Complex petal spot formation in the Beetle Daisy (*Gorteria diffusa*) relies on spot-specific accumulation of malonylated anthocyanin regulated by paralogous GdMYBSG6 transcription factors

Róisín Fattorini[1,2,3] (iD), **Farahnoz N. Khojayori[1]** (iD), **Gregory Mellers[1†]** (iD), **Edwige Moyroud[4,5]** (iD), **Eva Herrero[1]** (iD), **Roman T. Kellenberger[1]** (iD), **Rachel Walker[1]** (iD), **Qi Wang[1]** (iD), **Lionel Hill[6]** (iD) and **Beverley J. Glover[1]** (iD)

[1]Department of Plant Sciences, University of Cambridge, Downing St., Cambridge, CB2 3EA, UK; [2]Department of Biochemistry and Systems Biology, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, L69 7ZB, UK; [3]Department of Biology, University of Oxford, South Parks Road, Oxford, OX1 3RB, UK; [4]Sainsbury Laboratory Cambridge University, Bateman St., Cambridge, CB2 1LR, UK; [5]Department of Genetics, University of Cambridge, Downing St., Cambridge, CB2 3EH, UK; [6]Biomolecular Analysis Facility, John Innes Centre, Colney, Norwich, NR4 7UH, UK

Authors for correspondence:
Beverley J. Glover
Email: bjg26@cam.ac.uk

Róisín Fattorini
Email: roisin.fattorini@biology.ox.ac.uk

Summary

- *Gorteria diffusa* has elaborate petal spots that attract pollinators through sexual deception, but how *G. diffusa* controls spot development is largely unknown. Here, we investigate how pigmentation is regulated during spot formation.
- We determined the anthocyanin composition of *G. diffusa* petals and combined gene expression analysis with protein interaction assays to characterise R2R3-MYBs that likely regulate pigment production in *G. diffusa* petal spots.
- We found that cyanidin 3-glucoside pigments *G. diffusa* ray floret petals. Unlike other petal regions, spots contain a high proportion of malonylated anthocyanin. We identified three subgroup 6 R2R3-MYB transcription factors (GdMYBSG6-1,2,3) that likely activate the production of spot pigmentation. These genes are upregulated in developing spots and induce ectopic anthocyanin production upon heterologous expression in tobacco. Interaction assays suggest that these transcription factors regulate genes encoding three anthocyanin synthesis enzymes.
- We demonstrate that the elaboration of complex spots in *G. diffusa* begins with the accumulation of malonylated pigments at the base of ray floret petals, positively regulated by three paralogous R2R3-MYB transcription factors. Our results indicate that the functional diversification of these *GdMYBSG6s* involved changes in the spatial control of their transcription, and modification of the duration of *GdMYBSG6* gene expression contributes towards floral variation within the species.

Introduction

Pigmentation and structural effects are responsible for the huge variety of flower colouration present in the natural world (Moyroud & Glover, 2017). One function of brightly coloured flowers is to attract animal pollinators (Bradshaw & Schemske, 2003; Hoballah *et al.*, 2005; Sheehan *et al.*, 2012; Papiorek *et al.*, 2016). Many species produce patterned flowers, with stripes on petals, petal spots, or colour differences between floral organs (Moeller, 2005; Eckhart *et al.*, 2006; Gaskett, 2011; Leonard & Papaj, 2011; Shang *et al.*, 2011). Petal spots are composed of pigmented cells that form discrete aggregations, which contrast with the background petal colouration. Spots evolved several times

and are present in multiple plant families including Orchidaceae, Fabaceae, Liliaceae, and Asteraceae (Martins *et al.*, 2013). Petal spots can attract pollinators by forming visual or tactile nectar guides (Leonard & Papaj, 2011), by increasing the conspicuousness of flowers (de Jager *et al.*, 2017), and by prompting aggregation or mating behaviours (Johnson & Midgley, 1997; Ellis & Johnson, 2010).

In many species, anthocyanins, which produce red, pink, purple, and blue colours (Grotewold, 2006; Zhao & Tao, 2015), are the major class of petal spot pigments (e.g. Banba, 1967; Abe *et al.*, 2002; Martins *et al.*, 2013; Yuan *et al.*, 2014). The anthocyanin biosynthesis pathway has been characterised in a broad range of species and begins with chalcone synthase catalysing the formation of tetrahydroxychalcone (Winkel-Shirley, 2001;

Grotewold, 2005, 2006). Downstream enzymes in the core anthocyanin synthesis pathway are encoded by late biosynthesis genes, namely *DIHYDROFLAVONOL 4-REDUCTASE* (*DFR*), *ANTHOCYANIDIN SYNTHASE* (*ANS*), and *UDP-FLAVONOID GLUCOSYL TRANSFERASE* (*UFGT*).

R2R3-MYB transcription factors regulate anthocyanin biosynthesis in many systems (Grotewold, 2006; Schwinn *et al.*, 2006; Albert *et al.*, 2011). These MYBs contain the activation or repression domain at the C-terminus and two MYB repeats (R2 and R3) at the N-terminus, which constitute a binding domain for specific DNA sequences (Dubos *et al.*, 2010). R2R3-MYBs have been categorised into subgroups and the regulation of anthocyanin biosynthesis is strongly associated with subgroups 5, 6, and 7 (Stracke *et al.*, 2001; Feller *et al.*, 2011). Late biosynthesis genes are often regulated by subgroup 6 R2R3-MYB transcription factors that form an MBW complex with a basic helix–loop–helix (bHLH) transcription factor and a WD-repeat (WDR) protein (Spelt *et al.*, 2000; Ramsay & Glover, 2005; Quattrocchio *et al.*, 2006; Schwinn *et al.*, 2006; Stracke *et al.*, 2007; Gonzalez *et al.*, 2008; Dubos *et al.*, 2010; Petroni & Tonelli, 2011).

Petal pattern formation requires adjacent cells to adopt distinct fates, restricting pigment production and accumulation to certain regions of the epidermis (Fairnie *et al.*, 2022). The strict spatial control of anthocyanin synthesis is generally attained through transcriptional regulation. Multiple R2R3-MYB transcription factors often regulate anthocyanin production within a species, for example, in *Antirrhinum majus* (Schwinn *et al.*, 2006) and *Petunia* (Gerats *et al.*, 1984, 1985; Albert *et al.*, 2011, 2014). *Lilium* spp. anthocyanic tepal spots are regulated by LhMYB12-Lat which produces a 'splatter' and LhMYB6 that pigments 'raised' spots (Yamagishi *et al.*, 2010). In *Clarkia gracilis*, *CgMYB1* expression is restricted to a precise region where it activates the expression of *DFR* leading to spot production (Martins *et al.*, 2013).

The sexually deceptive petal spots of *Gorteria diffusa* Thunb. are highly complex, requiring coordinated regulation of multiple pigment pathways and different cell types (Thomas *et al.*, 2009). Thus, they constitute an excellent system to start unravelling the mechanisms regulating the production of sophisticated motifs on petal surfaces. *Gorteria diffusa* is a daisy (Asteraceae) endemic to southern Namibia and South Africa (Roessler, 1959; Duncan & Ellis, 2011). The Asteraceae capitulum is a compressed inflorescence containing a cluster of actinomorphic disc florets (flowers) surrounded by zygomorphic ray florets. In *G. diffusa*, the long ventral petals of the ray florets are fused into a ligule that displays dark anthocyanic petal spots (Karis, 2007; Thomas *et al.*, 2009; Bello *et al.*, 2013). Following quantitative phenotypic investigations, *G. diffusa* populations were assigned to floral morphotypes defined as geographically discrete floral forms identifiable by differences in capitulum phenotypes, including the number of spotted ray florets and spot micromorphology (Ellis & Johnson, 2009).

Gorteria diffusa petal spot types can be loosely categorised into 'simple' dark basal spots present on all ray florets, spots on all ray florets with a raised appearance due to the curvature of ray floret lamina, and raised spots present on one to four ray florets per capitulum. Raised spots are unusually complex compared with those of other daisies and eudicot species. In most floral morphotypes, they are deeply pigmented three-dimensional elaborations of the petal epidermis comprised of multiple specialised epidermal cell types: green interior cells, white highlight cells, and multicellular black papillae (Fig. 1; Thomas *et al.*, 2009).

Gorteria diffusa attracts a bee fly pollinator, *Megapalpus capensis* (Wiedemann), that feeds on pollen and nectar, performs brief inspection visits, or actively attempts to copulate with the petal spots (Johnson & Midgley, 1997; Ellis & Johnson, 2010). These pseudocopulatory responses are only observed on specific morphotypes (Spring, Buffels, and Nieuw) that have raised complex spots on a subset of ray florets (Thomas *et al.*, 2009; Ellis & Johnson, 2010) thought to mimic females, as only male flies attempt copulation (Johnson & Midgley, 1997; Ellis & Johnson, 2010).

In this study, we isolated *G. diffusa* R2R3-MYB transcription factors that promote and spatially restrict anthocyanin accumulation to developing spots. We discovered that malonylated anthocyanins are near exclusive to the spotted region. Of the four R2R3-MYB transcription factors isolated (*GdMYBSG6-1*, *GdMYBSG6-2*, *GdMYBSG6-3*, and *GdMYBSG6-4*), three were likely regulators of petal spot pigment production based on expression data. Their heterologous expression in *Nicotiana tabacum* L. under a strong constitutive promoter was sufficient to activate anthocyanin production. In parallel, we isolated the promoter regions of *ANS*, *DFR*, and *ANTHOCYANIN-3-O-GLUCOSIDE-6″-O-MALONYLTRANSFERASE* (*MAT1*) homologues in *G. diffusa*. We identified GdMYBSG6 binding sites possibly acting as *cis*-regulatory elements in these genomic regions and showed that the *GdDFR* and *GdMAT1* promoters were sufficient to allow some GdMYBSG6 proteins to activate transcription of these late biosynthesis genes. Finally, we suggest evolutionary paths accounting for the functional diversification of *GdMYBSG6* genes and the production of various morphotypes across *G. diffusa* populations.

Materials and Methods

Plant material

Seeds of *G. diffusa* Thunb. and *Gorteria personata* L. were collected in the Northern Cape of South Africa. *Nicotiana tabacum* L. and *Nicotiana benthamiana* Domin. were grown from laboratory lines maintained by selfing. Growth conditions are detailed in Supporting Information Methods S1.

Pigment extraction and quantification

Ray floret petals (from *G. diffusa*) and sepals, petals, and anthers (from *N. tabacum*) were snap-frozen in liquid nitrogen and pulverised in a tissue lyser at 30 Hz for 30 s (Qiagen TissueLyser II). Anthocyanin pigments were extracted using acidic methanol (Methods S2). The anthocyanin levels in the combined supernatants were detected using a spectrophotometer measuring absorbance at the wavelengths A530 and A657 and the overall anthocyanin content per gram of tissue was determined (Methods S2). One petal sample from transgenic *N. tabacum* and three
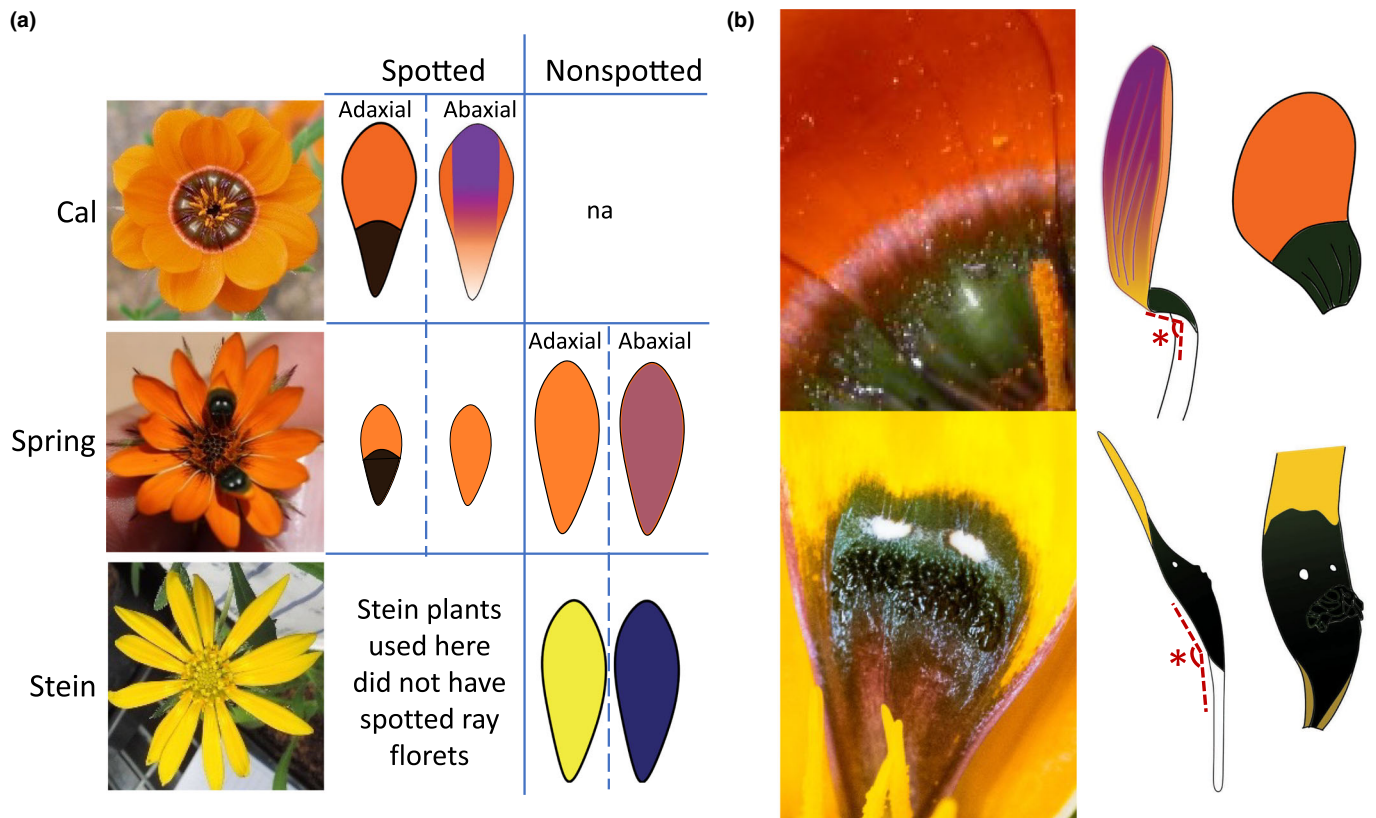
Fig. 1 Ray floret, petal, and spot morphology in the three selected morphotypes (Cal, Spring, and Stein) of *Gorteria diffusa*. (a) Capitulum phenotypes, ray floret types, and typical petal colouration. Cal capitula are 'ringed' with a complex spot forming across the petals of every ray floret. Spring capitula contain one to four ray florets which develop complex spots – these ray florets have a different overall morphology compared with 'nonspotted' Spring ray florets. Some Spring individuals have a small patch of black pigment at the base of all 'nonspotted' ray florets (not shown here). Some Stein individuals develop spots on multiple ray florets, the Stein plants used here did not develop any spots on any capitula. (b) Complex three-dimensional architectures of *G. diffusa* petal spots. Close-up view of Cal (top) and Spring (bottom) petal spots and profile view diagrams of the spotted ray florets. * is the angle of ray floret presentation as defined by Ellis *et al*. (2014).

independent samples for each *G. diffusa* tissue type were analysed through UHPLC-MS/MS at the Biomolecular Analysis Facility at John Innes Centre, Norwich, UK, using a Prominence/Nexera UHPLC system attached to an ion-trap ToF mass spectrometer (Shimadzu, Kyoto, Japan) (Methods S2). MS spectra were collected from *m/z* 220 to 2000 and MS2 spectra from *m/z* 50 to 2000. Compounds were identified as described in Methods S2. Total peak areas for each sample were automatically calculated by the software. The proportion of anthocyanin within the sample accounted for by each compound was calculated using peak areas relative to total peak areas. These proportions were multiplied by the mean total anthocyanin concentration for the identical tissue type determined through spectrophotometry and calculated as described in Methods S2. For subsequent comparisons, the data were grouped according to whether or not malonyl residues were present.

Isolation of candidate genes and promoter regions

Complete gene sequences were isolated from a combination of transcriptomic data (Walker, 2012; Kellenberger *et al.*, 2023) and gene hunting through PCR, 3′ RACE (Ren *et al.*, 2005), and genome walking (detailed in Methods S3, primers in Table S1). All PCRs were conducted with Phusion (New England Biolabs, Ipswich, MA, USA) according to the manufacturer's instructions. Sequencing data were formatted and analysed in GENEIOUS PRIME and BENCHLING (Biology Software) and deposited in GenBank (accession numbers in Methods S3).

Phylogenetic analyses

A gene tree was constructed using the coding sequences of Asteraceae subgroup 6 MYBs. Asteraceae sequences were taken from GenBank using BLAST analysis, the lettuce (*Lactuca sativa* L.) genome (Lettuce Genome Resource, https://lgr.genomecenter. ucdavis.edu/, Reyes-Chin-Wo *et al.*, 2017), the sunflower (*Helianthus annuus* L.) genome (Badouin *et al.*, 2017, https://www.sunflowergenome.org/), and from a literature search (Yue *et al.*, 2018). A subgroup 9 R2R3-MYB, *GdMYBSG9-1* (Thomas, 2009), was included as an outgroup. *Gorteria personata* gDNA was obtained from dried leaf field samples. Introns were predicted and removed from *G. personata* sequences by aligning the gDNA sequences with the corresponding *G. diffusa* coding sequences. The cDNA sequences were aligned with TRANSALIGN

(v.1.2) (Bininda-Emonds, 2005) using an amino acid alignment as a guide. Default parameters were used with one exception: no gaps were removed in the alignment because given the high homology of the sequences the gap regions are informative, rather than introducing noise to the phylogenetic reconstruction. The model of molecular evolution was selected using PARTITIONFINDER (v.2.1.1) (Lanfear et al., 2016) testing all nucleotide models available assuming each of the three codon positions is a data block. The phylogenetic tree was inferred using maximum-likelihood optimality criteria with RAxML-NG (v.1.1.0) (Kozlov et al., 2019) while calculating bootstrap branch support. The phylogenetic tree was visualised using FIGTREE (v.1.4.4).

Gene expression analyses

Gorteria diffusa ray floret tissue was sampled from three independent biological replicates at two developmental stages during spot development, each sample was comprised of multiple plants (Fig. S1). These stages were selected based on the developmental characterisation of the Nieuw morphotype (Thomas et al., 2009). *Nicotiana tabacum* petal tissue was harvested at developmental stage one (Pak Dek et al., 2017; Fig. S2) from each *N. tabacum* transgenic line, pooling tissue from multiple flowers of one T_1 individual. *Nicotiana tabacum* quantitative reverse transcription polymerase chain reaction analyses were conducted on wild-type (WT) plants and transgenic plants.

Quantitative reverse transcription polymerase chain reaction, cDNA synthesis, primer design, reference gene selection, and cycle conditions are detailed in Methods S4. Data were checked and analysed with the OPTICON MONITOR software package (Bio-Rad Laboratories Inc.). Expression levels were calculated relative to the reference gene/s using the modified common base method (Ganger et al., 2017) and primer efficiencies were calculated for each primer pair.

Stable transformation of *N. tabacum*

Nicotiana tabacum cv Samsun was stably transformed with *Agrobacterium tumefaciens* GV3101 transformed with pGREEN-GdMYBSG6 plasmids (Methods S5, S6), containing GdMYBSG6-1, GdMYBSG6-2 or GdMYBSG6-3 under the control of a constitutive double 35S CaMV promoter following a modified version of Horsch et al. (1985) (Methods S5). Several independent insertion lines were recovered for each construct (GdMYBSG6-1 $n = 4$, GdMYBSG6-2 $n = 8$, GdMYBSG6-3 $n = 8$). Quantitative reverse transcription polymerase chain reaction experiments and anthocyanin quantification were conducted on the next generation (T_1) of transformants, alongside WT controls.

Production and purification of recombinant GdMYBSG6 proteins

Gibson assembly (Gibson et al., 2009) was used to produce a pETM11 vector (Dirkmann et al., 2018), containing a sequence encoding a GdMYBSG6-2 protein tagged by six histidines at each end (Methods S7). This plasmid was introduced into *Escherichia coli* (Rosetta II strain). Following incubation, cultures were centrifuged, and bacteria pellets were resuspended in lysis buffer (Methods S8) and sonicated on ice at 20 amps for a duration of 30 s five times. After centrifugation, the protein was purified as detailed in Methods S8. The purified protein was concentrated, and buffer was exchanged into storage buffer (Methods S8) using protein concentrators (Pierce Protein Concentrator PES, 10K MWCO; Thermo Fisher, Waltham, MA, USA) according to the manufacturer's instructions (gel image in Fig. S3). Protein aliquots were snap-frozen in liquid N_2 and stored at $-80°C$.

Electromobility shift assays

Possible MYB binding motifs were identified in *GdANS*, *GdDFR*, and *GdMAT1* upstream sequences (Methods S9). A 29-bp oligonucleotide was designed containing each motif and the promoter region surrounding it and a G nucleotide was added to the 5′ end. This oligonucleotide was ordered along with a 29-bp reverse complement (excluding the added G) from Integrated DNA Technologies (San Diego, CA, USA). Complementary oligos were annealed and fluorescently labelled as detailed in Methods S10.

Purified protein was thawed on ice and diluted 1 : 5 with binding buffer (Methods S8). For the binding reaction, 8–17 µl of diluted protein, 2 µl of 100 ng µl^{-1} fish sperm DNA (Sigma-Aldrich), 1 µl Cy3-dCTP labelled oligos, and enough binding buffer to make a 20 µl reaction were mixed and incubated on ice for 30 min, alongside control reactions containing no protein. Gels (1.8 ml acrylamide (29 : 1), 600 µl 10× TBE, 120 µl 10× (v/v) APS, 12 µl TEMED, and 9.6 ml ddH$_2$O) were prerun for 30 min; then, samples were added and run at 90 V for 60–75 min at 4°C. Gel imaging was completed on a ChemiDoc MP (Bio-Rad Laboratories Inc.).

Yeast one-hybrid experiments

GdANS, *GdDFR*, and *GdMAT1* promoters were divided into overlapping fragments (162–200 bp) amplified from genomic DNA. Two additional fragments were synthesised (Genewiz, Azenta, South Plainfield, NJ, USA), consisting of six tandem repeats of either the *GdANS* promoter segment that was bound by GdMYBSG6-2 in the electromobility shift assay (EMSA) experiments or only the predicted binding motif from this segment. Fragments were integrated into the pHISi plasmid (Clontech, Mountain View, CA, USA) upstream of the *HIS3* reporter through digestion and ligation. Bait plasmids (Methods S11) were linearised using *Apa*I and integrated into the Y1H Gold strain on the *ura3* locus according to the Yeast Protocols Handbook (Clontech). Transformation and colony selection are detailed in Methods S12.

Prey plasmids were obtained by cloning the *GdMYBSG6-1,2,3* coding sequences into the pC-ACT.2 vector plasmids (Methods S11) in frame with the GAL4 activation domain. Prey plasmids were transformed into selected baits as described in the Yeast Protocols Handbook (Clontech). An empty pC-ACT.2 vector and a pC-ACT.2 vector with the coding sequence of Venus fused to

the GAL4 activation domain were used as negative controls. Successful transformations were selected on SD−URA/−LEU plates. Four colonies from each transformation were used as technical replicates for spot tests on SD−URA/−LEU plates and SD−URA/−LEU/−HIS plates containing different 3-AT concentrations (0, 1, 5, 10, 15, and 20 mM). Plates were incubated for 4–7 d at 30°C and photographs were taken.

Dual-luciferase assays

The $35S::GdMYBSG6$-2 pGREEN plasmid, used in stable transformation of *N. tabacum* (Methods S6), was used as an effector plasmid in a dual-luciferase assay. The reporter construct was obtained by cloning GdANSp (324 bp), GdDFRp (378 bp), and GdMAT1p (1016 bp) into the pGreen II 0800-LUC vector (Methods S13) upstream of the firefly-derived luciferase reporter gene (Hellens *et al.*, 2005). The pGreen II 0800-LUC vector also contained a *Renilla*-derived luciferase reporter under the control of the CaMV 35S promoter to normalise the values of the experimental reporter gene for variation caused by transfection efficiency.

Transformed *A. tumefaciens* GV3101 were cultured and prepared for infiltration as detailed in Methods S14. Cultures were infiltrated into the abaxial surface of 4-wk-old *N. benthamiana* young leaves. Four infiltrations per leaf were completed, with media containing a single reporter plasmid with the GdDFR promoter region, plus an effector plasmid encoding GdMYBSG6-1, 2, or 3, or an empty pGREEN as a control. For each combination, an injected leaf on a separate plant was treated as a biological replicate. Leaves were harvested 48 h after infiltration, and luciferase activity was measured immediately using the Dual-Glo Luciferase Assay Kit (Promega). Leaf discs that were 4 mm in diameter were cut and inserted into white 96-well plates with 75 µl of 1× PBS and 75 µl of luciferase assay reagent. Firefly luminescence was measured on the ClariostarPlus (BMG Lab Technologies, Ortenberg, Germany) after 15 min. Afterwards, 75 µl of Stop & Glo reagent was added to the samples, samples were incubated for 15 min, and the *Renilla* luminescence was measured as an internal control. Luminescence values from a control well on each plate (leaves infiltrated with no effector plasmid present) were subtracted from each reading. The results were expressed as the ratio of firefly to *Renilla* luciferase activity and normalised across plates.

Data analysis including the creation of graphs and statistical tests are detailed in Methods S15.

Results

Ray floret morphology in the *G. diffusa* morphotypes Calendula, Springbok, and Steinkopf

To capture the range of natural variation, Calendula (Cal), Springbok (Spring), and Steinkopf (Stein) morphotypes were used in this study (Fig. 1a). In Cal, each ray floret produces a spot creating a ringed capitulum (Fig. 1a). In Spring, one to four non-adjacent ray florets develop spots, varying between capitula within an individual. Spring ray florets lacking spots are larger and can have different background colouration than spotted counterparts (Figs 1a, S4). Within some Spring individuals, the nonspotted ray florets display a basal patch of dark pigment, or 'mark', with no cellular elaboration (Fig. S4). Spots sometimes develop on a subset of ray florets in certain Stein individuals, but plants used here did not have any spotted capitula (Fig. 1a).

Cal and Spring spots are vibrantly coloured and textured creating a complex and shimmering display. Cal spot colouration is overall deep green, while Spring spots appear black/purple/green (Fig. 1b). The 'angle of ray floret presentation' (i.e. the shape of the spotted ray florets) contributes more to the raised appearance of spotted ray florets in Cal compared with Spring (Fig. 1b; Ellis *et al.*, 2014; Mellers, 2016). In glasshouse and wild environments, the extent of abaxial petal purple/black pigmentation varies between plants within a morphotype (Fig. S5) and is particularly variable in Spring. Purple/black abaxial pigmentation in Cal and Stein is consistently darker and more extensive than that on nonspotted ray floret petals of Spring. Dark pigmentation was consistently absent from the abaxial side of spotted ray florets in Spring and spotted portions of Cal ray florets (Fig. S5). Only petal regions exposed when capitula close have abaxial anthocyanin pigmentation, possibly reducing flower conspicuousness as an anti-herbivore mechanism (Kemp & Ellis, 2019).

Cyanidin pigments ray floret petals and malonylated cyanidins are found almost exclusively in petal spots

We found that Spring spots had a higher anthocyanin content than any other tissue type (Fig. 2b; Table S2). Cal spots had equivalent anthocyanin concentrations to Cal and Stein nonspotted petal tissue, which have dark abaxial pigmentation. The nonspotted part of Cal ray floret petals contained significantly more anthocyanin than Spring nonspotted ray floret petals ($P = 0.005$), potentially due to greater abaxial pigmentation in Cal.

UHPLC-MS/MS analyses were conducted to identify anthocyanins in the three morphotypes. For every sample, a major peak was detected at 5.059 min in the chromatograms showing UV absorbance. For the nonspotted regions, this was the only major peak (Table S2; Fig. S6). The mass of the compound was 449 and fragmentation of the [M] + ion was induced (using collision-induced dissociation). This fragmentation at *m/z* 449 formed a base peak (the most intense peak, representing the ion with the greatest relevant abundance) at *m/z* 287, corresponding to aglycone cyanidin and a mass loss of 162 suggests loss of a glucose moiety (Table S3; Fig. S7). Cyanidin 3-glucoside was used as a reference compound by Schütz *et al.* (2006) during UHPLC-MS/MS analyses on acidic methanol extractions from *Cynara scolymus* L. (Asteraceae). The MS-MS analysis results of their study were identical to the major peak here identified; thus, we tentatively identify our major peak as cyanidin 3-glucoside.

Two additional major peaks were present in the UV absorbance chromatograms of all petal samples from spotted or marked regions (Cal spot, Spring spot, and Spring mark). The
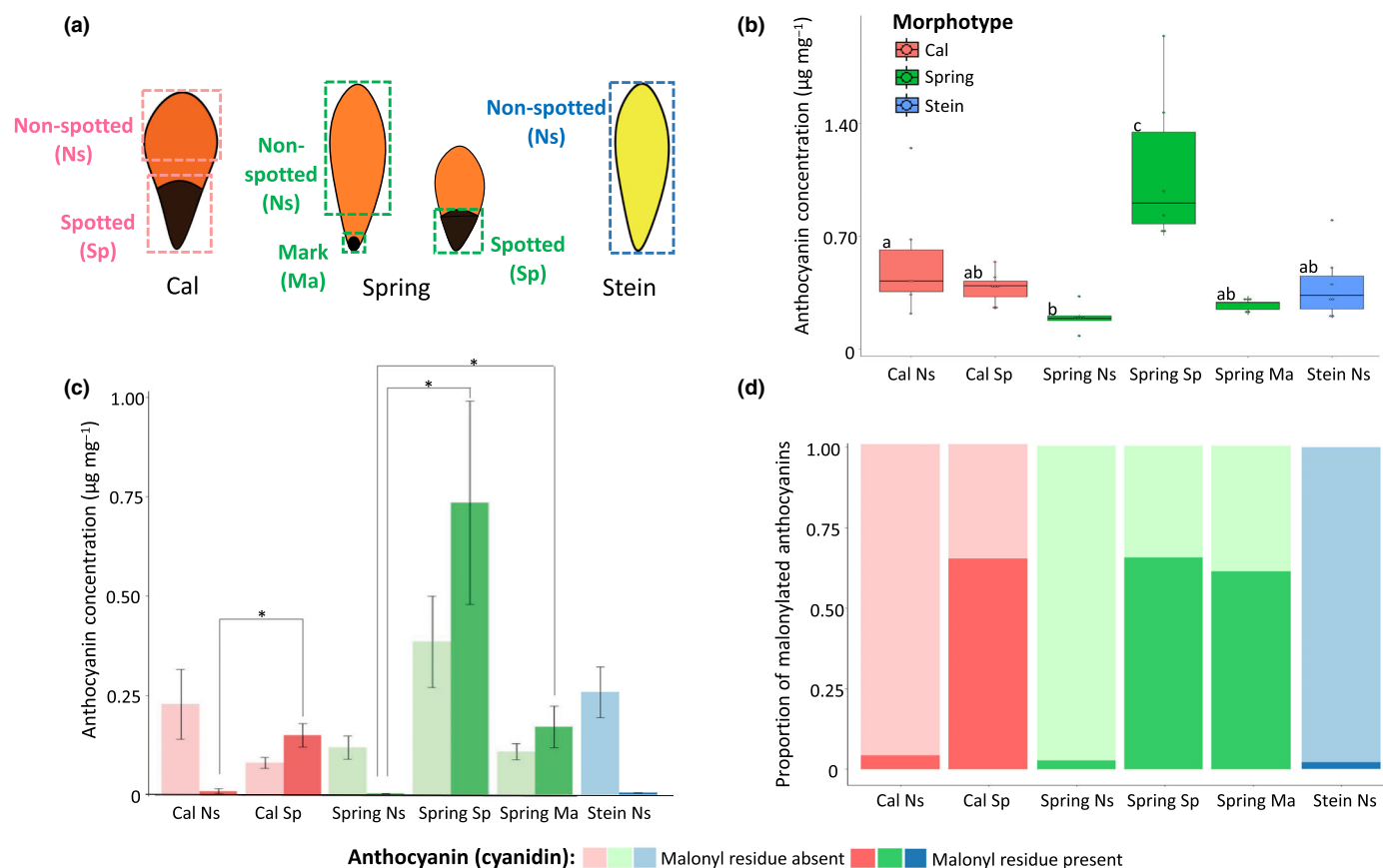
Fig. 2 Anthocyanin content of *Gorteria diffusa* ray florets in the morphotypes Cal, Spring, and (nonspotted) Stein. (a) Schematics of typical ray florets from each morphotype indicating regions used for pigment extraction (dashed boxes). (b) Anthocyanin content (µg cyanindin-3-glucoside equivalent per mg of fresh tissue) for tissue types depicted in (a). Individual data points are represented by black dots. Sample size $n = 5$–7, where n represents pooled tissue from a single individual. The black line in each box indicates the median value, and the whiskers represent the 25/75% quartile $\pm$ 1.5 $\times$ interquartile range. Boxes that do not share letters are significantly different from one another ($P \leq 0.05$, linear mixed model and Tukey's tests). (c) Summary high performance liquid chromatography-mass spectrometry analysis results. Approximate anthocyanin content for each tissue type is shown, grouped according to whether a malonate residue is present or absent. Sample size $n = 3$, where n represents pooled tissue from a single individual, error bars are $\pm$SD. *, $P < 0.0001$ (linear mixed model and Tukey's tests). (d) Proportion of anthocyanin for each sample type that contains a malonyl residue. For (c) and (d) presence or absence of a malonyl residue is colour-coded according to the key below the graphs.

retention times of these peaks were 7.133 and 8.283 min, and both had m/z 287 and m/z 449 consistent with a cyanidin gluco-side (Fig. S7; Table S3). Compound 4 (Table S3) had a mass loss of 44 (mass of 535 and a peak at m/z 491) corresponding to a likely loss of carbon dioxide from the terminal carboxylic acid group of a malonate. Compound 5 (Table S3) had a mass of 549 and a peak at m/z 517; the mass loss of 100 is appropriate for a methylmalonate. Additional minor peaks were occasionally detected in the UV absorbance chromatogram (Table S3). All peaks identified were cyanidin (although characterisation was inconclusive in one case) and contained a glucose moiety, except for one cyanidin that had a pentose sugar moiety.

Across morphotypes, concentrations of malonylated cyanidins were $\geq$ 12-fold higher in spotted compared with nonspotted petal tissue (Fig. 2c). Anthocyanins containing a malonyl residue accounted for 65% of anthocyanins in Spring and Cal spotted petal tissue and 61% in the Spring mark (Fig. 2d). Malonylated anthocyanins formed $\leq$ 6% of nonspotted petal tissue

anthocyanins. Anthocyanin composition of the Spring mark was similar to that of Cal and Spring complex spots (Table S2).

Isolation of *G. diffusa* R2R3-MYB candidate regulators of petal spot anthocyanin

Four R2R3-MYB transcription factors were identified as candidates for regulating petal spot pigmentation in Spring. *GdMYBSG6-1* and *GdMYBSG6-2* were found to be upregulated in spotted petal tissue within a *G. diffusa* transcriptome (Walker, 2012). The *GdMYBSG6-3* paralogue was isolated through gene hunting via PCR while a second transcriptome (Kellenberger *et al.*, 2023) recovered *GdMYBSG6-4*. All four genes encode the R2R3 domains and amino acid motif typical of subgroup 6 R2R3-MYBs (Fig. 3a). GdMYBSG6-1, GdMYBSG6-2, and GdMYBSG6-3 have 84–91% amino acid conservation. GdMYBSG6-4 is the most divergent paralogue with 68–69% sequence homology with GdMYBSG6-1,2,3
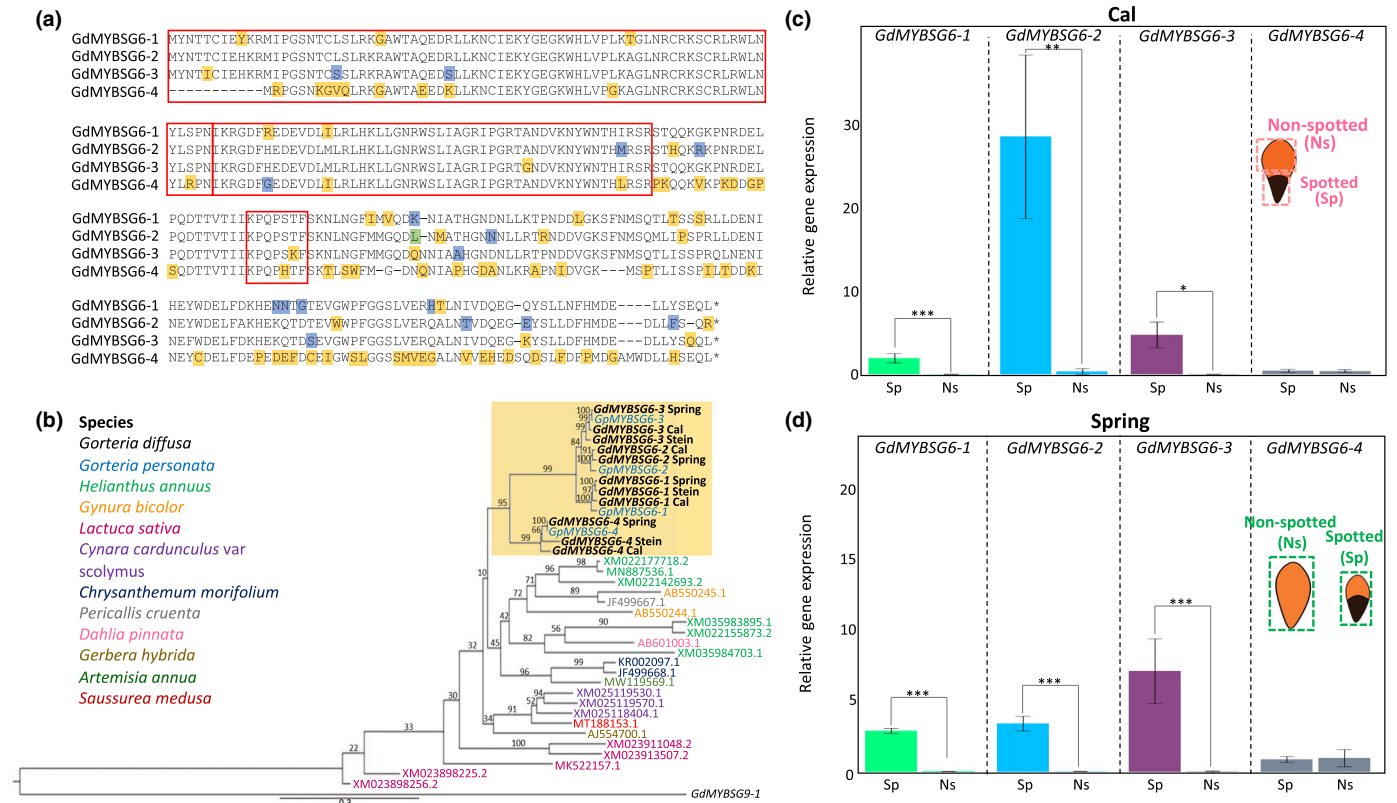
Fig. 3 *GdMYBSG6* genes are within subgroup 6 of the R2R3 MYB transcription factors and expressed within *Gorteria diffusa* ray floret petals. (a) Protein alignment of GdMYBSG6 paralogues from the *G. diffusa* Spring morphotype. The red boxes indicate the R2 and R3 domains, and the subgroup 6 motif, respectively. (b) Maximum-likelihood phylogeny of Asteraceae subgroup 6 R2R3 MYB cDNA sequences. The tree was rooted with *GdMYBSG9-1*, a gene encoding a *G. diffusa* subgroup 9 R2R3 MYB used as an outgroup (Thomas, 2009). Bootstrap values are given above each branch. *Gorteria MYBSG6* genes sequenced during this project are indicated by the yellow box. Accession numbers are given, and corresponding species are colour-coded as indicated in the key. (c) Quantitative reverse transcription polymerase chain reaction results showing the relative expression levels of *GdMYBSG6-1,2,3,4* in *G. diffusa* morphotypes Cal and (d) Spring. Cal spotted (Sp) and nonspotted (Ns) ray floret petal tissue and Spring whole spotted ray floret petals (Sp) and whole nonspotted ray floret petals (Ns) were sampled when spot formation initiates, as a small patch of pigment becomes visible. Error bars represent the mean $\pm$ SE of three biological replicates. Gene expression is relative to the reference gene *G. diffusa Elongation Factor 2* (*GdEF-2*). *, $P < 0.01$; **, $P < 0.001$; ***, $P < 0.0001$ (ANOVA and false discovery rate method).

(Fig. 3a). *GdMYBSG6* genes were highly structurally conserved in Cal and Stein morphotypes, including in length. Between morphotypes, intron 2 and exon 3 for all four genes vary in length, and the length of *GdMYBSG6-4* intron 1 also differs. Only a small section of *GdMYBSG6-2* was isolated in Stein, so this gene was excluded from comparisons.

Phylogenetic analysis confirms that *GdMYBSG6-1–4* encode subgroup 6 R2R3-MYBs

A maximum-likelihood phylogenetic reconstruction (Fig. 3b) confirmed that *Gorteria MYBSG6* genes clustere within subgroup 6, forming a single clade with high bootstrap support (95%). *Gorteria personata MYBSG6* sequences clustered with the corresponding *GdMYBSG6* orthologues, indicating that duplications occurred before the speciation event that yielded *G. diffusa* and *G. personata*. Within the *Gorteria* clade, *MYBSG6-4* diverged first (bootstrap value 99%). The remaining *MYBSG6* genes form a clade with *MYBSG6-2* and *MYBSG6-3* sister to one another (bootstrap value 84%).

Three *GdMYBSG6* paralogues are upregulated in developing petal spots

We set out to identify which *GdMYBSG6* genes were most likely to regulate pigment production during spot formation. The expression levels of each gene in spotted and nonspotted ray floret petal tissue were quantified across morphotypes (Fig. S1) during spot initiation (Fig. 3c,d), when a small patch of dark pigment is just visible, and later (Figs S8, S9) as anthocyanins accumulate and specialised cell types mature.

During spot initiation, *GdMYBSG6-4* had low uniform expression in spotted and nonspotted petal tissue in all morphotypes (Fig. 3c,d). Its expression significantly increased later in both spotted ($P < 0.001$) and nonspotted ($P < 0.001$) petal tissues of Cal (Fig. S9), and in Stein ($P = 0.042$) (Fig. S8). Thus, *GdMYBSG6-4* expression patterns do not correlate with spot pigment accumulation, suggesting it is unlikely to regulate spot anthocyanin production.

By contrast, *GdMYBSG6-1*, *GdMYBSG6-2*, and *GdMYBSG6-3* were more highly expressed in spotted tissue than nonspotted
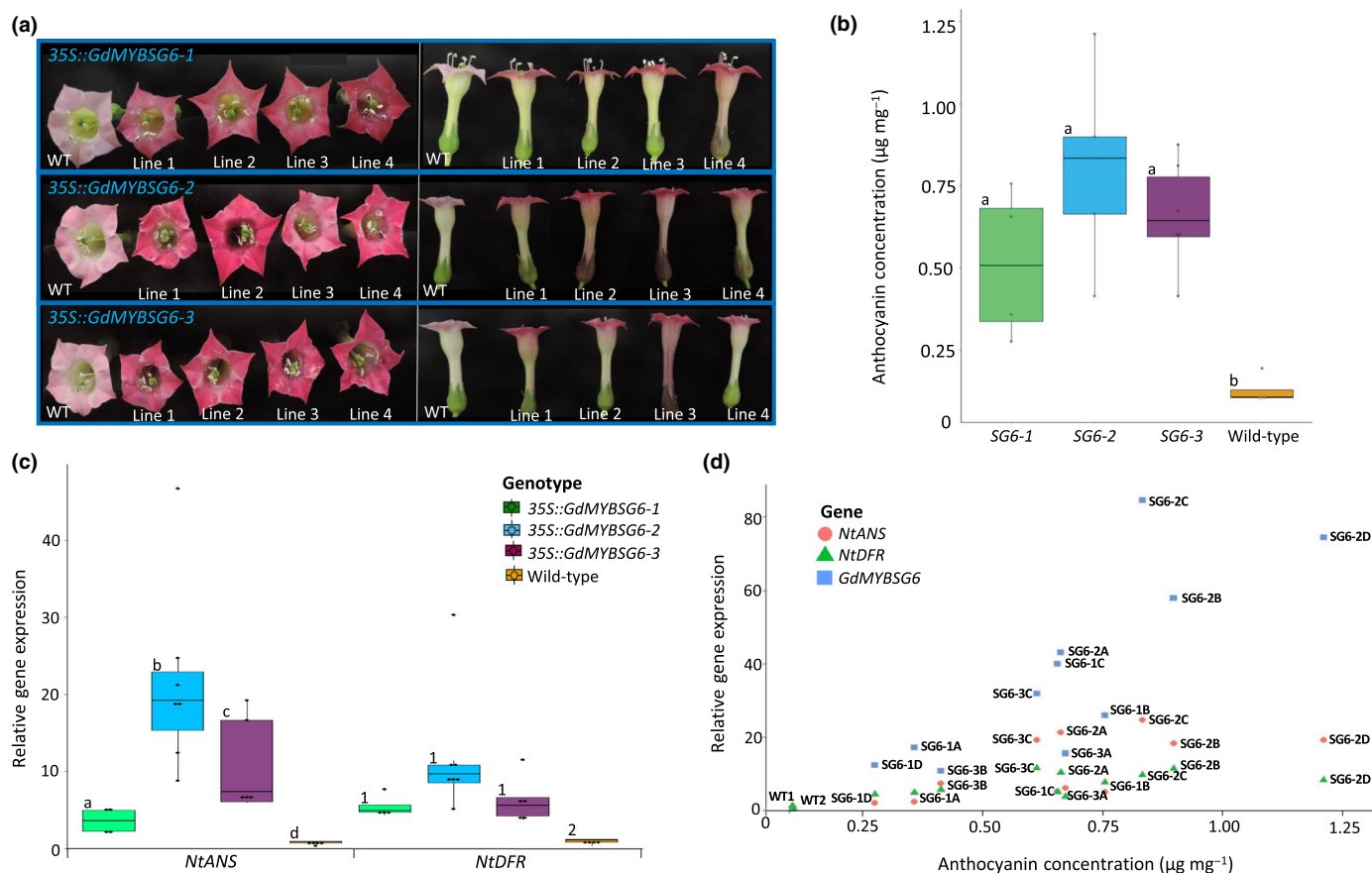
Fig. 4 Anthocyanin phenotypes and gene expression levels in T1 *Nicotiana tabacum* plants stably transformed with either *GdMYBSG6-1*, *GdMYBSG6-2*, or *GdMYBSG6-3* under the control of a strong constitutive promoter (35S). (a) Flowers from several independent lines carrying each construct. Wild-type (WT) flowers are labelled WT. (b) Anthocyanin concentration within petals. Each data point (black dot) is from a different independent line and represents the mean anthocyanin content of four to six samples. (c) Relative gene expression levels of *ANTHOCYANIDIN SYNTHASE* (*ANS*) and *DIHYDROFLAVONOL-4-REDUCTASE* (*DFR*) in the petals of *N. tabacum* transformants and WT plants. Each data point (black dot) is from a different independent line and sample size is $n = 4$–7. (d) Correlation between petal anthocyanin content and expression levels of genes of interest (*NtANS*, *NtDFR*, and *GdMYBSG6-1,2,3*) in individual plants from a subset of *N. tabacum* lines carrying each *35S::GdMYBSG6* construct. The line and construct are listed next to each data point (e.g. *35S::GdMYBSG6-1* line A = 'SG6-1A'). Quantitative reverse transcription polymerase chain reaction expression levels are shape and colour-coded according to which gene they correspond to (see key). Within a graph, tissues with statistically significant differences ($P \leq 0.05$) in expression levels or anthocyanin concentrations do not share a letter/number (ANOVA and Tukey's test/false discovery rate method). In boxplots, the black line in each box indicates the median value and the whiskers represent the 25/75% quartile $\pm$ 1.5 $\times$ interquartile range.

counterparts in Cal (*GdMYBSG6-1* $P < 0.001$, 137-fold difference; *GdMYBSG6-2* $P = 0.001$, 84-fold difference; *GdMYBSG6-3* $P = 0.002$, 292-fold difference) and Spring (*GdMYBSG6-1*: $P < 0.001$, 92-fold difference; *GdMYBSG6-2*: $P < 0.001$, 154-fold difference; *GdMYBSG6-3*: $P < 0.001$, 268-fold difference). Expression levels of *GdMYBSG6-2* in Cal spotted tissue were significantly higher than those of *GdMYBSG6-1* ($P < 0.001$) and *GdMYBSG6-3* ($P < 0.001$), while *GdMYBSG6-3* had the greatest expression levels in Spring (*GdMYBSG6-1* $P = 0.009$; *GdMYBSG6-2* $P = 0.027$) (Fig. 3c,d). *GdMYBSG6*-1,2,3 expression was extremely low or nondetectable in nonspotted petal tissue of Spring and Cal. Consistent with this, *GdMYBSG6-1* and *GdMYBSG6-3* expression levels were also very low in Stein nonspotted ray florets (Fig. S8). At the later stage, *GdMYBSG6-1* and *GdMYBSG6-3* remained preferentially expressed in spotted tissues of Cal and Spring, and in Spring their expression levels decreased. The expression level of *GdMYBSG6-2* in spotted tissues, however,

increased significantly at the second developmental stage in Spring and Cal (Fig. S9). Thus, these paralogues could act redundantly to promote anthocyanin production during spot formation, with GdMYBSG6-2 likely the main driver.

GdMYBSG6-1, 2, and 3 can induce anthocyanin production in tobacco

Stable transformations of tobacco plants (*N. tabacum*) were used to assess whether the GdMYBSG6 proteins could induce anthocyanin synthesis in a heterologous context (Fig. S10). Data were collected from the T1 generation corresponding to several (≥ 4) independent primary transformants for each transgene.

Constitutive expression of *GdMYBSG6-1*, *GdMYBSG6-2*, or *GdMYBSG6-3* was sufficient to induce anthocyanin synthesis within all transgenic lines. Petal tissue of all transformants had stronger anthocyanin phenotypes than WT plants (Fig. 4a).

Anthocyanin was ectopically produced within sepals, anthers, and leaves in at least a single line from each of *GdMYBSG6-1*, *GdMYBSG6-2*, and *GdMYBSG6-3* transformants (Figs 4a, S11). Transformed plants had significantly higher anthocyanin concentrations in petals (*35S::GdMYBSG6-1* $P = 0.004$; *35S::GdMYBSG6-2* $P < 0.001$; *35S::GdMYBSG6-3* $P < 0.001$) and anthers (*35S::GdMYBSG6-1* $P = 0.041$; *35S::GdMYBSG6-2* $P = 0.001$; *35S::GdMYBSG6-3* $P = 0.003$) compared with WT plants. Sepal anthocyanin content was also greater in *35S::GdMYBSG6-2* ($P < 0.001$) and *35S::GdMYBSG6-3* ($P = 0.012$) individuals compared with WT (Fig. S12). UHPLC-MS/MS analyses revealed no differences in anthocyanin composition compared with WT (Figs S13, S14).

The expression levels of *NtANS* and *NtDFR* were investigated through quantitative reverse transcription polymerase chain reaction to determine whether expression patterns were consistent with GdMYBSG6 transcriptional activation (*35S::GdMYBSG6-1* $n = 4$, *35S::GdMYBSG6-2* $n = 7$, *35S::GdMYBSG6-3* $n = 5$). Transgene expression was generally higher across *35S::GdMYBSG6-2* lines (Fig. S15). There was significant upregulation of *NtANS* (*35S::GdMYBSG6-1* $P = 0.03$; *35S::GdMYBSG6-2* $P < 0.001$; *35S::GdMYBSG6-3* $P < 0.001$) and *NtDFR* (*35S::GdMYBSG6-1* $P < 0.001$; *35S::GdMYBSG6-2* $P < 0.001$; *35S::GdMYBSG6-3* $P < 0.001$) compared with WT tobacco. *GdMYBSG6* expression was positively correlated with *NtANS* expression ($P = 0.001$, $R^2 = 0.77$) and *NtDFR* expression ($P < 0.001$, $R^2 = 0.50$) across transgenic lines (Fig. 4d).

The upregulation of *NtANS* in *35S::GdMYBSG6-1* transgenic lines was lower than in *35S::GdMYBSG6-2* ($P = 0.001$) and *35S::GdMYBSG6-3* ($P = 0.040$) transformants, even when variability resulting from transgene expression levels was accounted for in the statistical model. Despite similarities between certain lines regarding transgene expression levels and anthocyanin content, transgenic lines overexpressing *GdMYBSG6-1* had lower *NtANS* expression levels than *35S::GdMYBSG6-2* and *35S::GdMYBSG6-3* lines (Fig. 4c). Thus *GdMYBSG6-1–3* encode proteins with similar but not identical biochemical properties; further comparisons would be required to address this comprehensively.

Gorteria diffusa genes encoding anthocyanin synthesis enzymes are upregulated in developing petal spots

To test whether *ANS* and *DFR* were also regulatory targets of GdMYBSG6 proteins in *Gorteria*, *G. diffusa ANS* and *DFR* homologues were isolated. Single copies of *GdANS* and *GdDFR* were found, and copy number was later confirmed through an exhaustive BLAST search of the newly available *G. diffusa* genome (R. T. Kellenberger & B. J. Glover, unpublished). A malonyl transferase gene (*GdMAT1*) was also investigated because malonyl transferase (MAT) catalyses the addition of malonyl groups to anthocyanins (Nakayama *et al.*, 2003) and *G. diffusa* petal spot anthocyanins are largely malonylated (Fig. 2). Four *G. diffusa* malonyl transferase variants were recently identified (Kellenberger *et al.*, 2023) and BLAST searches confirmed that the *G. diffusa* genome contains three *MAT* genes (R. T. Kellenberger

& B. J. Glover, unpublished). All contained the motif associated with the subfamily for anthocyanin malonyl transferases (Unno *et al.*, 2007), but only one (*GdMAT1*) was upregulated in spotted tissue (Kellenberger *et al.*, 2023).

If *GdANS*, *GdDFR*, and *GdMAT1* are regulatory targets of GdMYBSG6s, their expression profiles are likely to follow those of *GdMYBSG6-1,2,3*. To test this, we characterised their expression patterns in the same tissue samples as those used to establish the transcriptional dynamics of *GdMYBSG6* genes. As spot formation initiates, *GdANS*, *GdDFR*, and *GdMAT1* are significantly upregulated in spotted compared with nonspotted petal tissue in Cal (*GdANS*: $P = 0.002$, 15-fold difference; *GdDFR*: $P < 0.001$, 53-fold difference, *GdMAT1*: $P = 0.003$, 46-fold difference) and Spring (*GdANS*: $P < 0.001$, 18-fold difference; *GdDFR*: $P < 0.001$, 217-fold difference, *GdMAT1*: $P = 0.002$, 9-fold difference) (Fig. 5a). At the later stage, expression levels of both *GdANS* and *GdDFR* significantly increased in nonspotted petal tissue of Cal (*GdANS*: $P < 0.001$, 66-fold increase; *GdDFR*: $P < 0.001$, 311-fold increase) and Spring (*GdANS*: $P < 0.0001$, 45-fold increase; *GdDFR*: $P < 0.0001$, 443-fold increase) coinciding with the production of anthocyanin pigmentation on the abaxial side of the ray floret petals (Figs S5, S16). *GdMAT1* was also upregulated in the spotted region at the later developmental stage in both Cal ($P = 0.01$) and Spring ($P = 0.005$) compared with nonspotted tissue, with expression in the latter remaining very low (Fig. S16). In nonspotted Stein petals, *GdMAT1* expression was very low and nondetectable for some biological replicates (mean ± SE: dev 1 0.01 ± 0.004, dev 2 0.012 ± 0.01).

GdMYBSG6-1,2,3 can interact with promoter regions of *GdDFR*, *GdANS*, and *GdMAT1*

To test whether GdMYBSG6-1,2,3 could act as direct activators of selected late biosynthesis genes, we isolated regions immediately upstream of the start codon of *GdMAT1* (1016 bp), *GdANS* (324 bp), and *GdDFR* (378 bp) and used complementary approaches to evaluate the ability of GdMYBSG6-1,2,3 to bind those regions and activate transcription.

As our expression data singled out GdMYBSG6-2 as the top candidate to regulate anthocyanin synthesis during spot development, we tested its ability to bind to predicted motifs *in vitro*. EMSA results showed that GdMYBSG6-2 can bind specific DNA motifs within the upstream regions of *GdANS*, *GdDFR*, and *GdMAT1* (Fig. 5b). Both motifs tested from *GdMAT1p* were bound by GdMYBSG6-2 (Fig. 5b). Out of three motifs tested in each of *GdANSp* and *GdDFRp*, GdMYBSG6-2 consistently bound to one motif in each upstream region (Fig. 5b). These motifs contained the sequence TTGAATG – previously identified by Wang *et al.* (2013) in the *Gerbera DFR2* promoter. Substituting 'AA' with 'GG' in the centre of these motifs, and the *GdMAT1* motifs, was sufficient to reduce or even prevent GdMYBSG6-2 binding in all cases.

Next, we used yeast one-hybrid (Y1H) experiments to assess the ability of GdMYBSG6s to bind *in vivo* to upstream regions of *GdANS*, *GdDFR*, and *GdMAT1*. All three GdMYBSG6s could bind *GdANSp* fragment 1 (*GdANSp*-1) and both *GdDFRp*
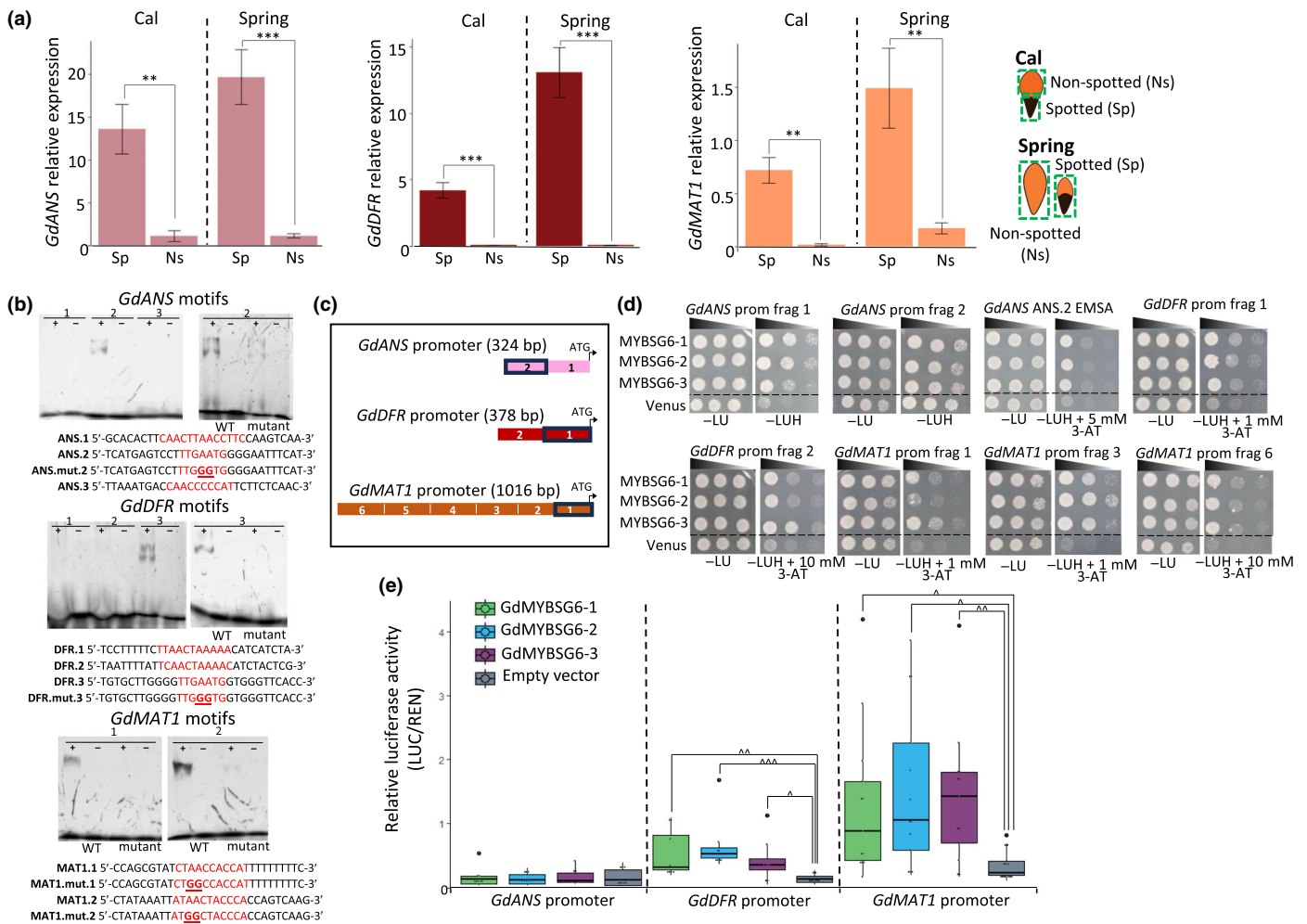
Fig. 5 GdMYBSG6 proteins are likely transcriptional regulators of genes encoding late anthocyanin synthesis enzymes in *Gorteria diffusa*. (a) Relative expression of *G. diffusa ANTHOCYANIDIN SYNTHASE* (*GdANS*), *DIHYDROFLAVONOL 4-REDUCTASE* (*GdDFR*) and *ANTHOCYANIN-3-O-GLUCOSIDE-6″-O-MALONYLTRANSFERASE 1* (*GdMAT1*) in spotted (Sp) and nonspotted (Ns) ray floret petal tissue from Cal and Spring morphotypes during spot initiation. Error bars represent the mean ± SE of three biological replicates. Gene expression is relative to the reference gene *G. diffusa Elongation Factor 2* (*GdEF-2*); **, $P \leq 0.02$; ***, $P \leq 0.0001$ (ANOVA and false discovery rate method). (b) Electromobility shift assays (EMSAs) testing whether GdMYBSG6-2 can bind to wild-type (WT) and mutated versions of selected DNA motifs from promoter regions of *GdANS*, *GdDFR*, and *GdMAT1*. Negative controls (−) consist of binding reactions with no protein added. DNA motif sequences are given under each gel, with predicted binding sites highlighted in red and changes made in mutated motifs indicated in underlined bold. The location of motifs which GdMYBSG6-2 binds to is indicated by black squares in the diagram in (c). (c) Diagram of promoter regions used in dual-luciferase and yeast one-hybrid (Y1H) experiments, the numbers along each promoter sequence correspond to the fragment numbers in (d). (d) Photographs of Y1H experiments. Each row of three is a dilution series (OD 1.5, 10-fold dilution, 100-fold dilution) of the same yeast colony transformed with either one of the *GdMYBSG6-1,2,3* genes or the negative control Venus. *GdANS* ANS.2 is a fragment containing six tandem repeats of the *GdANS* motif that GdMYBSG6-2 bound to in EMSA experiments. All colonies grow on −LEU −URA plates (left-hand side), with −LEU −URA −HIS plates +/−3-AT (right-hand side) indicating whether each GdMYBSG6 protein could bind to each promoter fragment. (e) Relative luciferase activity indicating whether GdMYBSG6-1, GdMYBSG6-2, and GdMYBSG6-3 can activate luciferase expression driven by the promoters of *GdANS*, *GdDFR*, and *GdMAT1*, with the empty effector plasmid used as a negative control. Individual data points are represented by black dots. The black line in each box indicates the median value and the whiskers represent the 25/75% quartile ± 1.5 × interquartile range. Sample size $n = 8$–13; ^, $P \leq 0.01$; ^^, $P \leq 0.001$; ^^^, $P \leq 0.0001$ (ANOVA and Tukey's test).

fragments, as well as *GdMAT1p*-1, *GdMAT1p*-3, and *GdMAT1p*-6 (Figs 5c,d, S17). While binding patterns were consistent, the strength of binding differed between paralogues for some promoter fragments. GdMYBSG6-3, for example, appears to have weaker binding to *GdANSp*-1 than GdMYBSG6-1 and GdMYBSG6-2 (Figs 5d, S17). *GdANSp*-2 contained the *GdANS* sequence that resulted in *in vitro* binding, but no binding to *GdANSp*-2 occurred within yeast. However, GdMYBSG6-1,2,3

did bind to a fragment consisting of tandem repeats of this positive EMSA sequence (Fig. 5d).

Finally, we tested whether *GdMYBSG6-1*, *GdMYBSG6-2*, or *GdMYBSG6-3* could activate the transcription of *GdANS*, *GdDFR*, and *GdMAT1 in planta* by dual-luciferase assay. We found that GdMYBSG6-1,2,3 could significantly induce expression of luciferase in tobacco leaves when driven by *GdDFRp* (GdMYBSG6-1: $P = 0.005$, 4.5-fold difference; GdMYBSG6-2:

$P = 0.0005$, 5.6-fold difference; GdMYBSG6-3: $P = 0.050$, 3.5-fold difference) or *GdMAT1p* (GdMYBSG6-1: $P = 0.025$, 3.8-fold difference; GdMYBSG6-2: $P = 0.005$, 4.6-fold difference; GdMYBSG6-3: $P = 0.008$, 4.1-fold difference) (Fig. 5e). No GdMYBSG6 protein could activate transcription from *GdANSp*.

Taken together, these experiments demonstrate that GdMYBSG6-1, GdMYBSG6-2, and GdMYBSG6-3 can bind to *GdDFR* and *GdMAT1* promoter regions and activate transcription in heterologous systems. Yeast one-hybrid results suggest additional binding motifs not tested in EMSA experiments exist in *GdDFRp*-2, *GdMAT1p*-3, and *GdMAT1p*-6. A binding motif for GdMYBSG6 was validated *in vitro* for *GdANSp* but the fragment containing this motif, *GdANSp*-2, did not lead to yeast reporter gene expression. Yeast one-hybrid experiments, however, demonstrated that all three GdMYBSG6 proteins can bind to the other *GdANSp* fragment (*GdANSp*-1), suggesting this fragment likely contains additional binding motifs not investigated through EMSAs. Whether those motifs are *bona fide* regulatory sites is not clear, as no GdMYBSG6 protein could induce luciferase expression driven by *GdANSp in planta*.

Discussion

Gorteria diffusa exhibits high levels of geographically defined floral variation within a narrow endemic range in the Succulent Karoo biodiversity hotspot of South Africa. Key components of this intraspecific diversity are unusually elaborate petal spots. Richly coloured and deeply textured, these spots are comprised of several specialised cell types that create three-dimensional elaborations and varying colours. Here, we investigated pigment production, one feature underpinning complex spot formation. We found that malonylated cyanidin 3-glucoside almost exclusively accumulates in *G. diffusa* spots and identified three paralogous *GdMYBSG6* genes likely to activate spot-specific anthocyanin production through transcriptional control of genes encoding late anthocyanin biosynthetic enzymes.

Morphotypes produce similar petal anthocyanins, but concentrations differ in the spot

Cultivar differences in anthocyanin composition have been reported in Asteraceae, including *Dahlia variabilis* and *Gerbera jamesonii* (Takeda *et al.*, 1986). Across the *G. diffusa* morphotypes examined, however, anthocyanin composition was largely consistent. Cyanidin 3-glucoside was predominant, and all anthocyanins identified were derived from cyanidin (with a single unconfirmed exception). Cyanidin derivatives contribute towards pink to red floral colouration, in *Chrysanthemum morifolium* and *Lilium* spp., and blue colouration in *Meconopsis grandis* and *Centaurea cyanus* (Yoshida *et al.*, 2006; Yoshida & Negishi, 2013; Hong *et al.*, 2015; Suzuki *et al.*, 2016).

Spring petal spots contained significantly more anthocyanin than those of Cal, perhaps due to differences in spot micromorphology. Papillae are groups of swollen epidermal cells that are larger, darker, and more densely distributed in the Spring morphotype compared with Cal. The raised appearance of Cal spots is largely due instead to ray floret petal curvature, which is exaggerated compared with Spring. Anthocyanins may also be less dominant in Cal petal spot pigmentation, with other pigments, like chlorophyll, potentially contributing towards the dark green colouration (Walker, 2012).

Anthocyanin malonylation is characteristic of petal spot pigments

In all three morphotypes, most *G. diffusa* spot anthocyanins are acylated by malonate (*c.* 60%), while only very small quantities of malonylated anthocyanins are present in nonspotted petal regions ($\leq 4\%$). Malonic acid is the most frequent aliphatic acyl group in acylated anthocyanins and is found throughout the Asteraceae, including *Senecio cruentus* and *Gerbera* (Harborne, 1964; Takeda *et al.*, 1986). The role of malonylated anthocyanins in *G. diffusa* petal spots is unknown. Studies in *Dahlia* and *Arabidopsis* suggest that malonic acid acylation increases anthocyanin stability (e.g. Suzuki *et al.*, 2002; Luo *et al.*, 2007) through, for example, the formation of zwitterions. Proton disassociation can decrease vacuolar sap pH protecting anthocyanin from degradation (Takeda *et al.*, 1986; Figueiredo *et al.*, 1999). Acylated anthocyanins can also be more stable against heat and light stress (Inami *et al.*, 1996; Sadilova *et al.*, 2006; Xu *et al.*, 2017; Zhao *et al.*, 2017), likely experienced by *G. diffusa* adaxial petal surfaces as the flowers open from mid-morning to mid-afternoon in a desert environment. Protection against anthocyanin degradation may be particularly important for petal spots because of their role in attracting pollinators (Johnson & Midgley, 1997; Ellis & Johnson, 2010). If *G. diffusa* also provide a heat reward for pollinators (Sapir *et al.*, 2006; Harrap *et al.*, 2017; Van Der Kooi *et al.*, 2019; Creux *et al.*, 2021), high anthocyanin stability in petal spots may counteract degradation resulting from increased capitula temperatures that are advantageous to the plant. Presumably heat and light stress is less severe for abaxial petal surfaces as they are only exposed to direct sunlight when the flowers are closed – in the early morning and late afternoon.

In Spring spots, qualitative modification of pigment identity favouring malonylation is accompanied by increased pigment production. Thus, spot development in Spring relies on two key events: a change in anthocyanin modification profile (as observed in Cal) and an overall increase in pigment production compared with nonspotted ray florets. The shift to higher proportions of malonylated anthocyanins is not unique to complex spot phenotypes and also occurs within Spring 'marks' (Fig. S4). Anthocyanin malonylation may, therefore, be the earliest event in spot specification.

GdMYBSG6-1,2, and *3* encode transcriptional activators of petal spot anthocyanin production

We isolated four subgroup 6 R2R3-MYB transcription factors expressed in *G. diffusa* ray florets. *GdMYBSG6-1*, *2* and *3* are strongly upregulated within spotted petal tissue during spot development and can trigger ectopic anthocyanin production when constitutively expressed in a tobacco heterologous context. EMSA and Y1H experiments demonstrated that petal spot-specific GdMYBSG6s bind specifically to DNA motifs found within the

GdDFR and *GdANS* promoters *in vivo* and *in vitro*. While GdMYBSG6s could bind to the *GdDFR* promoter and activate transcription in *N. benthamiana*, no transcriptional activation occurred during luciferase experiments using the *GdANS* upstream region. As such, *GdANS* may not be a direct target of GdMYBSG6s, or a larger *GdANS* promoter fragment is required for transcriptional activation. Evidently, *G. diffusa* anthocyanin petal patterning results from the spatially restricted transcription of MYB gene activators, consistent with findings in other systems (Yamagishi *et al.*, 2014; Yuan *et al.*, 2014; Martins *et al.*, 2017; Ding *et al.*, 2020).

GdMYBSG6-1, 2 and 3 likely regulate a malonyl transferase enzyme expressed within petal spots

As spot anthocyanins are largely malonylated, we investigated whether a gene encoding a malonyl transferase was upregulated in petal spots. We found that *GdMAT1* was more highly expressed in spotted compared with nonspotted tissue during spot initiation and elaboration in both Cal and Spring, while expression levels were very low in nonspotted Stein. These expression patterns mirror those of *GdMYBSG6-1*, *2* and *3*. EMSAs, Y1H, and luciferase assays confirmed that GdMYBSG6-1,2,3 can bind to *GdMAT1p* and activate transcription. The regulation of malonyl transferases by MYB transcription factors has been investigated in few studies to date. Leaf transcriptomics predicted a likely interaction in *Brassica rapa* (Rameneni *et al.*, 2020). Additionally, expression analyses conducted on berry tissues from vine plants (*Vitis vinifera*), overexpressing or silencing *VvMYBA*, indicated that VvMYBA regulates the transcription of the acyl-transferase *Vv3AT* (Rinaldo *et al.*, 2015). Here, we identified malonylation as a key event in *G. diffusa* petal spot formation along with potential regulators of this process. These components of the network controlling spot pigmentation may be under selection, contributing to the evolution of sexually deceptive spots through the production of particular hues.

Differences in *Gorteria MYBSG6* spatial expression patterns mirror evolutionary divergence

Phylogenetic analysis confirmed that *Gorteria MYBSG6* genes form a clade within subgroup 6 of the R2R3-MYB family. *Gorteria diffusa* and *G. personata* representatives cluster according to each paralogue indicating that duplications which yielded this clade predate speciation. *GdMYBSG6-4*, which lacks spot-specific expression, diverged first. This implies that the ancestor of *GdMYBSG6-1,2,3* likely already exhibited the spot-specific expression pattern characteristic of those paralogues. Whether expression was further restricted to the spotted region only along the lineage leading to the *GdMYBSG6-1,2,3* precursor, or whether *cis*-regulatory changes broadened the expression domain of *GdMYBSG6-4* sometime after the first duplication of the *GdMYBSG6* ancestor, is not clear (Fig. 6a).

In monkeyflowers, new floral pigmentation patterns occurred after duplications of MYB anthocyanin regulators: *Mimulus cupreus* and *Mimulus luteus* var *variegatus* evolved petal lobe

anthocyanin in parallel following duplication events affecting different MYB loci (Cooley *et al.*, 2011). Gene duplications leading to *GdMYBSG6-1,2,3* could have contributed towards the evolution of elaborate phenotypes. Perhaps, having three spot anthocyanin regulators increases gene dosage leading to enhanced anthocyanin production (Kondrashov & Kondrashov, 2006; Cheng *et al.*, 2018). Alternatively, GdMYBSG6-1, 2, and 3 may have divergent roles in spot anthocyanin production if their expression is localised to certain spot subregions or is specific to a specialised cell type. Hence, the divergence between *GdMYBSG6-4* and *GdMYBSG6-1,2,3* involved a clear change in spatial expression patterns within ray floret petals, but at this stage, we cannot rule out that sub- or neofunctionalisation between *GdMYBSG6-1,2,3* could also rely on divergence in protein properties.

Unlike the other *GdMYBSG6* paralogues, *GdMYBSG6-4* expression is not spot-specific. GdMYBSG6-4 is a likely regulator of petal abaxial pigmentation because the timing of its upregulation during ray floret development in nonspotted petal tissue coincides with the production of background anthocyanin pigmentation. The lack of abaxial malonylated anthocyanin, therefore, implies that within this region *GdMAT1* expression is not induced by GdMYBSG6-4. The GdMYBSG6-4 protein may have lost its ability to regulate *GdMAT1* (Fig. 6a) either because it cannot bind efficiently to the promoter of *GdMAT1* or because any binding at the *GdMAT1* locus does not allow subsequent activation of transcription. However, the high sequence homology between the DNA binding domains of GdMYBSG6-4 and the other three GdMYBSG6 paralogues (85–94%) implies potentially similar protein binding properties. Therefore, the lack of abaxial malonylated anthocyanin may instead result from external factors, such as a different chromatin landscape preventing access to the *GdMAT1* promoter in the abaxial epidermis and/or absence of a transcriptional partner for GdMYBSG6-4 (Fig. 6a). Subgroup 6 R2R3 MYB proteins often regulate anthocyanin biosynthesis through the formation of a MBW complex involving bHLH and WD40 partners (Ramsay & Glover, 2005). A bHLH protein required to partner with GdMYBSG6-4 for *GdMAT1* activation could be absent from the abaxial epidermis and this would be sufficient to account for the lack of *GdMAT1* expression within this tissue. A similar phenomenon occurs within *Antirrhinum*, where anthocyanin venation patterning within the corolla tube requires the R2R3 MYB VENOSA and a bHLH partner, DELILA. The restriction of *DELILA* expression to a specific region of the petal epidermis results in anthocyanin production only in epidermal cells overlapping veins – despite *VENOSA* transcription occurring across the entire epidermis (Shang *et al.*, 2011).

Temporal divergence in the expression of *GdMYBSG6* genes across morphotypes contributes to capitulum diversity

While a spatial restriction of *GdMYBSG6-1,2,3* expression to the basal portion of the adaxial epidermis accounts for the production of petal spots during evolution, the diversification of petal spot number in capitula across morphotypes coincides with
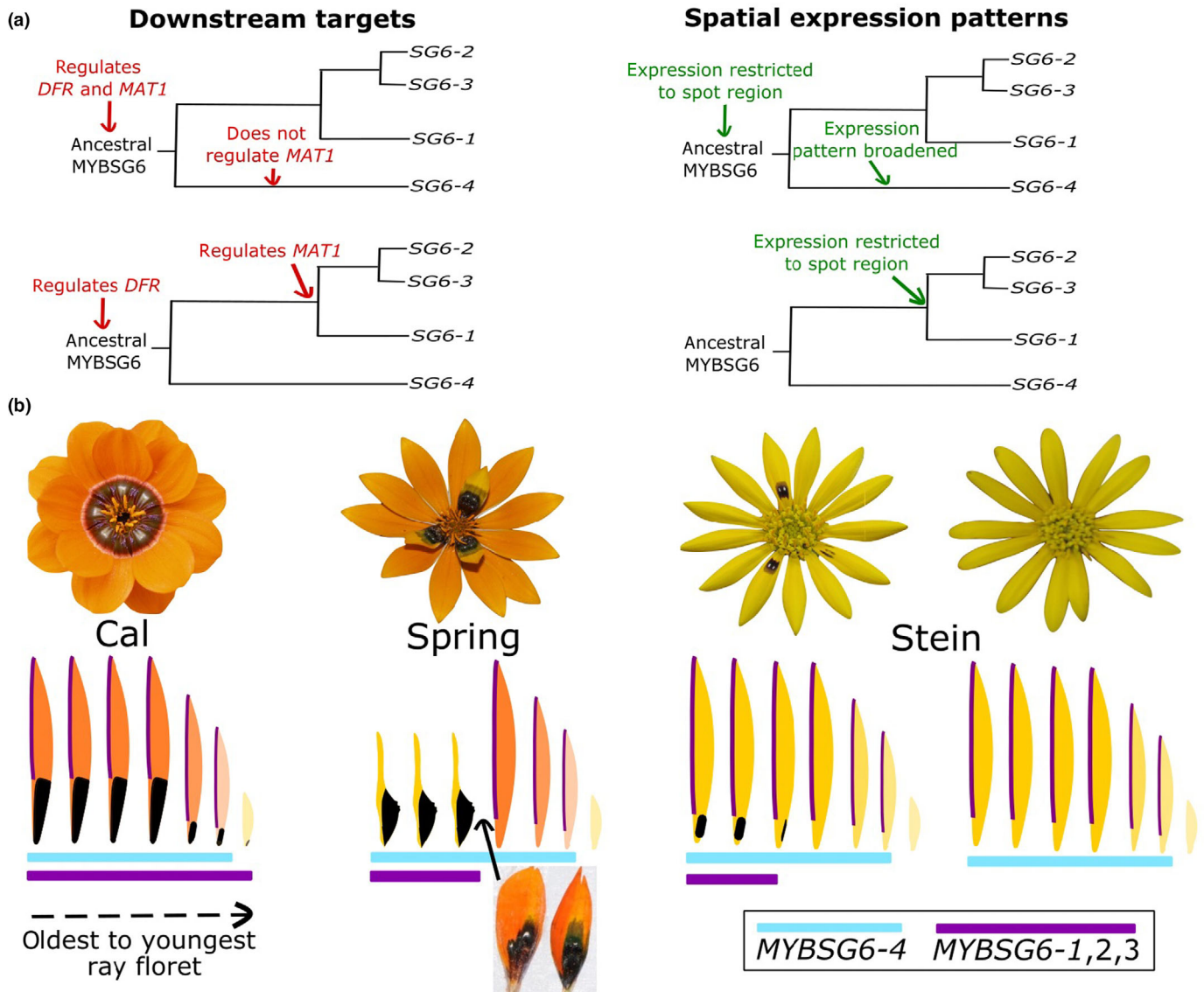
Fig. 6 Functional diversification between *GdMYBSG6* paralogues during evolution and the contribution of these genes to phenotypic variation across morphotypes. (a) Proposed evolutionary scenarios accounting for functional diversification across the *GdMYBSG6* clade based on changes in spatial expression and ability to regulate *GdMAT1* (either directly or with partners). Other variation is also possible. (b) Variation in floral pattern phenotype between morphotypes coincides with temporal changes in *GdMYBSG6-1,2,3* expression during capitulum development. Ray florets mature in a spiral phyllotactic pattern, with the inner florets being the first ones to mature. The purple vertical lines on the ray floret diagrams represent abaxial anthocyanin pigmentation. When expressed, *GdMYBSG6-1,2,3* transcription is restricted to the spotted region in all morphotypes. The duration of *GdMYBSG6-1,2,3* expression as the capitulum develops correlates with the proportion of spotted ray florets.

changes in the duration of *GdMYBSG6-1,2,3* expression during capitulum development (Fig. 6b; Thomas *et al.*, 2009; Kellenberger *et al.*, 2023). This is evident from placing our *GdMYBSG6-1,2,3* expression data from Cal, Spring, and Stein in the context of capitulum development. *Gorteria diffusa* ray florets mature in a spiral phyllotactic pattern with the first ray florets to mature in the centre of the capitulum. Within morphotypes with a few petal spots per capitulum, spots always initiate consecutively on the most mature ray florets (Thomas *et al.*, 2009). Therefore, *GdMYBSG6-1,2,3* expression is likely initiated at the same developmental timepoint across morphotypes, with the duration of

expression during capitulum development correlating with the difference in spot frequency (Fig. 6b).

It follows that consistent expression of *GdMYBSG6-1,2,3* throughout the development of all ray florets (i.e. expression during spot initiation at our developmental stage one in all ray florets) produces anthocyanic petal spots on every ray floret, creating a bullseye pattern like that in Cal. In Stein individuals with no petal spots, *GdMYBSG6-1,2,3* are not expressed at all during ray floret development. In Spring, spotted Stein individuals, and other morphotypes with petal spots on a subset of ray florets, the duration of *GdMYBSG6-1,2,3* expression is reduced compared

with that of Cal. As such, the younger outer ray florets are non-spotted because they mature after *GdMYBSG6-1,2,3* expression has ceased. Hence, the temporal expression patterns of *GdMYBSG6-1,2,3* identified in our study, and that of their upstream regulators, could account for the variation in petal spot frequency across all *G. diffusa* morphotypes.

In the Spring morphotype, petal spot number per capitulum can also vary within a single plant. It was previously suggested that this variation may contribute towards pollinator deception by making the female mimicking spots appear more randomly distributed across capitula (Thomas *et al.*, 2009). This phenotype implies that, in Spring, the duration of *GdMYBSG-1,2,3* expression during capitulum development is 'noisy'. Interestingly, within morphotypes where a subset of ray florets produce petal spots, the final petal spot produced by a capitulum is sometimes not fully developed (Fig. 6b). These 'partial spots' show signs of arrested development at different stages within each of the fused petals of a single ray floret. In Spring, the petals containing more developed spot sections also display other features characteristic of spotted ray floret petals. By contrast, the morphology of petals in the same ray floret with less developed spot sections matches the appearance of nonspotted ray floret petals (Fig. 6b). This implies that the frequency of petal spots within a capitulum can differ between morphotypes through variation in the temporal expression of upstream regulators of *GdMYBSG6-1,2,3* expression, some of which control the identity of the spotted ray floret as a whole. Such upstream regulators are expected to control anthocyanic spot production but should also control the expression of other targets specifying features such as petal shape. Identifying such regulator(s) is an exciting goal for future studies as it should both illuminate the mechanism accounting for spatial restriction of the expression of *R2R3-MYB* genes that control petal patterning, a process still not well understood in any angiosperm, but also uncover key players on which evolution acts to generate biodiversity.

Conclusion

Here, we identified three paralogous R2R3-MYB transcription factors (GdMYBSG6-1,2,3) that activate anthocyanin production during the development of elaborate petal spots in *G. diffusa*. We demonstrated that their expression is restricted to the petal region in which spots emerge and we present evidence that GdMYBSG6 directly activate *GdDFR*, while *GdANS* is an additional potential target. Our results also indicate that spatial and temporal changes in the expression of *GdMYBSG6* genes during evolution have played a key role in the creation of the spot and the production of diverse capitulum phenotypes across morphotypes, respectively. We discovered that anthocyanin malonylation is a characteristic of *G. diffusa* spot pigmentation and the GdMYBSG6 transcription factors we characterised also likely regulate the spot-specific malonyl transferase *GdMAT1*. Whether this is specific to *Gorteria* or a general chemical trick plants use to produce salient motifs on their petals, and whether the addition of a malonyl group plays a role in the formation of sexually deceptive spots, represent exciting avenues for future investigations.

Competing interests

None declared.

Author contributions

RF, GM and BJG designed the study. RW and RTK provided transcriptome data. RTK provided draft genome data. QW performed the phylogenetic analysis and genome searches. GM established a quantitative reverse transcription polymerase chain reaction reference gene. GM, RF and FNK gene hunted. GM and RF generated the plant expression vectors and transgenic tobacco. RF performed and analysed the quantitative reverse transcription polymerase chain reaction experiments. EH, EM and RF identified candidate binding motifs. RF and EM performed the EMSAs. RF genotyped, phenotyped, and imaged the tobacco transformants. FNK performed the luciferase experiments with input from EH. FNK, EH and RF performed the Y1H experiments. LH performed the UHPLC-MS/MS analysis and identified the anthocyanins. RF prepped UHPLC-MS/MS samples and performed and analysed *G. diffusa* spectrophotometer anthocyanin quantifications. RF completed all statistical analyses. EM, EH, GM and BJG advised on experimental design. RF prepared the figures and wrote the first draft of the manuscript used by RF, EM and BJG to prepare the subsequent versions. All authors read and provided feedback on the final draft and approved the final version, with the exception of GM who is deceased but whose insight from project discussions was highly valued and actively taken into account upon creation of the manuscript.

ORCID

Róisín Fattorini ID https://orcid.org/0000-0001-5596-5983

Beverley J. Glover https://orcid.org/0000-0002-6393-819X
Eva Herrero https://orcid.org/0000-0001-9881-3330
Lionel Hill https://orcid.org/0000-0001-9147-9130
Roman T. Kellenberger https://orcid.org/0000-0003-0274-2115
Farahnoz N. Khojayori https://orcid.org/0009-0008-6381-2942
Gregory Mellers https://orcid.org/0000-0001-7295-3407
Edwige Moyroud https://orcid.org/0000-0001-7908-3205
Rachel Walker https://orcid.org/0009-0002-5898-4463
Qi Wang https://orcid.org/0000-0002-6420-1720

Data availability

The data that support the findings of this study are available in the Supporting Information of this article.

References

Abe H, Nakano M, Nakatsuka A, Nakayama M, Koshioka M, Yamagishi M. 2002. Genetic analysis of floral anthocyanin pigmentation traits in Asiatic hybrid lily using molecular linkage maps. *Theoretical and Applied Genetics* **105**: 1175–1182.

Albert NW, Davies KM, Lewis DH, Zhang H, Montefiori M, Brendolise C, Boase MR, Ngo H, Jameson PE, Schwinn KE. 2014. A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. *Plant Cell* **26**: 962–980.

Albert NW, Lewis DH, Zhang H, Schwinn KE, Jameson PE, Davies KM. 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.

Badouin H, Gouzy J, Grassa CJ, Murat F, Staton SE, Cottret L, Lelandais-Brière C, Owens GL, Carrère S, Mayjonade B et al. 2017. The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. *Nature* **546**: 148–152.

Banba H. 1967. Pigments of lily flowers I. Survey of anthocyanin. *Journal of the Japanese Society for Horticultural Science* **36**: 61–65 (in Japanese).

Bello MA, Alvarez I, Torices R, Fuertes-Aguilar J. 2013. Floral development and evolution of capitulum structure in *Anacyclus* (Anthemideae, Asteraceae). *Annals of Botany* **112**: 1597–1612.

Bininda-Emonds OR. 2005. TRANSALIGN: using amino acids to facilitate the multiple alignment of protein-coding DNA sequences. *BMC Bioinformatics* **6**: 156.

Bradshaw HD Jr, Schemske DW. 2003. Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature* **426**: 176–178.

Cheng F, Wu J, Cai X, Liang J, Freeling M, Wang X. 2018. Gene retention, fractionation and subgenome differences in polyploid plants. *Nature Plants* **4**: 258–268.

Cooley AM, Modliszewski JL, Rommel ML, Willis JH. 2011. Gene duplication in *Mimulus* underlies parallel floral evolution via independent trans-regulatory changes. *Current Biology* **21**: 700–704.

Creux NM, Brown EA, Garner AG, Saeed S, Scher CL, Holalu SV, Yang D, Maloof JN, Blackman BK, Harmer SL. 2021. Flower orientation influences floral temperature, pollinator visits and plant fitness. *New Phytologist* **232**: 868–879.

Ding B, Patterson EL, Holalu SV, Li J, Johnson GA, Stanley LE, Greenlee AB, Peng F, Bradshaw HD Jr, Blinov ML et al. 2020. Two MYB proteins in a self-organizing activator-inhibitor system produce spotted pigmentation patterns. *Current Biology* **30**: 802–814.

Dirkmann M, Nowack J, Schulz F. 2018. An *in vitro* biosynthesis of sesquiterpenes starting from acetic acid. *ChemBioChem* **19**: 2146–2151.

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

Duncan G, Ellis AG. 2011. 723 *Gorteria diffusa*. *Curtis's Botanical Magazine* **28**: 341–348.

Eckhart VM, Rushing NS, Hart GM, Hansen JD. 2006. Frequency-dependent pollinator foraging in polymorphic *Clarkia Xantiana* ssp. *Xantiana* populations: implications for flower colour evolution and pollinator interactions. *Oikos* **112**: 412–421.

Ellis AG, Brockington SF, de Jager ML, Mellers G, Walker RH, Glover BJ. 2014. Floral trait variation and integration as a function of sexual deception in *Gorteria diffusa*. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **369**: 1–13.

Ellis AG, Johnson SD. 2009. The evolution of floral variation without pollinator shifts in *Gorteria diffusa* (Asteraceae). *American Journal of Botany* **96**: 793–801.

Ellis AG, Johnson SD. 2010. Floral mimicry enhances pollen export: the evolution of pollination by sexual deceit outside of the Orchidaceae. *The American Naturalist* **176**: E143–E151.

Fairnie AL, Yeo MT, Gatti S, Chan E, Travaglia V, Walker JF, Moyroud E. 2022. Eco-Evo-Devo of petal pigmentation patterning. *Essays in Biochemistry* **66**: 753–768.

Feller A, MacHemer K, Braun EL, Grotewold E. 2011. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *The Plant Journal* **66**: 94–116.

Figueiredo P, George F, Tatsuzawa F, Toki K, Saito N, Brouillard R. 1999. New features of intramolecular copigmentation by acylated anthocyanins. *Phytochemistry* **51**: 125–132.

Ganger MT, Dietz GD, Ewing SJ. 2017. A common base method for analysis of qPCR data and the application of simple blocking in qPCR experiments. *BMC Bioinformatics* **18**: 1–11.

Gaskett AC. 2011. Orchid pollination by sexual deception: pollinator perspectives. *Biological Reviews* **86**: 33–75.

Gerats AG, Farcy E, Wallroth M, Groot SP, Schram A. 1984. Control of anthocyanin synthesis in *Petunia hybrida* by multiple allelic series of the genes *An1* and *An2*. *Genetics* **106**: 501–508.

Gerats AGM, Vrijlandt E, Wallroth M, Schram AW. 1985. The influence of the genes *An1*, *An2*, and *An4* on the activity of the enzyme UDP-glucose:flavonoid 3-O-glucosyltransferase in flowers of *Petunia hybrida*. *Biochemical Genetics* **23**: 591–598.

Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**: 343–345.

Gonzalez A, Zhao M, Leavitt JM, Lloyd AM. 2008. Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in *Arabidopsis* seedlings. *The Plant Journal* **53**: 814–827.

Grotewold E. 2005. Plant metabolic diversity: a regulatory perspective. *Trends in Plant Science* **10**: 57–62.

Grotewold E. 2006. The genetics and biochemistry of floral pigments. *Annual Review of Plant Biology* **57**: 761–780.

Harborne JB. 1964. Plant polyphenols – XI.: The structure of acylated anthocyanins. *Phytochemistry* **3**: 151–160.

Harrap MJM, Rands SA, Hempel de Ibarra N, Whitney HM. 2017. The diversity of floral temperature patterns, and their use by pollinators. *eLife* **6**: 1–18.

Hellens RP, Allan AC, Friel EN, Bolitho K, Grafton K, Templeton MD, Karunairetnam S, Gleave AP, Laing WA. 2005. Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. *Plant Methods* **1**: 1–14.

Hoballah ME, Stuurman J, Turlings TC, Guerin PM, Connetable S, Kuhlemeier C. 2005. The composition and timing of flower odour emission by wild *Petunia axillaris* coincide with the antennal perception and nocturnal activity of the pollinator *Manduca sexta*. *Planta* **222**: 141–150.

Hong Y, Tang X, Huang H, Zhang Y, Dai S. 2015. Transcriptomic analyses reveal species-specific light-induced anthocyanin biosynthesis in *Chrysanthemum*. *BMC Genomics* **16**: 1–18.

Horsch RB, Fry JE, Hoffmann NL, Wallroth M, Eichholtz D, Rogers SG, Fraley RT. 1985. A simple and general method for transferring genes into plants. *Science* **227**: 1229–1231.

Inami O, Tamura I, Kikuzaki H, Nakatani N. 1996. Stability of anthocyanins of *Sambucus canadensis* and *Sambucus nigra*. *Journal of Agricultural and Food Chemistry* **44**: 3090–3096.

de Jager ML, Willis-Jones E, Critchley S, Glover BJ. 2017. The impact of floral spot and ring markings on pollinator foraging dynamics. *Evolutionary Ecology* 31: 193–204.

Johnson SD, Midgley JJ. 1997. Fly pollination of *Gorteria diffusa* (Asteraceae), and a possible mimetic function for dark spots on the capitulum. *American Journal of Botany* 84: 429–436.

Karis PO. 2007. Tribe Arctotideae Cass. In: Kubitski K, ed. *The families and genera of vascular plants, vol. VIII*. Berlin, Germany: Springer-Verlag, 223–229.

Kellenberger RT, Ponraj U, Delahaie B, Fattorini R, Balk J, Lopez-Gomollon S, Müller KH, Ellis AG, Glover BJ. 2023. Multiple gene co-options underlie the rapid evolution of sexually deceptive flowers in *Gorteria diffusa*. *Current Biology* 33: 1502–1512.

Kemp JE, Ellis AG. 2019. Cryptic petal coloration decreases floral apparency and herbivory in nocturnally closing daisies. *Functional Ecology* 33: 2130–2141.

Kondrashov FA, Kondrashov AS. 2006. Role of selection in fixation of gene duplications. *Journal of Theoretical Biology* 239: 141–151.

Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis ARN. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 35: 4453–4455.

Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2016. PARTITIONFINDER 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution* 34: 772–773.

Leonard AS, Papaj DR. 2011. 'X' marks the spot: the possible benefits of nectar guides to bees and plants. *Functional Ecology* 25: 1293–1301.

Luo J, Nishiyama Y, Fuell C, Taguchi G, Elliott K, Hill L, Tanaka Y, Kitayama M, Yamazaki M, Bailey P *et al*. 2007. Convergent evolution in the BAHD family of acyl transferases: identification and characterization of anthocyanin acyl transferases from *Arabidopsis thaliana*. *The Plant Journal* 50: 678–695.

Martins TR, Berg JJ, Blinka S, Rausher MD, Baum DA. 2013. Precise spatio-temporal regulation of the anthocyanin biosynthetic pathway leads to petal spot formation in *Clarkia gracilis* (Onagraceae). *New Phytologist* 197: 958–969.

Martins TR, Jiang P, Rausher MD. 2017. How petals change their spots: cis-regulatory re-wiring in *Clarkia* (Onagraceae). *New Phytologist* 216: 510–518.

Mellers G. 2016. *The evolution of morphological diversity in* Gorteria diffusa. Cambridge, UK: University of Cambridge.

Moeller DA. 2005. Pollinator community structure and sources of spatial variation in plant–pollinator interactions in *Clarkia xantiana* ssp. *xantiana*. *Oecologia* 142: 28–37.

Moyroud E, Glover BJ. 2017. The evolution of diverse floral morphologies. *Current Biology* 27: R941–R951.

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

Pak Dek MS, Padmanabhan P, Sherif S, Subramanian J, Paliyath G. 2017. Upregulation of phosphatidylinositol 3-kinase (PI3K) enhances ethylene biosynthesis and accelerates flower senescence in transgenic *Nicotiana tabacum* L. *International Journal of Molecular Sciences* 18: 1533.

Papiorek S, Junker RR, Alves-dos-Santos I, Melo GA, Amaral-Neto LP, Sazima M, Wolowski M, Freitas L, Lunau K. 2016. Bees, birds and yellow flowers: pollinator-dependent convergent evolution of UV patterns. *Plant Biology* 18: 46–55.

Petroni K, Tonelli C. 2011. Recent advances on the regulation of anthocyanin synthesis in reproductive organs. *Plant Science* 181: 219–229.

Quattrocchio F, Verweij W, Kroon A, Spelt C, Mo J, Koe 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. *Plant Cell* 18: 1274–1291.

Rameneni JJ, Cho SR, Chhapekar SS, Kim MS, Singh S, Yi SY, Oh SH, Kim H, Lee CY, Oh MH *et al*. 2020. Red Chinese cabbage transcriptome analysis reveals structural genes and multiple transcription factors regulating reddish purple color. *International Journal of Molecular Sciences* 21: 2901.

Ramsay NA, Glover BJ. 2005. MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends in Plant Science* 10: 63–70.

Ren M, Chen Q, Li L, Zhang R, Guo S. 2005. Successive chromosome walking by compatible ends ligation inverse PCR. *Molecular Biotechnology* 30: 95–101.

Reyes-Chin-Wo S, Wang Z, Yang X, Kozik A, Arikit S, Song C, Xia L, Froenicke L, Lavelle DO, Truco MJ *et al*. 2017. Genome assembly with *in vitro* proximity ligation data and whole-genome triplication in lettuce. *Nature Communications* 8: 1–11.

Rinaldo AR, Cavallini E, Jia Y, Moss SM, McDavid DA, Hooper LC, Robinson SP, Tornielli GB, Zenoni S, Ford CM *et al*. 2015. A grapevine anthocyanin acyltransferase, transcriptionally regulated by VvMYBA, can produce most acylated anthocyanins present in grape skins. *Plant Physiology* 169: 1897–1916.

Roessler H. 1959. *Revision der Arctotideae-Gorteriinae (Compositae)*. München, Germany: Mitteilungen der Botanischen Staatssammlung.

Sadilova E, Stintzing FC, Carle R. 2006. Thermal degradation of acylated and nonacylated anthocyanins. *Journal of Food Science* 71: C504–C512.

Sapir Y, Shmida A, Ne'eman G. 2006. Morning floral heat as a reward to the pollinators of the *Oncocyclus* irises. *Oecologia* 147: 53–59.

Schütz K, Persike M, Carle R, Schieber A. 2006. Characterization and quantification of anthocyanins in selected artichoke (*Cynara scolymus* L.) cultivars by HPLC-DAD-ESI-MS N. *Analytical and Bioanalytical Chemistry* 384: 1511–1517.

Schwinn K, Venail J, Shang Y, Mackay S, Alm V, Butelli E, Oyama R, Bailey P, Davies K, Martin C. 2006. A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the genus *Antirrhinum*. *Plant Cell* 18: 831–851.

Shang Y, Venail J, Mackay S, Bailey PC, Schwinn KE, Jameson PE, Martin CR, Davies KM. 2011. The molecular basis for venation patterning of pigmentation and its effect on pollinator attraction in flowers of *Antirrhinum*. *New Phytologist* 189: 602–615.

Sheehan H, Hermann K, Kuhlemeier C. 2012. Color and scent: how single genes influence pollinator attraction. *Cold Spring Harbor Symposia on Quantitative Biology* 77: 117–133.

Spelt C, Quattrocchio F, Mol JN, Koes RE. 2000. Anthocyanin1 of Petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. *Plant Cell* 12: 1619–1631.

Stracke R, Ishihara H, Huep G, Barsch A, Mehrtens F, Niehaus K, Weisshaar B. 2007. Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *The Plant Journal* 50: 660–677.

Stracke R, Werber M, Weisshaar B. 2001. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology* 4: 447–456.

Suzuki H, Nakayama T, Yonekura-Sakakibara K, Fukui Y, Nakamura N, Yamaguchi MA, Tanaka Y, Kusumi T, Nishino T. 2002. cDNA cloning, heterologous expressions, and functional characterization of malonyl-coenzyme A: anthocyanidin 3-O-glucoside-6″-O-malonyltransferase from Dahlia flowers. *Plant Physiology* 130: 2142–2151.

Suzuki K, Suzuki T, Nakatsuka T, Dohra H, Yamagishi M, Matsuyama K, Matsuura H. 2016. RNA-seq-based evaluation of bicolor tepal pigmentation in Asiatic hybrid lilies (*Lilium* spp.). *BMC Genomics* 17: 1–19.

Takeda K, Harborne JB, Self R. 1986. Identification and distribution of malonated anthocyanins in plants of the Compositae. *Phytochemistry* 25: 1337–1342.

Thomas MM. 2009. *Evolution of the "beetle daisy" petal spot*. Cambridge, UK: University of Cambridge.

Thomas MM, Rudall PJ, Ellis AG, Savolainen V, Glover BJ. 2009. Development of a complex floral trait: the pollinator-attracting petal spots of the beetle daisy, *Gorteria diffusa* (Asteraceae). *American Journal of Botany* 96: 2184–2196.

Unno H, Ichimaida F, Suzuki H, Takahashi S, Tanaka Y, Saito A, Nishino T, Kusunoki M, Nakayama T. 2007. Structural and mutational studies of anthocyanin malonyltransferases establish the features of BAHD enzyme catalysis. *Journal of Biological Chemistry* 282: 15812–15822.

Van Der Kooi CJ, Kevan PG, Koski MH. 2019. The thermal ecology of flowers. *Annals of Botany* 124: 343–353.

Walker RH. 2012. *Determining the regulators of petal spot development in* Gorteria diffusa. Cambridge, UK: University of Cambridge.

Wang H, Guan S, Zhu Z, Wang Y, Lu Y. 2013. A valid strategy for precise identifications of transcription factor binding sites in combinatorial regulation using bioinformatic and experimental approaches. *Plant Methods* 9: 1–11.

Winkel-Shirley B. 2001. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiology* **126**: 485–493.

Xu Q, Zhou Y, Luo L, Huang Z, Nie F, Gao G. 2017. Acylation of blueberry anthocyanins with aliphatic carbonyl acids and their stability analysis. *Advances in Engineering Research* **143**: 904–907. In *2017 6th International Conference on Energy and Environmental Protection (ICEEP 2017)*. Atlantis Press, Amsterdam, the Netherlands.

Yamagishi M, Shimoyamada Y, Nakatsuka T, Masuda K. 2010. Two R2R3-MYB genes, homologs of *Petunia AN2*, regulate anthocyanin biosyntheses in flower tepals, tepal spots and leaves of Asiatic hybrid lily. *Plant and Cell Physiology* **51**: 463–474.

Yamagishi M, Toda S, Tasaki K. 2014. The novel allele of the *LhMYB12* gene is involved in splatter-type spot formation on the flower tepals of Asiatic hybrid lilies (*Lilium* spp.). *New Phytologist* **201**: 1009–1020.

Yoshida K, Kitahara S, Ito D, Kondo T. 2006. Ferric ions involved in the flower color development of the Himalayan blue poppy, *Meconopsis grandis*. *Phytochemistry* **67**: 992–998.

Yoshida K, Negishi T. 2013. The identification of a vacuolar iron transporter involved in the blue coloration of cornflower petals. *Phytochemistry* **94**: 60–67.

Yuan YW, Sagawa JM, Frost L, Vela JP, Bradshaw HD Jr. 2014. Transcriptional control of floral anthocyanin pigmentation in monkeyflowers (*Mimulus*). *New Phytologist* **204**: 1013–1027.

Yue J, Zhu C, Zhou Y, Niu X, Miao M, Tang X, Chen F, Zhao W, Liu Y. 2018. Transcriptome analysis of differentially expressed unigenes involved in flavonoid biosynthesis during flower development of *Chrysanthemum morifolium* 'Chuju'. *Scientific Reports* **8**: 1–14.

Zhao CL, Yu YQ, Chen ZJ, Wen GS, Wei FG, Zheng Q, Wang CD, Xiao XL. 2017. Stability-increasing effects of anthocyanin glycosyl acylation. *Food Chemistry* **214**: 119–128.

Zhao D, Tao J. 2015. Recent advances on the development and regulation of flower color in ornamental plants. *Frontiers in Plant Science* **6**: 1–13.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 *Gorteria diffusa* ray floret developmental stages used in qRT-PCRs.

Fig. S2 Developmental stages used for qRT-PCR and pigment extraction in *Nicotiana tabacum*.

Fig. S3 SDS-PAGE gel of purified GdMYBSG6-2 recombinant protein used in EMSAs.

Fig. S4 Anthocyanin pigmentation phenotypes in the Spring morphotype.

Fig. S5 Pigmentation on the abaxial sides of *Gorteria diffusa* ray floret petals.

Fig. S6 Chromatograms from *Gorteria diffusa* UHPLC-MS/MS.

Fig. S7 Mass spectra from *Gorteria diffusa* UHPLC-MS/MS.

Fig. S8 Gene expression analysis in the nonspotted petals of the Stein morphotype.

Fig. S9 Gene expression analysis in the Cal and Spring morphotypes.

Fig. S10 Verification of transgene expression in transformed *Nicotiana tabacum* using RT-PCR.

Fig. S11 Leaves of *Nicotiana tabacum* transgenic lines constitutively expressing *GdMYBSG6* genes.

Fig. S12 Anthocyanin extractions from floral tissue of transgenic *Nicotiana tabacum* lines constitutively expressing *GdMYBSG6* genes.

Fig. S13 Chromatograms of anthocyanin petal extractions from transgenic *Nicotiana tabacum* constitutively expressing *GdMYBSG6* genes.

Fig. S14 Mass spectra of anthocyanin petal extractions from transgenic *Nicotiana tabacum* constitutively expressing *GdMYBSG6* genes.

Fig. S15 *GdMYBSG6* relative expression levels in floral tissue of transgenic *Nicotiana tabacum* carrying a *35S::GdMYBSG6* transgene.

Fig. S16 Expression analysis of *Gorteria diffusa* genes coding for late anthocyanin biosynthetic enzymes.

Fig. S17 Yeast one-hybrid experiments testing for binding of GdMYBSG6-1–3 to DNA sequences upstream of the coding region of selected anthocyanin biosynthetic enzymes.

Methods S1 Plant growth conditions.

Methods S2 Anthocyanin pigment extraction and quantification.

Methods S3 Isolating candidate genes and promoter regions.

Methods S4 qRT-PCR primer design and cycle conditions.

Methods S5 *Nicotiana tabacum* stable transformation procedure.

Methods S6 Map of plasmids used in *Nicotiana tabacum* transformation.

Methods S7 Map of plasmid used to produce the recombinant GdMYBSG6-2 protein used in EMSA.

Methods S8 Recombinant protein purification and EMSA.

Methods S9 Identifying possible GdMYBSG6 binding sites.

Methods S10 Preparing oligonucleotides for EMSA.

Methods S11 Map of plasmids used in yeast one-hybrid experiments.

Methods S12 Yeast one-hybrid bait strain selection.

Methods S13 Map of plasmid used in dual-luciferase assays.

Methods S14 Dual-luciferase assay culture preparation.

Methods S15 Data analysis.

Table S1 Primer sequences.

Table S2 Anthocyanin quantification and characterisation.

Table S3 Anthocyanins detected in the ray floret tissue of *Gorteria diffusa*.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.