



Genome-Wide Association Analysis of the Anthocyanin and Carotenoid Contents of Rose Petals

Dietmar F. Schulz¹, Rena T. Schott^{1†}, Roeland E. Voorrips², Marinus J. M. Smulders², Marcus Linde¹ and Thomas Debener^{1*}

¹ Abteilung Molekulare Pflanzenzüchtung, Institute for Plant Genetics, Leibniz University Hannover, Hannover, Germany,

² Wageningen University and Research Plant Breeding, Wageningen University and Research Centre, Wageningen, Netherlands

OPEN ACCESS

Edited by:

Soren K. Rasmussen,
University of Copenhagen, Denmark

Reviewed by:

Liezhao Liu,
Southwest University, China
Shuhua Yang,
Chinese Academy of Agricultural
Sciences, China

*Correspondence:

Thomas Debener
debener@genetik.uni-hannover.de

† Present Address:

Rena T. Schott,
Staatliches Museum für Naturkunde
Stuttgart, Stuttgart, Germany

Specialty section:

This article was submitted to
Crop Science and Horticulture,
a section of the journal
Frontiers in Plant Science

Received: 27 June 2016

Accepted: 15 November 2016

Published: 06 December 2016

Citation:

Schulz DF, Schott RT, Voorrips RE,
Smulders MJM, Linde M and
Debener T (2016) Genome-Wide
Association Analysis of the
Anthocyanin and Carotenoid Contents
of Rose Petals.
Front. Plant Sci. 7:1798.
doi: 10.3389/fpls.2016.01798

Petal color is one of the key characteristics determining the attractiveness and therefore the commercial value of an ornamental crop. Here, we present the first genome-wide association study for the important ornamental crop rose, focusing on the anthocyanin and carotenoid contents in petals of 96 diverse tetraploid garden rose genotypes. Cultivated roses display a vast phenotypic and genetic diversity and are therefore ideal targets for association genetics. For marker analysis, we used a recently designed Axiom SNP chip comprising 68,000 SNPs with additionally 281 SSRs, 400 AFLPs and 246 markers from candidate genes. An analysis of the structure of the rose population revealed three subpopulations with most of the genetic variation between individual genotypes rather than between clusters and with a high average proportion of heterozygous loci. The mapping of markers significantly associated with anthocyanin and carotenoid content to the related *Fragaria* and *Prunus* genomes revealed clusters of associated markers indicating five genomic regions associated with the total anthocyanin content and two large clusters associated with the carotenoid content. Among the marker clusters associated with the phenotypes, we found several candidate genes with known functions in either the anthocyanin or the carotenoid biosynthesis pathways. Among others, we identified a glutathione-S-transferase, 4CL, an auxin response factor and F3'H as candidate genes affecting anthocyanin concentration, and CCD4 and Zeaxanthine epoxidase as candidates affecting the concentration of carotenoids. These markers are starting points for future validation experiments in independent populations as well as for functional genomic studies to identify the causal factors for the observed color phenotypes. Furthermore, validated markers may be interesting tools for marker-assisted selection in commercial breeding programmes in that they provide the tools to identify superior parental combinations that combine several associated markers in higher dosages.

Keywords: petal color, anthocyanin, carotenoid, genome wide association study, tetraploid roses

INTRODUCTION

Rose is one of the most economically important ornamental crops and is sold as cut flowers, pot roses and garden roses. The genus *Rosa* comprises a vast amount of genetic resources represented by more than 100 wild species as well as more than 30,000 mostly tetraploid varieties bred for ornamental purposes (Gudin, 2000; Wissemann, 2003). Cultivated tetraploid rose genomes are complex mixtures of at least 10 species that have been used in ornamental rose breeding for more than two centuries (Gudin, 2000; Zhang et al., 2013). As a result, rose is highly diverse in many morphological and physiological characteristics. Despite its commercial importance as an ornamental plant, genomic resources for rose research and breeding remain scarce, and to date, no genome sequence is available. At the diploid level, genetic maps have been constructed, and a number of monogenic and quantitative traits have been localized on these maps (Debener and Linde, 2009; Spiller et al., 2011). However, as most diploid populations have derived from a few diploid genotypes, genetic variability is low for most horticultural traits. Therefore, these traits can only be analyzed at the tetraploid level. In tetraploid varieties, several monogenic traits have been analyzed, but only few QTL have been described, mostly by analyses of biparental populations (Debener and Linde, 2009; Spiller et al., 2011).

The esthetic features of the rose flower are of central importance for the ornamental quality of rose cultivars; therefore, commercial breeding pays special attention to floral characteristics. Flower traits, e.g., the number and color of petals, were among the first traits investigated in genetic studies (De Vries and Dubois, 1978; Debener, 1999, 2003).

The anthocyanin concentration in cells of rose petals is a major determinant of red and pink color variants, although the final hues are influenced by several other factors as e.g., pH, copigments, metal ions, types of glycosylation, etc. (Jay et al., 2003; Grotewold, 2006; Tanaka et al., 2008). Although single loci influencing color variation have been identified, other researchers have described a quantitative inheritance of the anthocyanin content (Cardoso et al., 2012; Cericola et al., 2014; Henz et al., 2015).

The levels of carotenoids, which produce yellow colors, are influenced by biotic and abiotic factors, including the developmental stage, the environment and stress (Eugster and Märki-Fischer, 1991; Deli et al., 1998; Kishimoto et al., 2004). In fully opened flowers of *Ipomoea*, the chromoplast-type carotenoids are β -cryptoxanthin, zeaxanthin, and β -carotene, whereas lutein, violaxanthin and β -carotene are predominant in the early stage of petal development, and the same compounds were found in the leaves (Yamamizo et al., 2010).

QTL Mapping and GWAS

All of the QTLs studied to date in rose have been mapped in biparental populations (Crespel et al., 2002; Linde et al., 2006; Spiller et al., 2011; Moghaddam et al., 2012; Roman et al., 2015) using AFLP and microsatellite markers. Tetraploid populations derived from crosses between ornamental varieties display complex patterns of inheritance that complicate not only

genetic analysis but also map construction and require many more markers (Bourke et al., 2015).

Association studies offer two main advantages over QTL studies based on biparental populations: a larger number of alleles per locus and a higher resolution of trait-marker associations due to a higher rate of recombination. In association genetics, genotyping can be restricted to candidate genes likely involved in the expression of the traits under study or markers covering the whole genome in genome-wide association studies (GWAS). Few studies cover polyploids, and very few GWAS have been performed on highly heterozygous polyploids, such as potato (D'hoop et al., 2010; Lindqvist-Kreuzer et al., 2014), switchgrass (Lu et al., 2013) and cotton (Abdurakhmonov et al., 2009). Rose is an interesting ornamental crop for association studies because its cultivars are extremely polymorphic, and many traits can be studied simultaneously in populations of moderate size.

Recently, an analysis of a large collection of rose ESTs and the development of an Axiom SNP array was described (Koning-Boucoiran et al., 2015). These resources are a significant extension of the genomic resources available for roses because they now permit the highly reproducible genotyping of rose genomes with approximately 68,000 SNP markers each represented by two probes. Hence, sufficient numbers of markers are now available for GWAS in tetraploid rose.

Aims of the Present Study

The aim of the present study was to exploit the enormous biodiversity of cultivated roses in flower-related traits for an analysis of the underlying genetic factors, focusing on the contents of anthocyanins and carotenoids, which are the main components of the rose petal color. This was accomplished using a combination of association genetics methods and markers on the SNP array and additional markers derived from candidate genes, SSRs and AFLPs. In addition, we tried to gain information about the variability of the genetic diversity and heterozygosity within our association panel.

MATERIALS AND METHODS

Plant Material

An association panel of 96 rose cultivars with code numbers from 1 to 141 (87 tetraploid, 8 triploid, and 1 diploid) was used for the present study. Most of the cultivars were commercially available or provided by German rose-breeding companies (**Table S1**). Based on known pedigrees, we attempted to minimize genotypic relatedness, which can result in spurious associations, while capturing the vast diversity of phenotypic traits, including different flower colors, plant architectures, etc. Clones of each cultivar, grafted on *R. corymbifera* "Laxa," were planted in three randomized blocks in a field at Hannover-Herrenhausen (Germany) in the spring of 2012. A second collection of cultivars was maintained under semi-controlled conditions as potted plants in three randomized blocks in a greenhouse (Federal Plant Variety Office, Hannover). The plants were initially cultivated in 3-l pots and then transferred to 7-l pots with the fertilized substrate Einheitserde T (Einheitserdewerke

Gebr. Patzer, Sinntal-Altengronau, D) under natural light with a day and night temperature of $22^{\circ} \pm 5^{\circ}\text{C}$.

Anthocyanin Content of Petals

Flowers were always sampled from 8 to 12 a.m. Opened buds at flower development stage 3 (Picone et al., 2004) were selected from each genotype and kept on ice until sample preparation on the same day. The anthocyanin content of petals was estimated according to Henz et al. (2015) with minor modifications. Three replicates (each 50 mg in fresh weight) from petals of each clone (3 biological replicates) were placed in 2 ml test tubes and extracted in 1 ml of methanol/HCl (99:1 v/v) (Figure S7). Following an overnight incubation (16 h) in the dark at 18°C , the total anthocyanin content in the solvent was determined based on the absorbance at 525 nm using a UV-Vis-Photometer UV mcSAFAS (Deelux Labortechnik GmbH, Germany). If necessary ($E_{525\text{nm}} > 1.0$), the anthocyanin extracts were diluted with the extraction solvent. Each clone was measured three times, and the overall mean was calculated for each cultivar. The anthocyanin content was recorded and evaluated in two environments: (i) in the field at Herrenhausen and (ii) in the greenhouse at the Federal Plant Variety Office, Hannover. The absorbance values were not used to calculate the levels of compounds, as this measurement involves variable mixtures of anthocyanidins.

Carotenoid Content of Petals

The content of carotenoids was evaluated from all cultivars cultivated in the greenhouse at the Federal Plant Variety Office Hannover and from 20 of the cultivars in the field at Herrenhausen. The accumulation of total carotenoids in rose petals was estimated according to de Vries et al. (1974) with modification. Petals (50 mg each) were extracted with 1 ml of a mixture of petroleum ether:acetone (3:2 v/v) for 4 h at 18°C in the dark, and the carotenoids in the samples were measured spectrophotometrically at a wavelength of 442 nm. The extracts showed three characteristic absorption maxima for carotenoids at 419, 442, and 471 nm (Figure S1). These maxima suggest Violaxanthin (420, 443, and 471 nm) or Neoxanthin (420, 442, and 473 nm) as components of the extract in addition to other possible compounds, respectively (Wellburn, 1994; Tinoi et al., 2006). The overlap in absorption at 442 nm is the reason that we did not determine the levels of a particular compound.

DNA Extraction

DNA was extracted from young rose leaves as described by Klie et al. (2013). The quality of DNA was checked on agarose gels, and quantification was performed using a Nanodrop 2000c spectrophotometer (PeQLab Biotechnologie GmbH, Erlangen, Germany).

Microsatellites (SSRs)

PCR was performed in 15 μl of Williams buffer (Williams et al., 1990) containing 0.2 mM dNTP, 0.5 μM forward and reverse primers, and 1 U of DCS Taq DNA polymerase (Enzymatics, Beverly, USA). PCR conditions were as follows: initial denaturation for 5 min. at 94°C , 28 cycles of 45 s at 94°C , 45 s at 50, 55 or 60°C (Table S2), 60 s at 72°C . SSR marker bands

were visually inspected and dominantly scored, and the data were transferred to a 0/1 matrix.

AFLP

AFLP markers were generated according to the protocol of Klie et al. (2013) with 250 ng of genomic DNA. We tested 21 *HindIII* and *MseI* primer pair combinations in the end reaction. Bands were scored dominantly and recorded in a 0/1 matrix for absence/presence of marker fragments.

SNPs

SNPs were analyzed using the Axiom WagRhSNP array, which contains 89,893 SNPs derived from cut roses and from garden roses (Koning-Boucoiran et al., 2015). The hybridisation intensities were interpreted as tetraploid SNP dosage scores (AAAA, AAAB, AABB, AB BB, and BBBB) using fitTetra (Voorrips et al., 2011) and were used to calculate the statistics required for the association study. SNP markers that were polymorphic and scorable were used for GWAS after filtering for minor allele frequency (MAF > 0.1) and missing data (<10%). Heterozygosity was calculated as the percentage of heterozygous loci (ABBB, AABBB, and ABBBB) compared to the total number of loci.

Population Structure

The population structure was modeled in STRUCTURE 2.3.4 (Pritchard et al., 2000; Falush et al., 2007) with a burn-in of 10,000 cycles and varied in the number of following Markov Chain Monte Carlo (MCMC) iterations (50,000 and 100,000) and the number of AFLP (400) and microsatellite and candidate gene markers (175 and 527). The SNP markers were not used for determining the population structure. We used the implemented admixture model with correlated allele frequencies and an initial alpha of 1 and accomplished three independent runs based on 10 repeats of the simulations for each K, from $K = 1$ to 10. Then, the most likely number of subpopulations was estimated based on the method of Evanno et al. (2005) using the InP (D) value (estimated likelihood) with the software StructureHarvester (<http://taylor0.biology.ucla.edu/structureHarvester/>) (Earl and von Holdt, 2012).

Furthermore, a PCoA (Principal Coordinate Analysis) was performed using DARwin 5.0.158 (Perrier and Jacquemoud-Collet, 2006) with the same subset of 927 AFLP, SSR and candidate gene markers.

Genetic Diversity Analysis

The genetic distance among the collection of rose genotypes was calculated with DARwin using a subset of 16,040 SNPs at the tetraploid dosage state from the filtered SNP set (MAF > 0.1 and missing data < 10%). An unweighted neighbor-joining (Saitou and Nei, 1987) dendrogram was constructed based on a distance dissimilarity matrix using a bootstrap analysis with 100 repetitions.

Kinship Matrix Calculation

A kinship matrix was used to establish and describe the relationship between the genotypes. Pairwise kinship coefficients were estimated using the programme SPAGeDi (Hardy and

Vekemans, 2002) based on the method of Hardy (2003) using 10,000 random SNP markers at the tetraploid dosage state from the 16,040 SNPs above. The diagonal of the matrix from SPAGeDi was set to two, and negative values were set to zero (Yu et al., 2006).

Trait-Marker Associations

Trait-marker association analysis was performed using the mixed linear model (MLM, $K + Q$) in TASSEL 3.0 (Bradbury et al., 2007). SNP markers with a minor allele frequency of less than 0.1 and with more than 10% missing data were excluded from further analysis. In TASSEL 3.0, marker allele configurations can only be used in a diploid configuration (e.g., AA, AB, or BB). Therefore, bi-allelic SNPs of the tetraploid rose cultivars were coded as diploids. For this, all possible heterozygous genotypes (AAAB, AABB, and ABBB) were coded as AB, similar to how Li et al. (2014) analyzed diploid and tetraploid Alfalfa genotypes and how Lindqvist-Kreuzer et al. (2014) analyzed potato genotypes. Associations were estimated including the Q-matrix for population effects based on the output from STRUCTURE 2.3.4 (based on AFLP, microsatellite and candidate gene data) and the kinship matrix (K) calculated with SPAGeDi (based on SNP data). Bonferroni adjustments of the p -values were made to correct for the number of independent tests and to establish a threshold (Johnson et al., 2010). For this, a total of 19,074 independent tests (number of contigs) were assumed because a precise estimation of the real number of independent tests could not be made due to unknown linkages between most of the markers. An SNP marker was considered associated if its $-\log_{10} p$ -value was greater than 5.58.

Statistical Analysis

The nonparametric Kruskal–Wallis rank-sum test was used to identify significant differences in the mean SNP effect in groups of cultivars. Spearman rank correlation was used to test the association between the anthocyanin content in petals from greenhouse and field-grown roses. Significant differences in the means of heterozygosity in different growth types of roses were calculated using the Wilcoxon signed rank test. The data were tested for normal distribution using the Shapiro-Wilk test ($\alpha = 0.05$). The data that were not normally distributed were transformed using log- and Box-Cox transformation (Wessa, 2016). The statistical calculations were performed in Excel 2007, MYSTAT 12 (Systat Software, Inc.) and QtiPlot 0.9.9 (Vasilief, 2015).

RESULTS

Population Structure

The population structure was analyzed based on three independent runs of STRUCTURE that varied in the number of MCMC iterations (50,000 and 100,000) and the number of AFLP, microsatellite and candidate gene markers (575 and 927) for $K = 1$ to $K = 10$. The optimum number of K can be identified according to the maximum value of $\text{LnP}(D)$ (Pritchard et al., 2000). In our data, the likelihood distribution increased slightly, leveled off and then decreased with a clear plateau from $K = 3$ to

$K = 5$ (Figure S2A). Using the method of Evanno et al. (2005) in two independent runs with a total of 575 AFLP and SSR markers (burn-in 10,000; MCMC 50,000, and 100,000) and one run with 627 markers (burn-in 10,000; MCMC 50,000), the maximum for ΔK was estimated at $K = 3$ (Figure S2B). The result was confirmed by additional independent runs (data not shown).

The structure of the population at $K = 3$ is visualized in Figure 1A. Below it are depicted the concordant results from the cluster analysis of the SNP data (Figure 1B). When using a threshold of 0.7 to assign individuals to a subpopulation or to classify them as a mix or as hybrid individuals (as (D'hoop et al., 2010) used in highly heterozygous tetraploid potato), subpopulation I, the largest group, consisted of 44 cultivars, which clustered according to their type or habit, particularly hybrid tea and floribunda roses. Subpopulation II contained 17 recently bred (1985–2011) cultivars, except for New Dawn (1930), which all have a groundcover habit. Subpopulation III, the smallest group, comprised only five cultivars belonging to the old garden type of roses: Damask ('Rose de Resht', before 1900), Alba ('Small Maidens Blush', 1797), Bourbon ('Louise Odier', 1851), Hybrid Perpetual ('Mrs. John Laing', 1885) and Portland ('Mme Knorr', 1858 und 'Mme Boll', 1858) roses. The positioning of the cultivars was supported by high bootstrap values, except for some in the first subpopulation. However, 30 cultivars could not be assigned to any of the three subpopulations using the threshold of 0.7 for the classification. The results of a principal coordinate analysis (PCoA; Figure S3) based on genetic distances agreed with this division in three subpopulations. The hybrid cultivars that shared part of subpopulations I and II can be observed as an intersection (black dots) between these subpopulations.

Genetic Diversity Analysis

The kinship estimates based on SNPs (Figure S4) indicated no familial relationship between most of the rose genotypes. Approximately 59% of the pairwise kinship coefficients had values near zero (<0.005). Higher kinship estimates (0.10–0.20) were found between climber and ground cover roses. The highest values (0.26–0.39) were found within the group of old garden roses (population III).

The heterozygosity was determined based on the SNP data without considering the dosage of the markers (i.e., AAAB, AABB, and ABBB are all classified as heterozygote). When defined this way, the percentage of heterozygous loci is identical to the percentage polymorphic loci. On average, varieties displayed 55.2% of heterozygous loci ranging from 27% for variety No. 105 and 66.9% for variety No. 2. No correlation between heterozygosity and the age of the variety was observed (Figure S5).

On the other hand, there were significant differences between ground cover, climber, bedding roses, Hybrid teas and shrub roses, both in the average level of heterozygosity and in the variation in heterozygosity within groups (Figure 2). Groundcover roses had the lowest heterozygosity (44.4% heterozygous loci). Hybrid teas were significantly higher in heterozygosity (60.1%) compared to climbers and ground cover roses.

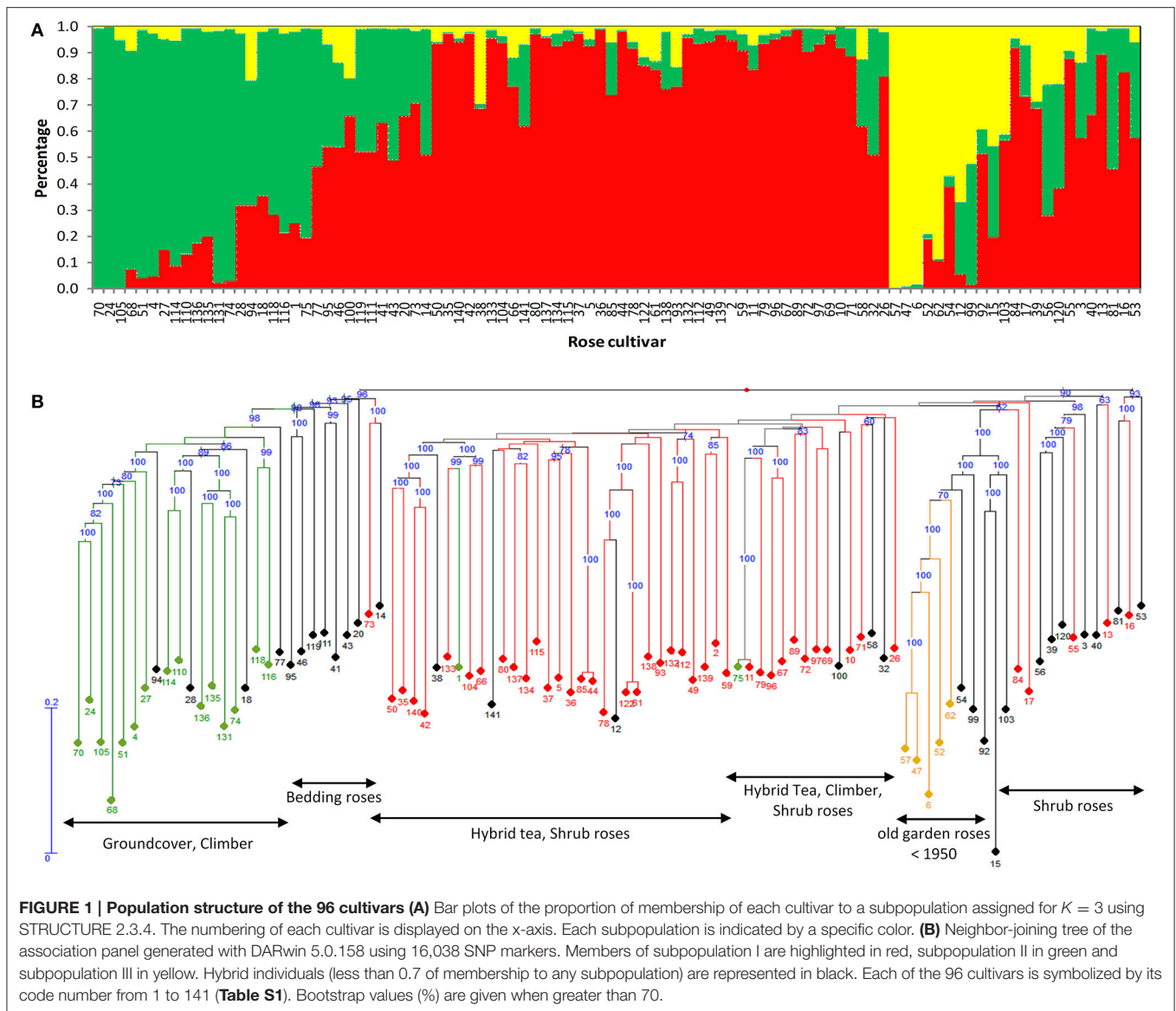


FIGURE 1 | Population structure of the 96 cultivars (A) Bar plots of the proportion of membership of each cultivar to a subpopulation assigned for $K = 3$ using STRUCTURE 2.3.4. The numbering of each cultivar is displayed on the x-axis. Each subpopulation is indicated by a specific color. **(B)** Neighbor-joining tree of the association panel generated with DARwin 5.0.158 using 16,038 SNP markers. Members of subpopulation I are highlighted in red, subpopulation II in green and subpopulation III in yellow. Hybrid individuals (less than 0.7 of membership to any subpopulation) are represented in black. Each of the 96 cultivars is symbolized by its code number from 1 to 141 (Table S1). Bootstrap values (%) are given when greater than 70.

Phenotypic Characterization of morphological Traits

Total Anthocyanins

In many cultivars, the anthocyanin content in the petals was low or not detectable, which was not unexpected because 23 cultivars had a white and yellow flower color. The measured anthocyanins were in the range from $E_{525\text{ nm}} = 0.35$ to $E_{525\text{ nm}} = 33.22$ in the greenhouse and from $E_{525\text{ nm}} = 0.33$ to $E_{525\text{ nm}} = 38.39$ in the field. The distribution was skewed to the left (Figure S6 and Table 1) and very similar for both environments ($r = 0.942$, Figure 3).

Carotenoids

Carotenoids in the rose petal extracts were measured by their characteristic absorbance at 442 nm (Figure S1) from all cultivars cultivated in the greenhouse at the Federal Plant Variety Office Hannover and from 20 of the cultivars in the field at

Herrenhausen. Because the Spearman rank correlation between the measured values was very high (Spearman's $\rho = 0.939$) the carotenoid contents of the additional 76 cultivars grown in the field were not estimated to avoid redundant data (Figure S8). In white flowers and in many red flowers, the yellowish to orange pigments were not present or only found as minor pigments in the petals. The maximum amount of carotenoids was detected in the yellow flowering cultivar "China Girl" ($E_{442\text{ nm}} = 0.8854$). In roses classified as orange, a balanced occurrence of anthocyanins and carotenoids was always measured. However, there was no overall correlation between the anthocyanin and carotenoid content in the rose petals ($r = -0.1803$, $p = 0.0836$).

Marker Trait Associations

Anthocyanins

Marker-trait associations were calculated in Tassel 3.0 with 39,831 SNPs based on their diploid allelic state (A:A, A:B,

and B:B). The tetraploid dosage of the SNP was not used for calculation because TASSEL 3.0 can only handle haploid and diploid genotypic data (user manual for Tassel 3.0, Buckler Lab, Cornell University, 2011).

The anthocyanin content in greenhouse-grown roses was significantly associated with 17 SNP markers, five of which were also associated with the anthocyanin content from field-grown roses (Table 2). These SNPs were Rh12GR_283_1910Q (in the Auxin response factor 8 gene), RhK5_1258_2078P (3 β -OH-steroid-dehydrogenase/decarboxylase isoform 2), RhK5_7371_202Q (Glutathione S-transferase), Rh12GR_20064_1031P and RhMCRND_20203_163Q (both are in the Medium-chain-fatty-acid-CoA ligase). We estimated the effects of the SNPs on the anthocyanin content in greenhouse-grown roses from 3.985 to 7.589 (Table 2). Under both conditions, the largest effect was found for the marker RhK5_1258_2078P. In Figure 4, two boxplots of anthocyanin content showed the direct effects of the markers. For the SNP in the auxin response factor 8 gene, the mean for the heterozygous genotypes A:B was 8.63 (E_{525nm}) and was significantly higher ($p = 5.03E-8$) than 1.93 (E_{525nm}) for the homozygous B:B genotypes. For the 3- β -OH-steroid-dehydrogenase SNP, the difference between the mean of the two groups (A:A = 10.26; A:B = 1.91) was also significant ($p = 8.56E-7$).

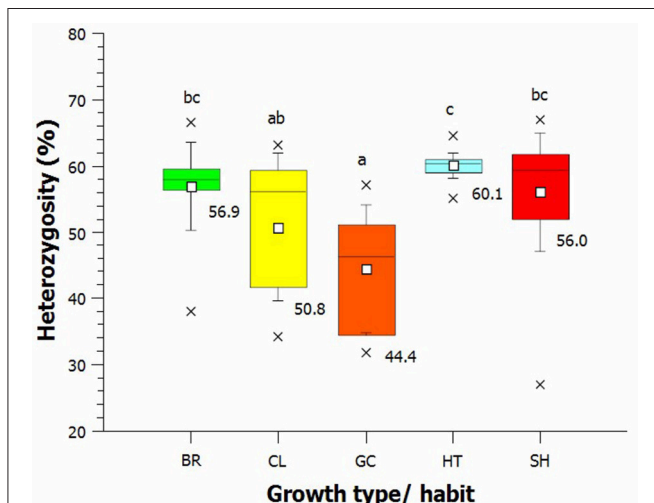


FIGURE 2 | Average heterozygosity of SNPs in different rose growth types (BR, bedding roses; CL, climber; GC, ground cover; HT, Hybrid tea; and SH, shrub roses). Small white square = mean; continuous line = median; asterisk = minimum, maximum; box = 1st and 3rd quartiles; and whisker = standard deviation.

Because the genome sequence of *Rosa sp.* is not complete, the contigs of 133 SNPs with hits just below the Bonferroni threshold plus the 17 significant SNPs were blasted against the closely related genomes of *Fragaria vesca* and *Prunus persica* and mapped on these genomes (Figures 6A,B, Table S6). The assumption is that both genomes display sufficient microsynteny to the rose genome. This assumption is supported by the fact that these rose SNPs clustered in distinct regions of the *Fragaria* genome, particularly in linkage groups Fvb1, Fvb2, Fvb4, Fvb5, and Fvb6, and in the partly homologous linkage groups Pp01, Pp03, Pp05, Pp06, and Pp08 of *Prunus persica* (https://www.rosaceae.org/gb/gbrowse_syn/peach_apple_strawberry/). A blast of the contigs located three of the SNP markers in the coding region of anthocyanin biosynthesis genes, 4-coumarate-ligase (4_CL), flavonoid 3'-hydroxylase (F3'H) and glutathione-S-transferase (GST). The positions of these genes of the anthocyanin biosynthesis pathway and of further transcription factors are shown in the genome plots as green dots (Figures 6A,B). The putative transcription factors that are associated with anthocyanin accumulation include a ubiquitin-like protein SMT3 (SUMO1), WRKY transcription factor 17 and UTP4/Cirhin, a WD40 repeat protein (Freed et al., 2012).

Carotenoids

Because as many as 351 SNPs were significantly associated with the accumulation of carotenoids in rose petals and surpassed the

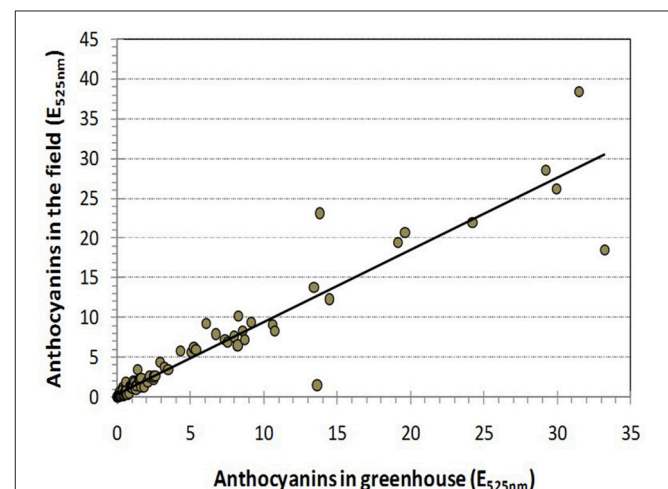


FIGURE 3 | Pearson's correlation between the total amount of anthocyanin that accumulated in rose cultivars grown in the field and in the greenhouse ($r = 0.942$).

TABLE 1 | Descriptive statistics of the investigated traits.

Trait	Mean	Median	Stdev	Stderr	Variance	Max	Min	Skewness	N
Anthocyanins (greenhouse)	4.78	1.19	7.639	0.784	58.35	33.22	0.035	2.261	95
Anthocyanins (field)	4.56	1.41	7.198	0.750	51.81	38.39	0.033	2.506	92
Carotenoids (greenhouse)	0.14	0.07	0.175	0.018	0.03	0.88	0.012	1.957	94

Stderr, standard error; Stdev, standard deviation; Max, maximum; min, minimum; N, number.

TABLE 2 | Significant SNPs for anthocyanin content in rose petals from greenhouse- and field-grown roses.

Greenhouse	p-value	Effect ^b (E _{525nm})	Allele			Function (identified gene)
			A:A	A:B	B:B	
Combined^a						
Rh12GR_283_1910Q	2.35E-09	-5.412	-	77	104	gene31631-v1.0-hybrid_Auxin_response_factor_8_ (putative)
RhMCRND_20203_163Q	7.20E-09	-4.091	-	81	102	gene31669-v1.0-hybrid_Medium-chain-fatty-acid-CoA_ligase_(probable)
Rh12GR_20064_1031P	1.86E-07	-4.521	102	75	-	gene31669-v1.0-hybrid_Medium-chain-fatty-acid-CoA_ligase_(probable)
RhK5_7371_202Q	2.02E-07	-4.492	-	79	104	gene31672-v1.0-hybrid_Glutathione_S-transferase_ (similar_to)
RhK5_1258_2078P	3.75E-07	-6.674	66	117	-	gene08692-v1.0-hybrid_3beta-hydroxysteroid-dehydrogenase/decarboxylase_isoform_2_(At3BETAHSD/D2)_ (similar_to)
RhMCRND_7128_1021Q	4.51E-07	-1.935	-	83	92	gene31668-v1.0-hybrid_Medium-chain-fatty-acid-CoA_ligase_(probable)
Rh12GR_64257_531Q	4.62E-07	-5.776	-	61	106	gene16750-v1.0-hybrid_hypothetical_protein
RhK5_2615_1401Q	5.58E-07	-5.549	-	76	105	gene09148-v1.0-hybrid_Cysteine_proteinase_RD21a_(RD21)_(precursor_(putative))
RhK5_15799_993Q	6.00E-07	-7.015	-	59	106	gene26062-v1.0-hybrid_Formin-like_protein_20_(AtFH20)_(similar_to)
RhMCRND_982_2342Q	1.06E-06	-5.538	-	60	117	gene26062-v1.0-hybrid_Formin-like_protein_20_(AtFH20)_(similar_to)
RhK5_1553_678P	1.10E-06	-4.719	119	65	-	gene26062-v1.0-hybrid_Formin-like_protein_20_(AtFH20)_(similar_to)
RhK5_9221_621Q	1.11E-06	6.179	76	85	-	gene16819-v1.0-hybrid_Thioredoxin_F-type_2,_chloroplastic_(Trx-F2)_(precursor_(similar_to))
Rh12GR_77973_217Q	1.17E-06	-4.794	108	75	-	-
RhK5_12663_103Q	1.32E-06	-3.628	-	69	114	gene24536-v1.0-hybrid_Possible_hemolysin_C_ (probable)
Rh12GR_17814_425Q	1.62E-06	-3.136	104	81	-	gene31679-v1.0-hybrid_Ubiquitin-like_protein_SMT3_ (probable)
RhK5_15799_993P	1.80E-06	-5.472	-	60	119	gene26062-v1.0-hybrid_Formin-like_protein_20_(AtFH20)_(similar_to)
RhMCRND_1369_1182Q	2.00E-06	5.215	-	62	121	gene26062-v1.0-hybrid_Formin-like_protein_20_(AtFH20)_(similar_to)
Rh12GR_283_1910P	2.38E-06	3.516	-	81	94	gene31631-v1.0-hybrid_Auxin_response_factor_8_ (putative)
RhK5_19460_153P	2.45E-06	4.855	123	61	-	gene04292-v1.0-hybrid_Cytokinin-O-glucosyltransferase_2_(AtZOG2)_(probable)
Rh12GR_38264_410P	2.58E-06	-4.584	88	79	-	-
Greenhouse	p-value	Effect^b (E_{525nm})	Allele A:A	Allele A:B	Allele B:B	Function (identified gene)
Rh12GR_283_1910Q	9.73E-11	-5.740	-	40	52	gene31631-v1.0-hybrid_Auxin_response_factor_8_ (putative)
RhK5_1439_806P	8.10E-09	-4.927	68	19	-	U3 small nucleolar RNA-associated protein 4/UTP4
RhMCRND_20203_163Q	1.17E-08	-5.135	-	42	51	gene31669-v1.0-hybrid_Medium-chain-fatty-acid-CoA_ligase_(probable)
RhK5_1258_2078P	9.44E-08	-7.589	33	60	-	gene08692-v1.0-hybrid_3beta-hydroxysteroid-dehydrogenase/decarboxylase_isoform_2_(At3BETAHSD/D2)_ (similar_to)
Rh12GR_20064_1031P	1.21E-07	-5.112	51	39	-	gene31669-v1.0-hybrid_Medium-chain-fatty-acid-CoA_ligase_(probable)
RhK5_7371_202Q	1.23E-07	-5.426	-	41	52	gene31672-v1.0-hybrid_Glutathione_S-transferase_ (similar_to)
RhMCRND_319_1197P	7.49E-07	4.398	-	60	26	Urease_(similar_to)
RhK5_12076_566Q	7.74E-07	-5.193	-	42	52	Photosystem_I_reaction_center_subunit_XI,_chloroplastic_(PSI-L)_(precursor_)
RhMCRND_982_2342Q	8.43E-07	-6.928	-	30	60	gene26062-v1.0-hybrid_Formin-like_protein_20_(AtFH20)_(similar_to)
Rh12GR_17814_425Q	8.45E-07	-4.391	52	42	-	gene31679-v1.0-hybrid_Ubiquitin-like_protein_SMT3_ (probable)

(Continued)

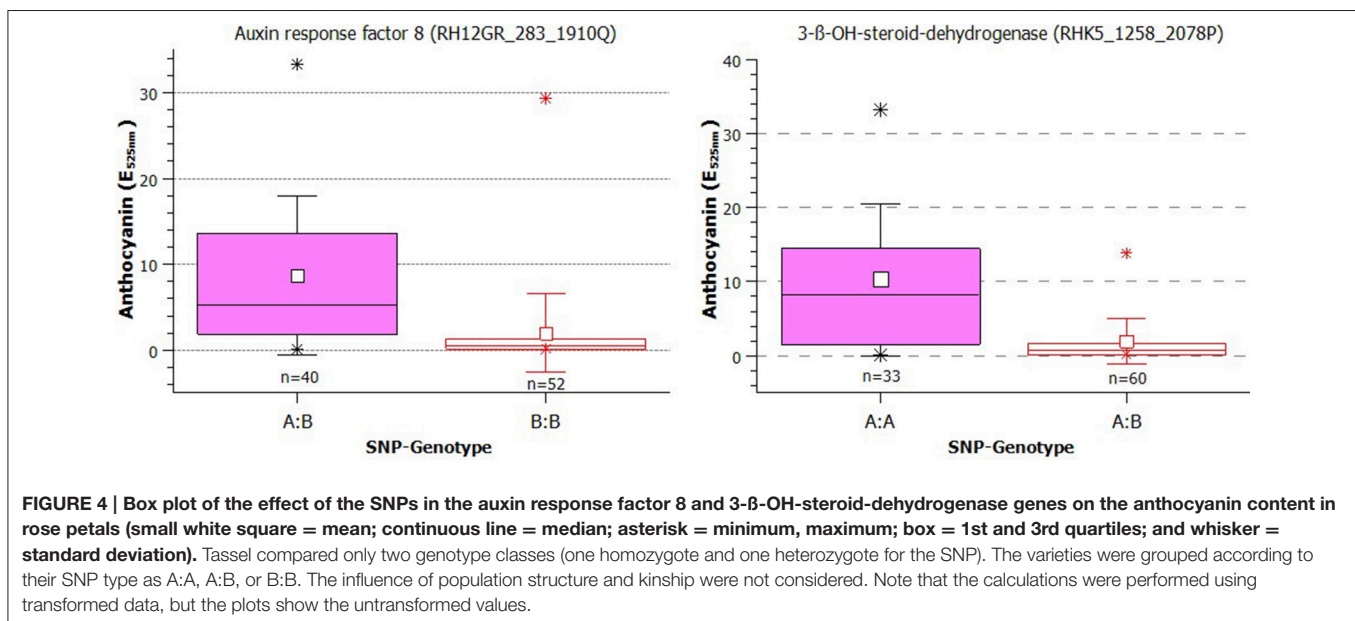
TABLE 2 | Continued

Greenhouse	p-value	Effect ^b	Function (identified gene)			
		(E _{525nm})	Allele A:A	Allele A:B	Allele B:B	
Rh12GR_92431_4144Q	8.48E-07	-3.985	-	35	58	gene31646-v1.0-hybrid_Serine/threonine-protein_kinase_PBS1_(probable)
RhK5_2615_1401Q	8.64E-07	-5.525	-	39	53	gene09148-v1.0-hybrid_Cysteine_proteinase_RD21a_(RD21)_Precursor_(putative)
RhK5_15799_993P	1.00E-06	-6.814	-	30	61	gene26062-v1.0-hybrid_Formin-like_protein_20_(AtFH20)_similar_to
RhK5_5774_854P	1.11E-06	4.321	-	62	26	Translation_initiation_factor_IF-2_
RhK5_12663_103Q	1.37E-06	4.258	-	33	57	gene24536-v1.0-hybrid_Possible_hemolysin_C_(probable)
RhK5_21626_409P	1.66E-06	4.064	-	61	25	gene07023-v1.0-hybrid_Anthranilate_phosphoribosyltransferase_(probable)
Rh12GR_3292_1365P	2.20E-06	4.122	18	68	-	gene10566-v1.0-hybrid_Putative_indole-3-acetic_acid-amido_synthetase_GH3.9_(AtGH3-9)

Field	p-value	Effect ^b	Function (identified gene)			
		(E _{525nm})	Allele A:A	Allele A:B	Allele B:B	
Rh12GR_283_1910Q	4.51E-08	-5.451	-	37	52	gene31631-v1.0-hybrid_Auxin_response_factor_8_(putative)
RhMCRND_20203_163Q	5.29E-08	-2.631	-	40	51	gene31669-v1.0-hybrid_Medium-chain-fatty-acid-CoA_ligase_(probable)
RhK5_1258_2078P	1.30E-07	-6.511	33	57	-	gene08692-v1.0-hybrid_3beta-hydroxysteroid-dehydrogenase/decarboxylase_isoform_2_(At3BETAHSD/D2)_similar_to
Rh12GR_20064_1031P	1.82E-06	5.051	51	36	-	gene31669-v1.0-hybrid_Medium-chain-fatty-acid-CoA_ligase_(probable)
RhK5_7371_202Q	1.94E-06	5.100	-	38	52	gene31672-v1.0-hybrid_Glutathione_S-transferase_(similar_to)
RhMCRND_7128_1021Q	2.37E-06	-1.901	-	39	48	gene31668-v1.0-hybrid_Medium-chain-fatty-acid-CoA_ligase_(probable)

^aTotal number of A:A, A:B, and B:B genotypes was counted in two environments and exceeded the number of 96 cultivars.

^bThe effect was presented as untransformed values.



Bonferroni threshold of $\alpha = 2.62e-6$ (Table S7), the effects of the significant SNPs on carotenoid content ranged from 0.00015 to 0.259 (E_{442nm}). Most of the significant SNPs formed two

large clusters in linkage group 5 of *F. vesca* and *P. persica* with more than 250 SNPs (Figures 7A,B). They may be located on a part of the chromosome with low recombination. Two of

the SNPs were located on contigs encoding genes of the MEP (methylerythritol 4-phosphate) pathway, CMS (2-c-methyl-d-erythritol-cyclodiphosphatase) and DXR (1-deoxy-d-xylulose 5-phosphate reducto-isomerase). A third enzyme, more upstream in the carotenoid biosynthesis, is Zeaxanthin epoxidase (ZEP, $p = 8.77e-6$). It was located in linkage groups Fvb1 and Pp07. ZEP is a part of the branch of the β -carotenoid biosynthesis that catalyzes the step from zeaxanthin to violaxanthin. Additionally, several significant SNPs were mapped to linkage group four of *F. vesca* and to linkage group one of *P. persica*. The positions of these SNPs were close to the assumed position of a carotenoid cleavage dioxygenase gene, CCD4. Other significant SNPs were located in two cytochrome P450 monooxygenases (CP450): cytochrome P450_71A24 (Pp04, Fvb5) and cytochrome CP450_CYP749A22 (Pp01, Fvb4).

The effects of SNPs on the carotenoid biosynthesis genes CMS and DXR are shown as box plots in **Figure 5**. For the DXR-SNP, the mean of the homozygous A:A genotypes was 0.074 ($E_{442\text{nm}}$), whereas the mean for the A:B genotypes was 0.325 ($E_{442\text{nm}}$). The effect of the SNP for CMS, the subsequent gene after DXR in the carotenoid pathway, was 0.206 ($E_{442\text{nm}}$), and the mean was also obviously higher in the A:B group [A:A = 0.0475 ($E_{442\text{nm}}$); A:B = 0.253 ($E_{442\text{nm}}$)].

DISCUSSION

Floral traits in ornamental roses are determined by a number of quantitative traits, e.g., petal number, flower size and a large number of secondary metabolites that constitute flower color and flower fragrance. Here, we applied GWAS based on the rose WagRhSNP array to analyse factors influencing the amount of anthocyanins and carotenoids in petals. In this study, we tried for the first time to utilize association genetics to exploit the vast phenotypic variation

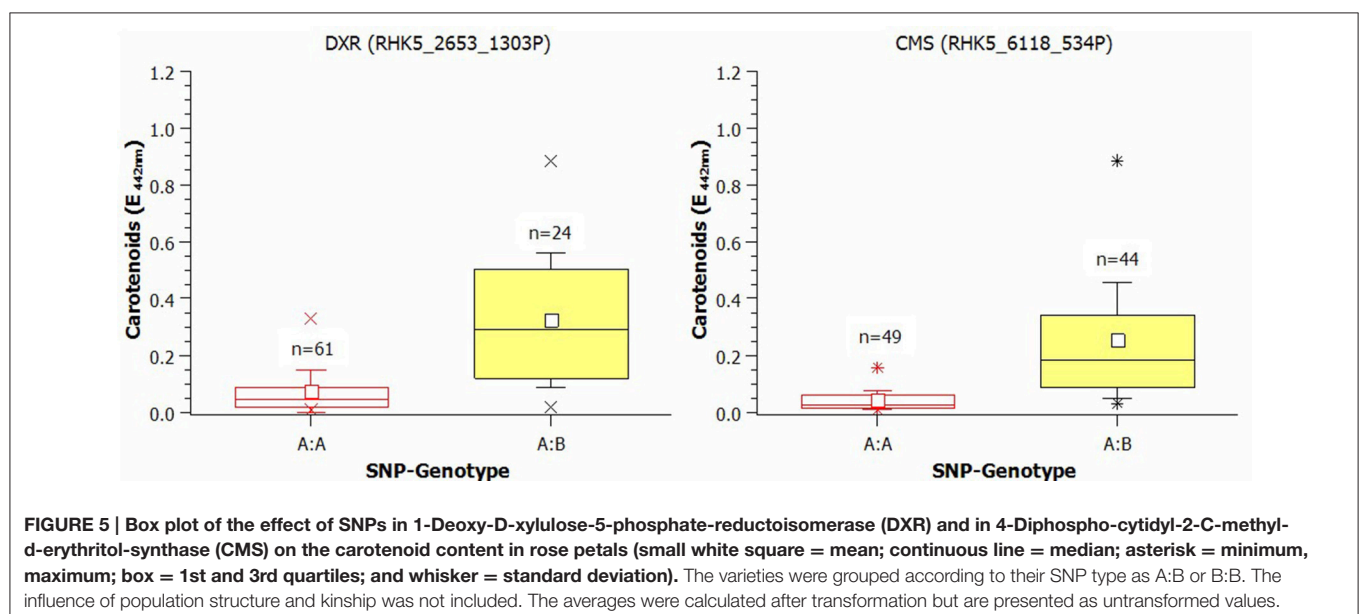
in mainly tetraploid cultivated roses for an analysis of quantitative traits.

Heterozygosity is not Influenced by Cultivar Age

The large number of markers used to genotype the association panel revealed a high average heterozygosity of 55.2% in the population, which is in agreement with previous studies on marker diversity in cultivated roses (Debener et al., 1996; Esselink et al., 2003). The value of heterozygosity is in fact the percentage of polymorphic SNPs; thus, it is not surprising that the value is somewhat lower than that found with 24 SSR markers in garden roses (Vukosavljev et al., 2013). Unlike many other cultivated plant species (Kilian et al., 2007; Gil-Ariza et al., 2009; Wang et al., 2010; Gross et al., 2014), we did not see a reduction over time when we considered the year of release of the variety. This may corroborate observations on the sensitivity of rose for inbreeding depression, which makes selfing unsuitable as a breeding strategy (Pipino et al., 2011). It might also be the result of the large range of ornamental traits that breeders have selected for, causing selection to never be consistently in the same direction. We found a lower percentage of polymorphic loci in the climber (40.9%) and ground cover genotypes (50.2%). This could indicate some degree of ascertainment bias if the genetic background of these groups of roses is partly different from those of the cut and garden roses from which the SNPs derived (Koning-Boucoiran et al., 2015; Smulders et al., 2015). This would require further studies, for instance, a study that searches for homologous regions in the genomes of diploid species that have contributed to the tetraploid groups.

Population Structure

Population structure and relatedness between genotypes can be confounding factors in association mapping (Nordborg and Weigel, 2008). Minimal population structure or relatedness will



result in high statistical power, but larger collections offer more power, and a collection of 100–500 individuals is recommended (Hirschhorn and Daly, 2005; Rafalski, 2010). The STRUCTURE software with the implemented Bayesian clustering approach is a common tool to assess population structure with a moderate number of markers. Together with the estimation of kinship, this tool can reduce the rate of false positives in association mapping (Pritchard et al., 2000; Rafalski, 2010). With 400 AFLP and 175 SSR markers, we identified three subpopulations in our rose association panel of 96 different cultivars. This is comparable to the sample sizes used in various association mapping studies with other crop species. For instance, the association panel of Lindqvist-Kreuzer et al. (2014) comprised 103 potato genotypes, and population structure was estimated with 120 SNP markers. Seventy-one almond cultivars were the basis of an association study on kernel phytosterol content (Font i Forcada et al., 2015) in which population structure was corrected using 40 SSRs. Simko et al. (2009) used 68 lettuce cultivars for association with disease resistance and validated the detected marker trait association in a second set of 132 cultivars.

Using Synteny

A major problem of association studies is filtering true marker-trait associations out of a large number of false-positive associations. Next to p -values and the extent of the observed effects, the clustering of significant markers in particular genomic regions is a criterion for true marker trait associations. However, for rose, no completed genome sequence is available and relative positions are known only for a small subset of the markers that we applied in this study. Therefore, we used two related rosaceous genomes assuming sufficient microsynteny to the rose genome: the very closely related genome of strawberry and the peach genome, which represents the next-closest relative (Shulaev et al., 2008). In several conventional marker mapping studies, strawberry was shown to be highly similar in its genome structure and marker order with only minor differences to roses (Gar et al., 2011; Spiller et al., 2011; Terefe-Ayana et al., 2012). This strategy proved informative, as we found a couple of clear clusters of significantly associated SNPs in both genomes. Our attempt to locate candidate markers to a genomic region failed for some markers (e.g., SUMO1) in one of the two genomes, but many others were located in both genomes (Figures 6A,B).

SNPs Associated With Anthocyanin Accumulation in Rose Petals

As a simple measure for red and pink flower colors, we used the total amounts of anthocyanins determined by spectrophotometry in extracts prepared at defined flower stages. This has the advantage that other factors, e.g., cellular pH, cofactors, or flower age, which influence the visual characteristics of the anthocyanins in the natural context, are excluded; therefore, the phenotypic complexity can be partially reduced. This strategy was successful, as evidenced by the high correlation between the greenhouse and the field environment, which

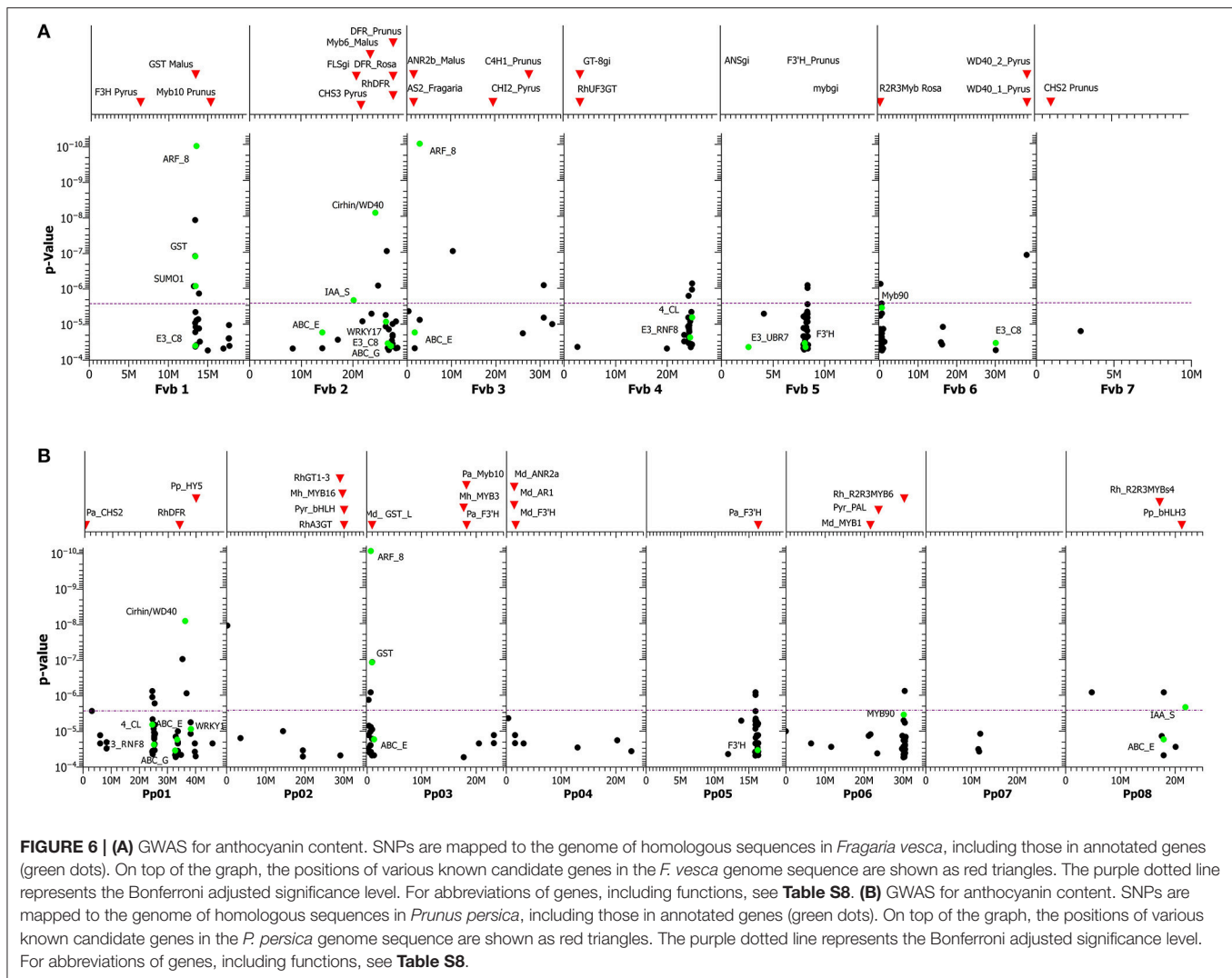
differ significantly in terms of temperature profiles and UV radiation.

Our study indicates that at least five genomic regions contain factors influencing anthocyanin concentration. Interestingly, all of these regions contained either SNP markers from genes with known functions in anthocyanin metabolism or candidate genes mapping to these regions (Table S3). The cluster in linkage group Fvb1 of *Fragaria* comprised the marker with the lowest p -value, a homolog to an auxin response factor known to influence anthocyanin concentration by regulating auxin expression in apple (Ji et al., 2015), Arabidopsis (Liu et al., 2014), tobacco (Zhu et al., 2013) and cabbage (Kang and Burg, 1973), and a family member of GSTs, which have important functions in anthocyanin transport from the cytosol to the vacuole. GSTs are responsible for color variation in a number of ornamental species, including petunia and carnation (Zhao, 2015). Homologues of the SUMO-1 transcription factor on *Fragaria* chromosome 1 are regulators of signal transduction in auxin signaling in plants (del Pozo et al., 1998; Vierstra and Callis, 1999). Because chromosome 1 of *Fragaria* is mostly collinear with the ICM linkage group 2 of *Rosa* (Gar et al., 2011), this region is likely to be close to the QTL for anthocyanin content in a diploid biparental rose population (Henz et al., 2015).

Several transcription factors on *Fragaria* chromosome 2 were associated with anthocyanin accumulation: WRKY transcription factor 17, a ubiquitin-like protein_SMT3 (SUMO1), and UTP4/Cirhin, a WD40 repeat protein (Freed et al., 2012). Zorrilla-Fontanesi et al. (2011) detected in strawberry associations between a putative R2R3 Myb transcription factor and QTLs for anthocyanin accumulation in linkage group 2. The *Fragaria* chromosome 2 is largely syntenic with rose the ICM linkage group 6 harboring major QTLs for the anthocyanin content, which were also stable across several environments (Henz et al., 2015). The cluster of SNPs mapped on *Fragaria* chromosome 4 included SNPs in the rose 4_CL gene and various Myb transcription factors, which are known to regulate anthocyanin biosynthesis. The transcription factor myb90-like (myb90), also known as “Production of Anthocyanin Pigment 2” (PAP2), was identified on *Fragaria* chromosome 6 and is a member of the MBW-complex (Maier and Hoecker, 2015). The MBW-complex activates anthocyanin biosynthetic genes and is a complex of the transcription factors R2R3-Myb, basic Helix-Loop-Helix (bHLH), and WD40 proteins (Petroni and Tonelli, 2011; Maier and Hoecker, 2015).

SNPs Associated with Carotenoid Accumulation in Rose Petals

Eugster and Märki-Fischer (1991) could identify nearly 40 different carotenoids in the extract of rose petals, most prominently Violaxanthin together with Auroxanthin, Luteoxanthin and β -Carotene (Ohmiya, 2011). Glick (2009) identified as main components Violaxanthin and Neoxanthin, which comprised 85% of the total carotenoid content in the petals of the rose cultivar “Frisco.” The biosynthetic pathway to Violaxanthin occurs via Zeaxanthin, and the modification is catalyzed by the enzyme Zeaxanthin epoxidase (ZEP). The carotenoids in the rose petals in our association study panel

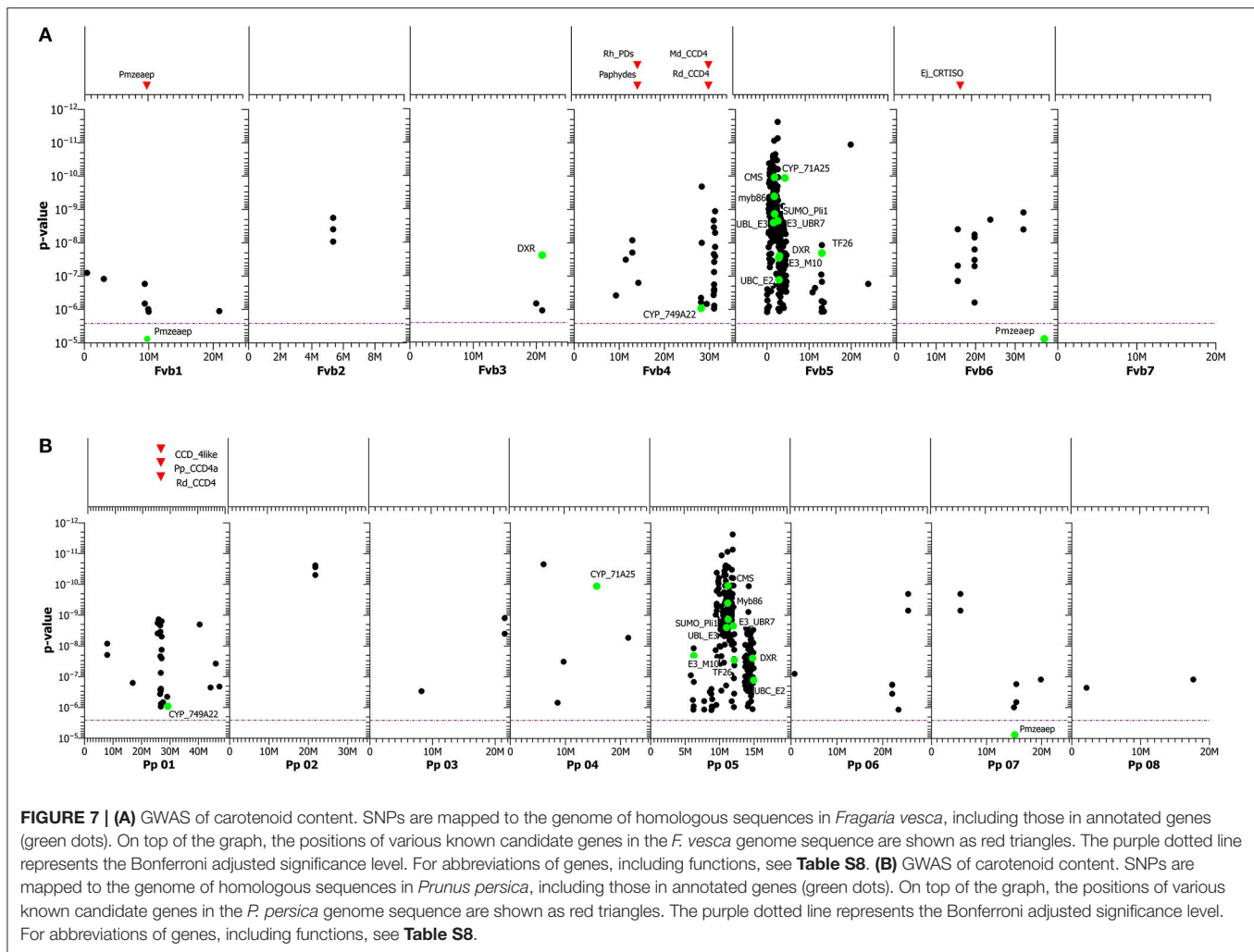


were characterized spectroscopically at 442 nm, which does not distinguish between different carotenoids. We detected as many as 303 significant SNPs associated with carotenoid content, 250 of which clustered in two positions on chromosome 5 of both *Prunus* and *Fragaria*. The majority of these 250 significant SNPs are not located in the genes causing the effect themselves, but they are closely linked to one or few of such genes located on the same chromosomal regions with low recombination rates. This extreme clustering was most probably due to the high linkage disequilibrium around the two clusters. The causes of this LD are not clear yet. Possible reasons might be either linkage to factors suppressing recombination in roses or the presence of genes under high selection pressure in cultivated roses. However, the fact that these two large clusters of significant SNPs were detected on both the *Fragaria* and the *Prunus* genome independently indicates that this is not due to a computational artifact of the GWAS or potential assembly errors in the target genome regions, but a real effect of the chromosomal region. Both target sequences have been independently assembled by different

research groups and our results on rose sequences that match the same region in both heterologous genomes indicate synteny for the location of these sequences in all three genomes. Due to the huge number of significant carotenoid SNPs, we further discuss only SNPs in potential candidate genes (**Table S4**).

Carotenoids are synthesized from isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) via the MEP pathway. Enzymes of the MEP pathway significantly influence carotenoid production (Lois, 2000; Moehs et al., 2001; Rodríguez-Concepción et al., 2001, 2003; Carretero-Paulet et al., 2002, 2006). We identified two SNPs for carotenoid accumulation in roses in the coding region of genes of the MEP pathway: CMS and DXR. The importance of DXR was shown with *Arabidopsis*, as down-regulation resulted in reduced pigmentation and defects in chloroplast development, whereas overexpression led to the accumulation of isoprenoids, such as chlorophylls and carotenoids (Carretero-Paulet et al., 2006).

Carotenoid accumulation is also influenced by degradation. Campbell et al. (2010) showed that a reduced expression of the



carotenoid cleavage dioxygenase 4 (CCD4) gene increased the carotenoid level in mature potato tubers 2- to 5-fold. Similarly, Glick (2009) found a high correlation between carotenoid degradation in the rose cultivars “Frisco” and “Golden gate” and the expression of RhCCD4. In chrysanthemum, the loss of the CmCCD4a gene caused a change in petal color from white to yellow; the level of color mutation from white to yellow may depend on the copy number of the CmCCD4a gene (Ohmiya et al., 2012; Yoshioka et al., 2012). We detected several significant SNPs for carotenoid accumulation on *Fragaria* chromosome 4 and *Prunus* chromosome 1, close to a CCD4 gene (**Figures 7A,B**, **Table S5**). It is obvious that the activity of the CCD4 genes affects the carotenoid content in different plants, but the role of the CCD4 genes in the degradation of carotenoids in roses needs more research.

Beside these SNPs located in the two regions on *Fragaria* chromosomes 4 and 5 and the corresponding *Prunus* chromosomes 1 and 5, we detected a SNP in the coding region of ZEP with a p -value of $8.77e-6$ for the association located on *Fragaria* chromosome 1.

CONCLUSIONS

This is the first association mapping study in rose. We focussed on the anthocyanin and carotenoid contents, which largely determine petal color. The phenotype data were collected in the field and in the greenhouse, and the overall levels of these compounds were not influenced by the differences in environment. To analyse the GWAS-associated SNPs in the absence of a rose genome sequence, we mapped the underlying rose contigs to the genome sequence of the related species *Fragaria vesca* and *Prunus persica*. Clusters of hits on these sequenced genomes in regions with known candidate genes confirmed that these genomes are probably largely syntenic and suggested that we identified 17 (anthocyanins) to 351 (carotenoids) trait-marker associations. Some of these had large effect sizes: these QTLs may be useful in breeding for intense flower colors in that parental breeding lines with combinations of several markers with high SNP dosages (duplex to quadruplex) might now be selected using the validated SNPs.

AUTHOR CONTRIBUTIONS

DS conducted the experiments, did the statistical analysis and wrote most parts of the manuscript. RS conducted parts of the experiments. RV conducted part of the data analysis and contributed in writing the manuscript. MS contributed some of the data and contributed in writing the manuscript. ML contributed to the experimental setup, contributed data to some of the experiments and contributed in writing the manuscript. TD was involved in planning the experiments, in statistical analysis, and wrote parts of the manuscript.

FUNDING

The research was in part funded within the program “Zentrales Innovationsprogramm Mittelstand (ZIM) of the German Bundesministerium für Wirtschaft und Energie (BMWi).”

ACKNOWLEDGMENTS

This study was carried out in part with financial support from the Aif (ZIM: KF 2554802MD0). We are grateful to the German rose breeders W. Kordes’Söhne, Rosen Tantau and Noack Rosen for providing plant material and carrying out part of the field trials. We thank Klaus Dreier and the gardeners from the Department of Molecular Plant Breeding for excellent assistance, especially Burkhard Spellerberg and Jörn Klocke, and the Federal Plant Variety Office for providing greenhouse facilities in Hannover.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2016.01798/full#supplementary-material>

Figure S1 | Frequency distribution of carotenoids (A) in greenhouse (Federal Plant Variety Office, Hannover) and absorption spectra of carotenoid extracts of three rose cultivars with the main carotenoid peak at 442 nm **(B)**.

Figure S2 | Estimation of the number of subpopulations. (A) The mean log likelihood of data $L(K)$ (\pm SD) (y -axis) as a function of K (x -axis) over ten repetitions for three independent runs in Structure 2.3.4. **(B)** ΔK (y -axis) as a function of K

(mean \pm SD) (x -axis) estimated with the method of Evanno et al. (2005) for the same runs.

Figure S3 | Principal component analysis of the association panel generated in DARwin 5.0.158 using 926 filtered AFLP and SSR markers.

The defined subpopulation I is circled in red, subpopulation II in green and subpopulation III in yellow. The cultivars are colored according to their share of belonging to a subpopulation ($\geq 70\%$) as in the neighbor joining tree **(Figure 1B)**.

Figure S4 | Distribution of relatedness in the association panel estimated in SPAGeDi.

Figure S5 | Heterozygosity of the 96 rose varieties of the association panel plotted against the age of the varieties.

Figure S6 | Frequency distribution of the total amount of anthocyanins in the greenhouse at the Federal Plant Variety Office (left) and in the field at Herrenhausen (right).

Figure S7 | Extracts from rose petals containing anthocyanins and carotenoids (above) and rose leaves containing carotenoids only (below).

Figure S8 | Spearman rank correlation between the total amount of carotenoids in rose petals from 20 cultivars grown in the field and in the greenhouse (Spearman’s rho = 0.939).

Table S1 | Rose Association panel: Cultivars, breeder, origin, and breeding year of roses, code number (1–141) and flower color. Cultivars with ploidy levels differing from tetraploid are labeled as ⁽¹⁾ or ⁽²⁾.

Table S2 | Primer sequences, annealing temperature and linkage map location in rose (according to Spiller et al. 2011) of 27 SSR markers used in the association studies with the 96 rose cultivars. Rh primers are from Esselink et al. (2003); RMS primers are published at <http://www.wipo.int/pctdb/en/wo.jsp?wo=2003097869&IA=WO2003097869&DISPLAY=STATUS>.

Table S3 | Candidate genes for anthocyanin biosynthesis localized in the genomes of *F. vesca* and *P. persica*.

Table S4 | Candidate genes for carotenoid biosynthesis in the genomes of *F. vesca* and *P. persica*.

Table S5 | SNPs significantly associated with the anthocyanin and carotenoid contents after Bonferroni correction at a significance level of $\alpha = 2.62E-6$.

Table S6 | Anthocyanin SNP markers and their position on the genome of *Fragaria vesca* 2.0.

Table S7 | SNPs significantly associated with carotenoid content and their function, p -value, and effect (untransformed value) on carotenoid content in rose petals.

Table S8 | Abbreviations for known transcription factors, carotenoid and anthocyanin biosynthetic pathway genes.

REFERENCES

- Abdurakhmonov, I. Y., Saha, S., Jenkins, J. N., Buriev, Z. T., Shermatov, S. E., Scheffler, B. E., et al. (2009). Linkage disequilibrium based association mapping of fiber quality traits in *G. hirsutum* L. variety germplasm. *Genetica* 136, 401–417. doi: 10.1007/s10709-008-9337-8
- Bourke, P. M., Voorrips, R. E., Visser, R. G. F., and Maliepaard, C. (2015). The double reduction landscape in tetraploid potato as revealed by a high-density linkage map. *Genetics* 201, 853–863. doi: 10.1534/genetics.115.181008
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308
- Campbell, R., Ducreux, L. J. M., Morris, W. L., Morris, J. A., Suttle, J. C., Ramsay, G., et al. (2010). The metabolic and developmental roles of carotenoid cleavage Dioxygenase4 from potato. *Plant Physiol.* 154, 656–664. doi: 10.1104/pp.110.158733
- Cardoso, S., Lau, W., Eiras Dias, J., Feveteiro, P., and Maniatis, N. (2012). A candidate-gene association study for berry colour and anthocyanin content in *Vitis vinifera* L. *PLoS ONE* 7:e46021. doi: 10.1371/journal.pone.0046021
- Carretero-Paulet, L., Ahumada, I., Cunillera, N., Rodríguez-Concepción, M., Ferrer, A., Boronat, A., et al. (2002). Expression and molecular analysis of the Arabidopsis DXR gene encoding 1-deoxy-D-xylulose 5-phosphate reductoisomerase, the first committed enzyme of the 2-C-methyl-D-erythritol 4-phosphate pathway. *Plant Physiol.* 129, 581–1591. doi: 10.1104/pp.003798
- Carretero-Paulet, L., Cairó, A., Botella-Pavía, P., Besumbes, O., Campos, N., Boronat, A., et al. (2006). Enhanced flux through the methylerythritol 4-phosphate pathway in Arabidopsis plants overexpressing deoxyxylulose 5-phosphate reductoisomerase. *Plant Mol. Biol.* 62, 683–695. doi: 10.1007/s11103-006-9051-9
- Cericola, F., Portis, E., Lanteri, S., Toppino, L., Barchi, L., Acciarri, N., et al. (2014). Linkage disequilibrium and genome-wide association analysis for anthocyanin pigmentation and fruit color in eggplant. *BMC Genomics* 15:896. doi: 10.1186/1471-2164-15-896

- Crespel, L., Chirollet, M., Durel, C. E., Zhang, D., Meynet, J., and Gudín, S. (2002). Mapping of qualitative and quantitative phenotypic traits in *Rosa* using AFLP markers. *Theor. Appl. Genet.* 105, 1207–1214. doi: 10.1007/s00122-002-1102-2
- De Vries, D. P., and Dubois, L. A. M. (1978). On the transmission of the yellow flower colour from *Rosa foetida* to recurrent flowering hybrid tea-roses. *Euphytica* 27, 205–210. doi: 10.1007/BF00039136
- de Vries, D. P., Van Keulen, H. A., and de Bruyn, J. W. (1974). Breeding research on rose pigments. 1. The occurrence of flavonoids and carotenoids in rose petals. *Euphytica* 27, 205–210. doi: 10.1007/bf00035892
- Debener, T. (1999). Genetic analysis of horticulturally important morphological and physiological characters in diploid roses. *Gartenbauwissenschaft* 64, 14–20.
- Debener, T. (2003). “Genetics: inheritance of characteristics,” in *Encyclopedia of Rose Science*, eds A. V. Roberts, T. Debener, and S. Gudín (Oxford: Elsevier; Academic Press), 286–292.
- Debener, T., Bartels, C., and Mattiesch, L. (1996). RAPD analysis of genetic variation between a group of rose cultivars and selected wild rose species. *Mol. Breed.* 2, 321–327. doi: 10.1007/BF00437910
- Debener, T., and Linde, M. (2009). Exploring complex ornamental genomes. The rose as a model plant. *Crit. Rev. Plant Sci.* 28, 267–280. doi: 10.1080/07352680903035481
- Deli, J., Molnár, P., Matus, Z., Tóth, G., Steck, A., and Pfander, H. (1998). Isolation and characterization of 3,5,6-trihydroxy-carotenoids from petals of *Lilium tigrinum*. *Chromatographia* 48, 27–31. doi: 10.1007/BF02467511
- del Pozo, J. C., Timpote, C., Tan, S., Callis, J., and Estelle, M. (1998). The ubiquitin-related protein RUB1 and auxin response in *Arabidopsis*. *Science* 280, 1760–1763. doi: 10.1126/science.280.5370.1760
- D’hoop, B. B., Paulo, M. J., Kowitwanich, K., Sengers, M., Visser, R. G. F., van Eck, H. J., et al. (2010). Population structure and linkage disequilibrium unravelled in tetraploid potato. *Theor. Appl. Genet.* 121, 1151–1170. doi: 10.1007/s00122-010-1379-5
- Earl, D. A., and von Holdt, B. M. (2012). Structure Harvester: a website and program for visualizing structure output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359–361. doi: 10.1007/s12686-011-9548-7
- Esselink, G. D., Smulders, M. J. M., and Vosman, B. (2003). Identification of cut rose (*Rosa hybrida*) and rootstock varieties using robust sequence tagged microsatellite site markers. *Theor. Appl. Genet.* 106, 277–286. doi: 10.1007/s00122-002-1122-y
- Eugster, C. H., and Märki-Fischer, E. (1991). The Chemistry of Rose Pigments. *Angew. Chem. Int. Ed. Engl.* 30, 654–672. doi: 10.1002/anie.199106541
- Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure. A simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x
- Falush, D., Stephens, M., and Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol. Ecol. Notes* 7, 574–578. doi: 10.1111/j.1471-8286.2007.01758.x
- Font i Forcada, C., Velasco, L., Socias i Company, R., and Fernández i Martí, Á. (2015). Association mapping for kernel phytosterol content in almond. *Front. Plant Sci.* 6:530. doi: 10.3389/fpls.2015.00530
- Freed, E. F., Prieto, J.-L., McCann, K. L., McStay, B., Baserga, S. J., and Ellis, S. R. (2012). NOL11, implicated in the pathogenesis of North American Indian childhood cirrhosis, is required for pre-rRNA transcription and processing. *PLoS Genet.* 8:e1002892. doi: 10.1371/journal.pgen.1002892
- Gar, O., Sargent, D. J., Tsai, C.-J., Pleban, T., Shalev, G., and Byrne, D. H. (2011). An autotetraploid linkage map of rose (*Rosa hybrida*) validated using the strawberry (*Fragaria vesca*) genome sequence. *PLoS ONE* 6:e20463. doi: 10.1371/journal.pone.0020463
- Gil-Ariza, D. J., Amaya, I., López-Aranda, J. M., and Sánchez-Sevilla, J. F. (2009). Impact of plant breeding on the genetic diversity of cultivated strawberry as revealed by expressed sequence tag-derived simple sequence repeat markers. *J. Am. Soc. Hortic. Sci.* 134, 337–347.
- Glick, A. (2009). *Synthesis and Degradation of Carotenoids in Cut Rose Petals during Vase Life, and Characterization of the Effect of Methyljasmonate Treatment on the Processes*. Thesis, Hebrew University of Jerusalem.
- Gross, B. L., Henk, A. D., Richards, C. M., Fazio, G., and Volk, G. M. (2014). Genetic diversity in *Malus x domestica* (Rosaceae) through time in response to domestication. *Am. J. Bot.* 101, 1770–1779. doi: 10.3732/ajb.1400297
- Grotewold, E. (2006). The genetics and biochemistry of floral pigments. *Annu. Rev. Plant Biol.* 57, 761–780. doi: 10.1146/annurev.arplant.57.032905.105248
- Gudín, S. (2000). Rose: genetics and breeding. *Plant Breed.* 17, 159–189. doi: 10.1002/9780470650134.CH3
- Hardy, O. J. (2003). Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant genetic markers. *Mol. Ecol.* 12, 1577–1588. doi: 10.1046/j.1365-294X.2003.01835.x
- Hardy, O. J., and Vekemans, X. (2002). SPADeDi. A versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Ecol. Notes* 2, 618–620. doi: 10.1046/j.1471-8286.2002.00305.x
- Henz, A., Debener, T., and Linde, M. (2015). Identification of major stable QTLs for flower color in roses. *Mol. Breed.* 35, 190. doi: 10.1007/s11032-015-0382-6
- Hirschhorn, J. N., and Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nat. Rev. Genet.* 6, 95–108. doi: 10.1038/nrg1521
- Jay, M., Biolley, J.-P., Fiasson, J.-L., Fiasson, K., Gonnet, J.-F., Grossi, C., et al. (2003). “Anthocyanins and other Flavonoid Pigments,” in *Encyclopedia of Rose Science*, Vol. 1. Elsevier, eds A. V. T. Roberts, T. Debener, and S. Gudín (Oxford: Elsevier Ltd.), 248–255.
- Ji, X.-H., Wang, Y.-T., Zhang, R., Wu, S.-J., An, M.-M., Li, M., et al. (2015). Effect of auxin, cytokinin and nitrogen on anthocyanin biosynthesis in callus cultures of red-fleshed apple (*Malus sieversii* f. niedzwetzkyana). *Plant Cell Tiss. Organ Culture* 120, 325–337. doi: 10.1007/s11240-014-0609-y
- Johnson, R. C., Nelson, G. W., Troyer, J. L., Lautenberger, J. A., Kessing, B. D., Winkler, C. A., et al. (2010). Accounting for multiple comparisons in a genome-wide association study (GWAS). *BMC Genomics* 11:724. doi: 10.1186/1471-2164-11-724
- Kang, B. G., and Burg, S. P. (1973). Role of ethylene in phytochrome-induced anthocyanin synthesis. *Planta* 110, 27–235. doi: 10.1007/BF00387635
- Kilian, B., Ozkan, H., Walther, A., Kohl, J., Dagan, T., and Salamini, F. (2007). Molecular diversity at 18 loci in 321 wild and 92 domesticated lines reveal no reduction of nucleotide diversity during *Triticum monococcum* (Einkorn) domestication. Implications for the origin of agriculture. *Mol. Biol. Evol.* 24, 2657–2668. doi: 10.1093/molbev/msm192
- Kishimoto, S., Maoka, T., Nakayama, M., and Ohmiya, A. (2004). Carotenoid composition in petals of chrysanthemum (*Dendranthema grandiflorum* (Ramat.) Kitamura). *Phytochemistry* 65, 2781–2787. doi: 10.1016/j.phytochem.2004.08.038
- Klie, M., Menz, I., Linde, M., and Debener, T. (2013). Lack of structure in the gene pool of the highly polyploid ornamental chrysanthemum. *Mol. Breed.* 32, 339–348. doi: 10.1007/s11032-013-9874-4
- Koning-Boucoiran, C. F. S., Esselink, G. D., Vukosavljev, M., van ’t Westende, W. P. C., Gitonga, V. W., Krens, F. A., et al. (2015). Using RNA-Seq to assemble a rose transcriptome with more than 13,000 full-length expressed genes and to develop the WagRhSNP 68k Axiom SNP array for rose (*Rosa* L.). *Front. Plant Sci.* 6:249. doi: 10.3389/fpls.2015.00249
- Li, X., Han, Y., Wei, Y., Acharya, A., Farmer, A. D., Ho, J., et al. (2014). Population of an alfalfa SNP array and its use to evaluate patterns of population structure and linkage disequilibrium. *PLoS ONE* 9:e84329. doi: 10.1371/journal.pone.0084329
- Linde, M., Hattendorf, A., Kaufmann, H., and Debener, T. (2006). Powdery mildew resistance in roses: QTL mapping in different environments using selective genotyping. *Theor. Appl. Genet.* 113, 1081–1092. doi: 10.1007/s00122-006-0367-2
- Lindqvist-Kreuzer, H., Gastelo, M., Perez, W., Forbes, G. A., de Koeber, D., and Bonierbale, M. (2014). Phenotypic stability and genome-wide association study of late blight resistance in potato genotypes adapted to the tropical highlands. *Phytopathology* 104, 624–633. doi: 10.1094/PHYTO-10-13-0270-R
- Liu, Z., Shi, M.-Z., and Xie, D.-Y. (2014). Regulation of anthocyanin biosynthesis in *Arabidopsis thaliana* red pap1-D cells metabolically programmed by auxins. *Planta* 239, 765–781. doi: 10.1007/s00425-013-2011-0
- Lois, L. M. (2000). Carotenoid biosynthesis during tomato fruit development: regulatory role of 1-deoxy-D-xylulose 5-phosphate synthase. *Plant J.* 22, 503–513. doi: 10.1046/j.1365-313x.2000.00764.x
- Lu, F., Lipka, A. E., Glaubitz, J., Elshire, R., Cherney, J. H., Casler, M. D., et al. (2013). Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. *PLoS Genet.* 9:e1003215. doi: 10.1371/journal.pgen.1003215
- Maier, A., and Hoecker, U. (2015). COP1/SPA ubiquitin ligase complexes repress anthocyanin accumulation under low light and high light

- conditions. *Plant Signal. Behav.* 10:e970440. doi: 10.4161/15592316.2014.970440
- Moehs, C. P., Tian, L., Osteryoung, K. W., and DellaPenna, D. (2001). Analysis of carotenoid biosynthetic gene expression during marigold petal development. *Plant Mol. Biol.* 45, 281–293. doi: 10.1023/A:1006417009203
- Moghaddam, H., Leus, L., de Riek, J., van Huylbroeck, J., and van Bockstaele, E. (2012). Construction of a genetic linkage map with SSR, AFLP and morphological markers to locate QTLs controlling pathotype-specific powdery mildew resistance in diploid roses. *Euphytica* 184, 413–427. doi: 10.1007/s10681-011-0616-6
- Nordborg, M., and Weigel, D. (2008). Next-generation genetics in plants. *Nature* 456, 720–723. doi: 10.1038/nature07629
- Ohmiya, A. (2011). Diversity of carotenoid composition in flower petals. *JARQ* 45, 163–171. doi: 10.6090/jarq.45.163
- Ohmiya, A., Toyoda, T., Watanabe, H., Emoto, K., Hase, Y., and Yoshioka, S. (2012). Mechanism behind petal color mutation induced by heavy-ion-beam irradiation of recalcitrant chrysanthemum cultivar. *J. Japan. Soc. Hort. Sci.* 81, 269–274. doi: 10.2503/jjshs.1.81.269
- Perrier, X., and Jacquemoud-Collet, J. P. (2006). *DARwin Software*. Available online at: <http://darwin.cirad.fr/darwin>
- Petroni, K., and Tonelli, C. (2011). Recent advances on the regulation of anthocyanin synthesis in reproductive organs. *Plant Sci.* 181, 219–229. doi: 10.1016/j.plantsci.2011.05.009
- Picone, J. M., Clery, R. A., Watanabe, N., MacTavish, H. S., and Turnbull, C. G. N. (2004). Rhythmic emission of floral volatiles from *Rosa damascena* semperflorens cv. 'Quatre Saisons'. *Planta* 219, 468–478. doi: 10.1007/s00425-004-1250-5
- Pipino, L., van Labeke, M. C., Mansuino, A., Scariot, V., Giovannini, A., and Leus, L. (2011). Pollen morphology as fertility predictor in hybrid tea roses. *Euphytica* 178, 203–214. doi: 10.1007/s10681-010-0298-5
- Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Rafalski, J. A. (2010). Association genetics in crop improvement. *Curr. Opin. Plant Biol.* 13, 174–180. doi: 10.1016/j.pbi.2009.12.004
- Rodríguez-Concepción, M., Ahumada, I., Diez-Jueves, E., Sauret-GuÉto, S., Lois, L. M., Gallego, F., et al. (2001). 1-Deoxy-d-xylulose 5-phosphate reductoisomerase and plastid isoprenoid biosynthesis during tomato fruit ripening. *Plant J.* 27, 213–222. doi: 10.1046/j.1365-313x.2001.01089.x
- Rodríguez-Concepción, M., Querol, J., Lois, L. M., Santiago Imperial, S., and Boronat, A. (2003). Bioinformatic and molecular analysis of hydroxymethylbutenyl diphosphate synthase (GCPE) gene expression during carotenoid accumulation in ripening tomato fruit. *Planta* 217, 476–482. doi: 10.1007/s00425-003-1008-5
- Roman, H., Rapicault, M., Miclot, A. S., Larenaudie, M., Kawamura, K., Thouroude, T., et al. (2015). Genetic analysis of the flowering date and number of petals in rose. *Tree Genet. Genomes* 11, 85. doi: 10.1007/s11295-015-0906-6
- Saitou, N., and Nei, M. (1987). The Neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.
- Shulaev, V., Korban, S. S., Sosinski, B., Abbott, A. G., Aldwinckle, H. S., Folta, K. M., et al. (2008). Multiple models for Rosaceae genomics. *Plant Physiol.* 147, 985–1003. doi: 10.1104/pp.107.115618
- Simko, I., Pechenick, D. A., McHale, L. K., Truco, M. J., Ochoa, O. E., Michelmore, R. W., et al. (2009). Association mapping and marker-assisted selection of the lettuce dieback resistance gene Tvr1. *BMC Plant Biol.* 9:135. doi: 10.1186/1471-2229-9-135
- Smulders, M. J. M., Voorrips, R. E., Esselink, G. D., Santos Leonardo, T. M., van 't Westende, W. P. C., Vukosavljev, M., et al. (2015). Development of the WagRhSNP Axiom SNP array based on sequences from tetraploid cut roses and garden roses. *Acta Hort.* 1064, 177–184. doi: 10.17660/ActaHortic.2015.1064.20
- Spiller, M., Linde, M., Hibrand-Saint Oyant, L., Tsai, C.-J., Byrne, D. H., Smulders, M. J. M., et al. (2011). Towards a unified genetic map for diploid roses. *Theor. Appl. Genet.* 122, 489–500. doi: 10.1007/s00122-010-1463-x
- Tanaka, Y., Sasaki, N., and Ohmiya, A. (2008). Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J.* 54, 733–749. doi: 10.1111/j.1365-313X.2008.03447.x
- Terefe-Ayana, D., Kaufmann, H., Linde, M., and Debener, T. (2012). Evolution of the Rdr1 TNL-cluster in roses and other Rosaceous species. *BMC Genomics* 13:409. doi: 10.1186/1471-2164-13-409
- Tinoi, J., Rakariyatham, N., and Deming, R. L. (2006). Determination of major carotenoid constituents in petal extracts of eight selected flowering plants in the north of Thailand. *Chiang. Mai. J. Sci.* 33, 327–334.
- Vasilief, I. (2015). *QtiPlot - Data Analysis and Scientific Visualisation Version 0.9.9*. Available online at: <http://soft.proindependent.com/qtiplot.html>
- Vierstra, R. D., and Callis, J. (1999). Polypeptide tags, ubiquitous modifiers for plant protein regulation. *Plant Mol. Biol.* 41, 435–442. doi: 10.1023/A:1006323317890
- Voorrips, R. E., Gort, G., and Vosman, B. (2011). Genotype calling in tetraploid species from bi-allelic marker data using mixture models. *BMC Bioinform.* 12:172. doi: 10.1186/1471-2105-12-172
- Vukosavljev, M., Zhang, J., Esselink, G. D., van 't Westende, W. P. C., Cox, P., Visser, R. G. F., et al. (2013). Genetic diversity and differentiation in roses. A garden rose perspective. *Sci. Hortic.* 162, 320–332. doi: 10.1016/j.scienta.2013.08.015
- Wang, C., Chen, J., Zhi, H., Yang, L., Li, W., Wang, Y., et al. (2010). Population genetics of foxtail millet and its wild ancestor. *BMC Genet.* 11:90. doi: 10.1186/1471-2156-11-90
- Wellburn, A. R. (1994). The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. *J. Plant Physiol.* 144, 307–313. doi: 10.1016/S0176-1617(11)81192-2
- Wessa, P. (2016). *Free Statistics Software, Office for Research Development and Education, version 1.1.23-r7*. Available online at: <http://www.wessa.net/>
- Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A., and Tingey, S. V. (1990). DNA polymorphism amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* 18, 6531–6535. doi: 10.1093/nar/18.22.6531
- Wissemann, V. (2003). "Classification / Conventional taxonomy (wild roses)," in *Encyclopedia of Rose Science*, eds A. V. Roberts, T. Debener, and S. Gudin (Oxford: Elsevier Ltd.), 111–117.
- Yamamizo, C., Kishimoto, S., and Ohmiya, A. (2010). Carotenoid composition and carotenogenic gene expression during Ipomoea petal development. *J. Exp. Bot.* 61, 709–719. doi: 10.1093/jxb/erp335
- Yoshioka, S., Aida, R., Yamamizo, C., Shibata, M., and Ohmiya, A. (2012). The carotenoid cleavage dioxygenase 4 (CmCCD4a) gene family encodes a key regulator of petal color mutation in chrysanthemum. *Euphytica* 184, 377–387. doi: 10.1007/s10681-011-0602-z
- Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38, 203–208. doi: 10.1038/ng1702
- Zhang, J., Esselink, G. D., Che, D., Fougère-Danezan, M., Arens, P., and Smulders, M. J. M. (2013). The diploid origins of allopolyploid rose species studied using single nucleotide polymorphism haplotypes flanking a microsatellite repeat. *J. Hortic. Sci. Biotechnol.* 88, 85–92. doi: 10.1080/14620316.2013.11512940
- Zhao, J. (2015). Flavonoid transport mechanisms: how to go, and with whom. *Trends Plant Sci.* 20, 576–585. doi: 10.1016/j.tplants.2015.06.007
- Zhu, Q., Li, B., Mu, S., Han, B., Cui, R., Xu, M., et al. (2013). TTG2-regulated development is related to expression of putative AUXIN RESPONSE FACTOR genes in tobacco. *BMC Genomics* 14:806. doi: 10.1186/1471-2164-14-806
- Zorrilla-Fontanesi, Y., Cabeza, A., Domínguez, P., Medina, J. J., Valpuesta, V., Denoyes-Rothan, B., et al. (2011). Quantitative trait loci and underlying candidate genes controlling agronomical and fruit quality traits in octoploid strawberry (*Fragaria × ananassa*). *Theor. Appl. Genet.* 123, 755–778. doi: 10.1007/s00122-011-1624-6

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Schulz, Schott, Voorrips, Smulders, Linde and Debener. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.