Supplemental Methods 1. DNA sequencing and reference genomes.

Reference draft genomes of *P. axillaris* and *P. exserta* used were taken from DNAZoo (https://www.dnazoo.org/assemblies/Petunia_axillaris and https://www.dnazoo.org/assemblies/Petunia_exserta) and modified as follows: For *Petunia exserta*, the first seven scaffolds were renamed as follows: HiC_scaffold_1 was renamed to PeexChr1, HiC_scaffold_2 to PeexChr3, HiC_scaffold_3 to PeexChr6, HiC_scaffold_4 to PeexChr4, HiC_scaffold_5 to PeexChr5, HiC_scaffold_6 to PeexChr7 and HiC_scaffold7 to PeexChr2. HiC_scaffold_4 and HiC_scaffold_4 were reverse complemented to conform to existing genetic maps. For *Petunia axillaris,* HiC_scaffold_1 was renamed to PeaxChr2, HiC_scaffold_2 to PeaxChr1, HiC_scaffold_3 to PeaxChr4, HiC_scaffold_4 to PeaxChr3, HiC_scaffold_5 to PeaxChr6, HiC_scaffold_6 to PeaxChr7 and HiC_scaffold7 to PeaxChr5. HiC_scaffold_1, HiC_scaffold_2 HiC_scaffold_4, and HiC_scaffold_5 were reverse complemented to conform to existing genetic maps. In addition, all scaffolds with fewer than 1000 bp were removed. These updated versions of each genome are *P. axillaris* v3.04 and *P. exserta* v3.04.

Structural annotations of *P. exserta* version 3.04 genome and *P. axillaris* genome 3.04 were performed with GMAP version 2019.03.15 (Wu and Watanabe, 2005). The *P. axillaris* v.1.6.2 structural annotations were mapped to *P. axillaris* genome v3.04 and the *Petunia exserta* v1.1.3 annotations to the *Petunia exserta* 3.04 genome. *P. exserta* and *P. axillaris* gene, mRNA and CDS annotations are available on Zenodo (https://doi.org/ 10.5281/zenodo.4384741). The reference genome for *P. inflata* was v1.0.1 from (Bombarely et al., 2016), assembly and annotation files accessible on the SolGenomics website (https://solgenomics.net/organism/Petunia_inflata/genome).

We completed a draft assembly of *Petunia secreta* for this manuscript. *P. secreta* plants were grown under sterile conditions in tissue culture containers and used for DNA extraction as described in Bombarely et al. (2016). A PCR-free library was prepared producing 150 bp paired-end reads with 400 bp library insert size and sequenced on an Illumina HiSeq3000 platform at the University of Bern Next Generation Sequencing (NGS) Platform (Bern, Switzerland). Reads were quality-controlled using Trimmomatic v.0.36 (Bolger et al., 2014) and the results were evaluated using FastQC software v.0.11.5

(https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). For the assembly, an optimal kmer value of 119 was estimated with KmerGenie v1.7048 (Chikhi and Medvedev, 2013) and a genome size of 1.169 Gbp was estimated with Jellyfish v.2.1.0 (Marçais and Kingsford, 2011). Assembly of quality-controlled reads was performed by ABySS v.2.0.2 using abyss-pe (parameters: k=119, B=40G H=6 kc=3 v=-v) (Simpson et al., 2009; Jackman et al., 2017). This achieved an average read depth of 75x, n=2436959, n:500 = 199133, L50 = 13943, N80 = 20002, N50 = 41755.

P. secreta structural annotation was performed using MAKER v2.31.9 (Campbell et al., 2014). SNAP v2013.11.29 with a *P. secreta* model and Augustus v3.2.3 with the tomato model were used as *ab initio* gene predictors (Korf, 2004; Stanke et al., 2004). Exonerate was to improve alignment around splice cites. Evidence used as input included *P. secreta* transcriptomes and the proteins from the *P. axillaris* genome (Slater and Birney, 2005). Masking was performed with RepeatMasker v4.0.7 (Smit et al., 2015). BLAST algorithms were used from NBCI BLAST v2.7.1 (Altschul et al., 1997). *P. secreta* RNAseq Illumina data (described below) were assembled into predicted transcripts using Trinity v2.4.0 in genome-guided mode with *P. axillaris* v3.0.4 as the reference genome (Grabherr et al., 2011; Haas et al., 2013). tRNAs were annotated using tRNAscan-SE v2.0.0 and snoRNAs using Snoscan v.0.9.1 (Lowe and Eddy, 1999; Chan and Lowe, 2019). Gene functional annotation was performed by sequence homology searches against the Uniprot dataset using BLASTP and protein domains search using InterProScan v5.33.72.0 (Jones et al., 2014). The *P. secreta* genome assembly and gene annotation files are uploaded to the NCBI SRA database under PRJNA674325.

Supplemental Methods 2. RIL population generation, GBS sequencing, and genetic map estimation.

An F7 mapping population of 195 progenies were bred from selfed progeny of an F2 population of the parental species *P. axillaris* x *P. exserta*. DNA samples were extracted from leaf material using a CTAB protocol (Murray and Thompson, 1980). DNA samples were sent to the Institute for Genomic Diversity (Cornell University, Ithaca, NY, USA) for library preparation and pooled sequencing following protocols described in Elshire et al. (2011). *PstI* was used as digestion enzyme. Libraries were sequenced on six lanes of a HiSeq2000 machine for 101 cycles single-ended resulting in 318.4 million raw reads (195 samples). Three biological replicates of each parental species were also sequenced. Demultiplexing (http://sourceforge.net/projects/gbsbarcode/) and removal of low-quality reads with fastq-mcf (Aronesty, 2011) discarded 11.9 % of the reads. Remaining reads were mapped to the draft genome of *Petunia axillaris* (v.1.6.2 from Bombarely et al. (2016)) using bowtie2 (Langmead and Salzberg, 2012) and 1998 SNP-markers were extracted from alignments with the ANGSD toolkit (Korneliussen et al., 2014) with the following parameters: -doCounts 1 -dumpCounts 2 -GL 1 -doMaf 2 -doMajorMinor 1 -doGeno 21 -doPost 1 -postCutoff 0.7 -geno_minDepth 3 -geno_maxDepth 1000000 -minMAF 0.05 –minInd 100. Inhouse python scripts were used to assign parental genotypes to SNP markers (available at https://github.com/MichelMoser/GBS2map).

The physical locations of the genetic map from the *P. axillaris* draft genome v1.6.2 were updated to draft genome v.3.0.4 (chromosome-level assembly) using a series of custom scripts (https://github.com/ginac/GeneticMap). Briefly, a 201bp region including each SNP marker and 100 bp up and downstream from the marker were extracted from the *P. axillaris* v.1.6.2 genome and BLASTED against the *P. axillaris* v.3.0.4 genome. Sequences were extracted with Seqret from the EMBOSS toolkit (Rice et al., 2000). These sequences were then BLASTED to the *P. axillaris* v3.0.4 genome. Markers with perfect size matches to the *P. axillaris* v3.0.4 genome were used in subsequent analyses. The physical information provided by this step was used to identify the linkage groups as chromosomes.

Supplemental Methods 3. *Stable* Agrobacterium-*mediated transformation of* P. axillaris

The construct used for transformation, pNWA12 (*35S$_{pro}$*:*DPL*) from Albert et al. (2011), was kindly gifted by Nick Albert. Stable transgenic *P. axillaris* N lines were generated by leaf disc transformation with *Agrobacterium tumefaciens* strain LBA4404 following an adapted version of the protocol described by Conner et al. (2009). The *A. tumefaciens* culture (LBA4404-pNWA12) was grown at 28°C with vigorous shaking in YEB broth (5 g/L meat extract, 5 g/L Bacto Peptone, 5 g/L sucrose, 1 g/L yeast extract, 240 mg/L $MgSO_4$-$7H_2O$) containing rifampicin, streptomycin, and kanamycin antibiotics (50 mg/L, 50 mg/L, 100 mg/L respectively) for bacterial selection. A dilution of this culture (1:1 of culture:water) was used for transformation. Young leaves from five to six weeks old *P. axillaris* N plants were sterilized in 70% ethanol and dried on filter paper. Single leaves were cut into squares of 1 cm^2 using a sharp scalpel. Leaf squares were immediately incubated in the transformation culture in the dark for 15-30 min while shaking.

After incubation, 30-40 dried leaf squares were placed onto Murashige and Skoog (MS) medium consisting of 4.4 g/L MS salts with B5 vitamins (Duchefa), 20 g/L sucrose, 10 g/L glucose, 100 ul $CuSO_4$, 200 uM acetosyringone, 3 mg/L 6-benzylaminopurine (BAP), 0.2 mg/L indole-3-acetic acid (IAA), 400 mg/L cefotaxime, 10 mg/L natamycin and 0.68% Micro Agar (Duchefa) at pH 5.8. Plates were then placed in a growth chamber (18h light/6 hrs dark, 24°C/18°C) in a darkened box to eliminate light. After 24 hours, the box was removed and samples were transferred onto MS medium containing 150 mg/L kanamycin for specific selection of transformed cells carrying the pNWA12 vector. After four weeks, regenerating shoots were harvested and placed onto MS medium with 50 mg/L kanamycin, 3 mg/L indole-3-butyric acid (IBA), 1 mg/L folic acid, 400 mg/L cefotaxime, and 10 mg/L natamycin. Shoots that developed roots were excised, put onto soil, and analysed for the presence of the *DPL* transgene by PCR and Sanger sequencing. Nine independent lines from the T0 generation were established with two T1 lines (L2 and L7) selected for characterization.

Supplemental Methods 4. LC-UV and LC-MS chromatographic methods.

Quantitative LC-MS-UV detection for P. exserta *flavonoids*

To determine the precise anthocyanidins and flavonol aglycones in *P. exserta*, we used LC-MS-UV for three petal limbs from each of three biological replicates. A Waters Aquity I-Class UPLC system was used with a binary solvent manager, Sample Manager-FL autosampler, eLambda Detector 800 nm photodiode array, and QDa Detector (single quadrupole mass spectrometer with ESI). Samples were cooled at 4°C before injection in the autosampler, and 5 µl of each sample was injected onto an ACQUITY UPLC BEH C18 1.7 µm column kept at 40°C with a flow rate of 0.2 mL/min. Solvent A = H_2O: acetonitrile: formic acid 98.9:1:0.1; Solvent B = Acetonitrile: formic acid 99.9: 0.1. A 23-minute gradient with linear and nonlinear ramps (exponential increases represented as Waters gradient curve numbers) was established with the following conditions: 0-2 min 99% A, linear; 2-10 min 85% A, curve 7; 10-13 min 80% A, curve 9; 13-19 min 40% A, linear; 19-21 min hold at 0% A, 21.1-23 min 99% A, linear. Peaks were detected using a coupled Waters Acquity PDA (model: UPLC eLambda 800 nm) at 520 nm (anthocyanidins) and 365 nm (flavonols), scanning from 190 to 790 nm at a resolution of 1.2 nm and a sampling rate of 20 points/sec. Mass spectrometry with a Waters QDa detector coupled to the UPLC system was used to verify compound identity, in addition to retention time and UV spectra. Flavonoids were detected in positive electrospray ionization mode (m/z 150-400) from 0-21 minutes, with data collection in centroid mode, scan time 0.195 sec, capillary voltage 0.80 kV, cone voltage 10V, ion source temperature 120°C and probe temperature 600°C. As the pre-selected ionization parameters caused only little fragmentation, only the precursor ions were considered ([M+H]+ in ESI+ mode) for compound identification. Data was visualized using MassLynx v4.1 SCN888 and peaks were integrated with the included TargetLynx.

Qualitative LC-MS for anthocyanin identification in P. exserta, P. secreta, P. inflata

Methanolic extracts (1 mL MeOH with 1% formic acid) were prepared on ice with three whole petal limbs each of *P. exserta, P. secreta,* and *P. inflata*, and diluted 1:40, 1:50, and 1:100. Samples were analyzed immediately to prevent anthocyanin degradation. We based the separation method after (Züst et al., 2020). We analyzed extracts on an Acquity UHPLC system coupled to a Xevo G2-XS QTOF mass

spectrometer with electrospray ionization (Waters, Milford MA, USA). Extracts were separated on a Waters Acquity BEH C18 100 × 2.1 mm column with 1.7 µm pore size, fitted with a BEH guard column. The column was maintained at 40°C and injections of 1 µl were eluted at a constant flow rate of 0.4 mL/min with a gradient of 0.1% formic acid in H_2O (A) and 0.1% formic acid in acetonitrile (B) as follows: 0-5.5 min from 10% to 65% B, 5.5-6 min from 65% to 100% B, followed by a 2 min wash phase at 100% B, and 2 min reconditioning at 10% B. Compounds were ionized in positive mode and ion data were acquired over an m/z range of 50 to 1200 Da in MSE mode using alternating scans of 0.15 s at low collision energy of 6 eV and 0.15 s and at high collision energy ramped from 10 to 40 eV. The electrospray capillary voltage was set to 2 kV and the cone voltage was set to 20 V. The source temperature was maintained at 140°C and the desolvation gas temperature at 400°C. The desolvation gas flow was set to 1000 L/h, and argon was used as a collision gas. The mobile phase was diverted to waste during the wash and reconditioning phase at the end of each gradient. Accurate mass measurements were obtained by infusing a solution of leucine-enkephalin at 200 ng/mL at a flow rate of 10 µL/min through the LockSpray probe.

We also ran an extended method on the *P. exserta* extracts to separate co-eluting compounds with identical flow rate, solvents, and injection volume as follows: 0-1 min from 1% to 5% B, 1-6.5 min from 5% to 50% B, 6.5-7.5 min from 50% to 95% B, followed by a 2 min wash phase at 100% B, and 2 min reconditioning at 1% B. For the extended method, compounds were detected with a PDA at 520 nm and scanning from 190-790nm at a resolution of 1.2 nm at a sampling rate of 20 points/sec. Compounds were ionized in positive mode and ion data were acquired over an m/z range of 100 to 1200 Da in MSE mode using alternating scans of 0.15 s at low collision energy of 40 eV and 0.15 s and at high collision energy ramped from 50 to 100 eV. The electrospray capillary voltage was set to 3 kV and the cone voltage was set to 20 V. All other MS parameters were the same as above.

Compounds were identified as in Berardi et al. (2016). We extracted ion chromatograms from the high energy data of masses corresponding to anthocyanidins (cyanidin, delphinidin, petunidin, peonidin, malvidin) and subsequently extracted ion chromatograms of the respective parent masses. We then extracted ion chromatograms of corresponding fragments to verify the likely fragmentation pattern of the anthocyanin. Care was taken to estimate m/z as

accurately as possible as the difference between molecular ion masses of many sugar and acyl moieties is very small (e.g. glucose 162.053, *p*-coumaric acid 162.032).

Although LC-MS cannot confirm precise flavonoid structure, we assigned putative compound identities with comparisons to mass spectral databases. The *m/z* values of each compound were submitted to the Fragment Search function at the ReSpect database (Sawada et al., 2012) and a putative identification was made based on the best-hit compound. In tandem, we consulted the KNApSAck database (Afendi et al., 2012) as well as literature on *Petunia* and other related Solanaceae (Griesbach and Asen, 1990; Griesbach, 1993; Ando et al., 1999; Griesbach et al., 1999; Tatsuzawa et al., 1999; Ando et al., 2000; Jin et al., 2015)

Quantitative LC-UV for P. axillaris 35S$_{pro}$:DPL and VIGS *samples*

For all other samples, including stable *P. axillaris 35S$_{pro}$:DPL* lines and VIGS samples, anthocyanidin and flavonol algycones were quantitated using standard HPLC-UV. Note that VIGS tissues were sampled by sector when possible (as in *AAT* experiment) but otherwise whole petal limbs were collected. Hydrolyzed extracts were resuspended in 50 µl methanol:water:formic acid (75:24:1 [v/v]). Injections (5 µl) were analyzed with a Dionex Ultimate 3000 Series HPLC with a Phenomenex Kinetix LC C18 column (1.7 µm particle size, 150 × 2.1 mm). Flavonoids were separated by gradient elution at a flow rate of 0.4 ml/min at 35°C using solvents A (H_2O with 0.1% formic acid [v/v]) and B (acetonitrile with 0.1% formic acid) with the following protocol: 0–2 min 10% B, 2–7 min 20%, 7–11 min 60% B, 11–12 min 100% B, 12–16 min 100%, 16–16.1 min 10% B and 21 min 10% B. Peaks were detected at 520 nm and 365 nm using a photodiode array detector (DAD) scanning from 190 to 790 nm. Chromatograms were visualized using Dionex Chromeleon (Version 7.1.0.898). Quantification of each compound was conducted by creating a dilution series of a standard compound of each flavonoid aglycone. Anthocyanidins, flavonols and flavones were identified by comparison with retention times and UV spectra of the following standard solutions: delphinidin, cyanidin, pelargonidin, malvidin, peonidin, petunidin, kaempferol, quercetin and myricetin from Extrasynthese (Genay, France).

Supplemental References

Afendi, F.M., Okada, T., Yamazaki, M., Hirai-Morita, A., Nakamura, Y., Nakamura, K., Ikeda, S., Takahashi, H., Altaf-Ul-Amin, M., Darusman, L.K., Saito, K., and Kanaya, S. (2012). KNApSAcK family databases: integrated metabolite-plant species databases for multifaceted plant research. Plant & Cell Physiology **53,** e1-e1.

Albert, N.W., Lewis, D.H., Zhang, H., Schwinn, K.E., Jameson, P.E., and Davies, K.M. (2011). Members of an R2R3-MYB transcription factor family in *Petunia* are developmentally and environmentally regulated to control complex floral and vegetative pigmentation patterning. The Plant Journal **65,** 771-784.

Albert, N.W., Davies, K.M., Lewis, D.H., Zhang, H., Montefiori, M., Brendolise, C., Boase, M.R., Ngo, H., Jameson, P.E., and Schwinn, K.E. (2014). A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. The Plant Cell **26,** 962-980.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research **25,** 3389-3402.

Ando, T., Tatsuzawa, F., Saito, N., Takahashi, M., Tsunashima, Y., Numajir, H., Watanabe, H., Kokubun, H., Hara, R., Seki, H., and Hashimoto, G. (2000). Differences in the floral anthocyanin content of red petunias and *Petunia exserta*. Phytochemistry **54,** 495-501.

Ando, T., Saito, N., Tatsuzawa, F., Kakefuda, T., Yamakage, K., Ohtani, E., Koshi-ishi, M., Matsusake, Y., Kokubun, H., Watanabe, H., Tsukamoto, T., Ueda, Y., Hashimoto, G., Marchesi, E., Asakura, K., Hara, R., and Seki, H. (1999). Floral anthocyanins in wild taxa of *Petunia* (Solanaceae). Biochemical Systematics and Ecology **27,** 623-650.

Aronesty, E. (2011). ea-utils: Command-line tools for processing biological sequencing data (Durham, NC).

Berardi, A.E., Hildreth, S.B., Helm, R.F., Winkel, B.S.J., and Smith, S.D. (2016). Evolutionary correlations in flavonoid production across flowers and leaves in the Iochrominae (Solanaceae). Phytochemistry **130,** 119-127.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30,** 2114-2120.

Bombarely, A., Moser, M., Amrad, A., Bao, M., Bapaume, L., Barry, C.S., Bliek, M., Boersma, M.R., Borghi, L., Bruggmann, R., Bucher, M., D'Agostino, N., Davies, K., Druege, U., Dudareva, N., Egea-Cortines, M., Delledonne, M., Fernandez-Pozo, N., Franken, P., Grandont, L., Heslop-Harrison, J.S., Hintzsche, J., Johns, M., Koes, R., Lv, X., Lyons, E., Malla, D., Martinoia, E., Mattson, N.S., Morel, P., Mueller, L.A., Muhlemann, J., Nouri, E., Passeri, V., Pezzotti, M., Qi, Q., Reinhardt, D., Rich, M., Richert-Poggeler, K.R., Robbins, T.P., Schatz, M.C., Schranz, M.E., Schuurink, R.C., Schwarzacher, T., Spelt, K., Tang, H., Urbanus, S.L., Vandenbussche, M., Vijverberg, K., Villarino, G.H., Warner, R.M., Weiss, J., Yue, Z., Zethof, J., Quattrocchio, F., Sims, T.L., and Kuhlemeier, C. (2016). Insight into the evolution of the Solanaceae from the parental genomes of *Petunia hybrida*. Nature Plants **2,** 16074.

Campbell, M.S., Holt, C., Moore, B., and Yandell, M. (2014). Genome Annotation and Curation Using MAKER and MAKER-P. Current protocols in bioinformatics **48,** 4.11.11-39.

Chan, P.P., and Lowe, T.M. (2019). tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. Methods in Molecular Biology **1962,** 1-14.

Chikhi, R., and Medvedev, P. (2013). Informed and automated k-mer size selection for genome assembly. Bioinformatics **30,** 31-37.

Conner, A.J., Albert, N.W., and Deroles, S.C. (2009). Transformation and Regeneration of *Petunia*. In *Petunia*: Evolutionary, Developmental and Physiological Genetics, T. Gerats and J. Strommer, eds (New York, NY: Springer New York), pp. 395-409.

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in *Arabidopsis*. Trends in Plant Science **15,** 573-581.

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., and Mitchell, S.E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLOS ONE **6,** e19379.

Esfeld, K., Berardi, A.E., Moser, M., Bossolini, E., Freitas, L., and Kuhlemeier, C. (2018). Pseudogenization and resurrection of a speciation gene. Current Biology **28,** 3776-3786.

Gates, D.J., Olson, B.J.S.C., Clemente, T.E., and Smith, S.D. (2018). A novel R3 MYB transcriptional repressor associated with the loss of floral pigmentation in *Iochroma*. New Phytologist **217,** 1346-1356.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology **29,** 644-652.

Griesbach, R.J. (1993). Characterization of the flavonoids from *Petunia x hybrida* flowers expression the A1 gene of *Zea mays*. HortScience **28,** 659-660.

Griesbach, R.J., and Asen, S. (1990). Characterization of the flavonol glycosides in *Petunia*. Plant Science **70,** 49-56.

Griesbach, R.J., Stehmann, J.R., and Meyer, F. (1999). Anthocyanins in the "red" flowers of *Petunia exserta*. Phytochemistry **51,** 525–528.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., LeDuc, R.D., Friedman, N., and Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature Protocols **8,** 1494-1512.

Hichri, I., Barrieu, F., Bogs, J., Kappel, C., Delrot, S., and Lauvergeat, V. (2011). Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. Journal of Experimental Botany **62,** 2465-2483.

Horai, H., Arita, M., Kanaya, S., Nihei, Y., Ikeda, T., Suwa, K., Ojima, Y., Tanaka, K., Tanaka, S., and Aoshima, K. (2010). MassBank: a public repository for sharing mass spectral data for life sciences. Journal of Mass Spectrometry **45,** 703-714.

Jackman, S.D., Vandervalk, B.P., Mohamadi, H., Chu, J., Yeo, S., Hammond, S.A., Jahesh, G., Khan, H., Coombe, L., Warren, R.L., and Birol, I. (2017). ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter. Genome Research **27,** 768-777.

Jin, H., Liu, Y., Yang, F., Wang, J., Fu, D., Zhang, X., Peng, X., and Liang, X. (2015). Characterization of anthocyanins in wild *Lycium ruthenicum* Murray by HPLC-DAD/QTOF-MS/MS. Analytical Methods **7,** 4947-4956.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A.,

Scheremetjew, M., Yong, S.-Y., Lopez, R., and Hunter, S. (2014). InterProScan 5: genome-scale protein function classification. Bioinformatics **30,** 1236-1240.

Korf, I. (2004). Gene finding in novel genomes. BMC Bioinformatics **5,** 59.

Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics **15,** 356.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods **9,** 357-U354.

Lin-Wang, K., Bolitho, K., Grafton, K., Kortstee, A., Karunairetnam, S., McGhie, T.K., Espley, R.V., Hellens, R.P., and Allan, A.C. (2010). An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in Rosaceae. BMC Plant Biology **10,** 50.

Liu, Y., Lin-Wang, K., Espley, R.V., Wang, L., Yang, H., Yu, B., Dare, A., Varkonyi-Gasic, E., Wang, J., Zhang, J., Wang, D., and Allan, A.C. (2016). Functional diversification of the potato R2R3 MYB anthocyanin activators AN1, MYBA1, and MYB113 and their interaction with basic helix-loop-helix cofactors. Journal of Experimental Botany **67,** 2159-2176.

Lowe, T.M., and Eddy, S.R. (1999). A computational screen for methylation guide snoRNAs in yeast. Science **283,** 1168-1171.

Mallona, I., Lischewski, S., Weiss, J., Hause, B., and Egea-Cortines, M. (2010). Validation of reference genes for quantitative real-time PCR during leaf and flower development in *Petunia hybrida*. BMC Plant Biology **10,** 11.

Marçais, G., and Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics **27,** 764-770.

Murray, M.G., and Thompson, W.F. (1980). Rapid isolation of high molecular weight plant DNA. Nucleic Acids Research **8,** 4321-4325.

Nakayama, T., Suzuki, H., and Nishino, T. (2003). Anthocyanin acyltransferases: specificities, mechanism, phylogenetics, and applications. Journal of Molecular Catalysis B: Enzymatic **23,** 117-132.

Petit, P., Granier, T., d'Estaintot, B.L., Manigand, C., Bathany, K., Schmitter, J.-M., Lauvergeat, V., Hamdi, S., and Gallois, B. (2007). Crystal structure of grape dihydroflavonol 4-reductase, a key enzyme in flavonoid biosynthesis. Journal of Molecular Biology **368,** 1345-1357.

Quattrocchio, F., Verweij, W., Kroon, A., Spelt, C., Mol, J., and Koes, R. (2006). PH4 of *Petunia* is an R2R3 MYB protein that activates vacuolar acidification through interactions with basic-helix-loop-helix transcription factors of the anthocyanin pathway. The Plant Cell **18,** 1274-1291.

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. Trends in Genetics **16,** 276-277.

Sawada, Y., Nakabayashi, R., Yamada, Y., Suzuki, M., Sato, M., Sakata, A., Akiyama, K., Sakurai, T., Matsuda, F., and Aoki, T. (2012). RIKEN tandem mass spectral database (ReSpect) for phytochemicals: a plant-specific MS/MS-based data resource and database. Phytochemistry **82,** 38-45.

Seitz, C., Ameres, S., and Forkmann, G. (2007). Identification of the molecular basis for the functional difference between flavonoid 3'-hydroxylase and flavonoid 3',5'-hydroxylase. FEBS Letters **581,** 3429-3434.

Sheehan, H., Moser, M., Klahre, U., Esfeld, K., Dell'Olivo, A., Mandel, T., Metzger, S., Vandenbussche, M., Freitas, L., and Kuhlemeier, C. (2016). MYB-FL controls gain and loss of floral UV absorbance, a key trait affecting pollinator preference and reproductive isolation. Nature Genetics **48**.

Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J.M., and Birol, I. (2009). ABySS: a parallel assembler for short read sequence data. Genome Research **19,** 1117-1123.

Slater, G.S.C., and Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics **6,** 31.

Smit, A., Hubley, R., and Green, P. (2015). RepeatMasker Open-4.0.

Smith, C.A., O'Maille, G., Want, E.J., Qin, C., Trauger, S.A., Brandon, T.R., Custodio, D.E., Abagyan, R., and Siuzdak, G. (2005). METLIN: a metabolite mass spectral database. Therapeutic Drug Monitoring **27,** 747-751.

Stanke, M., Steinkamp, R., Waack, S., and Morgenstern, B. (2004). AUGUSTUS: a web server for gene finding in eukaryotes. Nucleic Acids Research **32,** W309-312.

Tatsuzawa, F., Ando, T., Saito, N., Yoda, K., Kokubun, H., Watanabe, H., Hashimoto, G., Asakura, K., Hara, R., and Seki, H. (1999). Acylated malvidin 3-rutinosides in dusky violet flowers of *Petunia integrifolia* subsp *inflata*. Phytochemistry **52,** 351-355.

Wu, T.D., and Watanabe, C.K. (2005). GMAP: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics **21,** 1859-1875.

Zhang, H., Koes, R., Shang, H., Fu, Z., Wang, L., Dong, X., Zhang, J., Passeri, V., Li, Y., Jiang, H., Gao, J., Li, Y., Wang, H., and Quattrocchio, F.M. (2019). Identification and functional analysis of three new anthocyanin R2R3-MYB genes in *Petunia*. Plant Direct **3,** e00114.

Zimmermann, I.M., Heim, M.A., Weisshaar, B., and Uhrig, J.F. (2004). Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like BHLH proteins. The Plant Journal **40,** 22-34.

Züst, T., Strickler, S.R., Powell, A.F., Mabry, M.E., An, H., Mirzaei, M., York, T., Holland, C.K., Kumar, P., Erb, M., Petschenka, G., Gomez, J.M., Perfectti, F., Muller, C., Pires, J.C., Mueller, L.A., and Jander, G. (2020). Independent evolution of ancestral and novel defenses in a genus of toxic plants (*Erysimum*, Brassicaceae). Elife **9**.