, Christensen BJ, Bradshaw HD (2013) Genetic dissection of a major anthocyanin QTL contributing to pollinator-mediated reproductive isolation between sister species of *Mimulus*. Genetics 194:255–263