

# A Highly Mutable *GST* is Essential for Bract Colouration in *Euphorbia Pulcherrima* Willd. Ex Klotsch

**Vinicius Vilperte**

Leibniz Universität Hannover: Leibniz Universität Hannover

**Robert Boehm**

Klemm + Sohn GmbH & Co., KG

**Thomas Debener** (✉ [Debener@genetik.uni-hannover.de](mailto:Debener@genetik.uni-hannover.de))

Leibniz Universität Hannover

---

## Research article

**Keywords:** anthocyanin, *Euphorbia pulcherrima*, ionizing radiation, glutathione S-transferase, mutation breeding, poinsettia, short repeat sequences

**Posted Date:** October 26th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-96082/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at BMC Genomics on March 23rd, 2021. See the published version at <https://doi.org/10.1186/s12864-021-07527-z>.

# Abstract

**Background:** Mutation breeding is an extraordinary tool in plant breeding to increase the genetic variability, where mutations in anthocyanin biosynthesis are targets to generate distinctive phenotypes in ornamental species. In poinsettia, ionizing radiation is routinely applied in breeding programs to obtaining a range of colours, with nearly all pink and white varieties being obtained after  $\gamma$ - or X-ray mutagenesis of red varieties. In the present study we performed a thorough characterization of a potential mutagenesis target gene as the main responsible for the '*white paradox*' in poinsettia

**Results:** We identified a *GST* gene in poinsettia (*Bract1*) as an essential factor for the expression of anthocyanin-based red colouration of bracts, which presents a high phylogenetic similarity to known anthocyanin-related GSTs. Red poinsettia varieties and white mutants generated from these varieties by X-ray were analysed for polymorphisms related to the '*white paradox*' in the species. A 4 bp mutation in a short repeat within the coding region of *Bract1* is most likely responsible for the appearance of white phenotypes upon irradiation treatment. The polymorphism between wild-type and mutant alleles co-segregates with the phenotype in progeny from heterozygous red and white parents. Moreover, overexpression of *Bract1* wild-type allele in *Arabidopsis tt19* mutants restored the anthocyanin phenotype, while the *Bract1* mutated allele showed to be non-functional.

**Conclusions:** The identified repeat seems to be highly unstable, since mutated plants can be easily detected among fewer than 200 shoots derived from 10 mutated plants. Our data indicate that particular short repeat sequences, similar to microsatellite sequences or so-called dynamic mutations, might be hot spots for genetic variability. Moreover, the identification of the *Bract1* mutation fills a gap on the understanding on the molecular mechanism of colour formation in poinsettia

## Background

Poinsettia, *Euphorbia pulcherrima* Willd. ex Klotsch, commonly known as Christmas Star, is an important ornamental crop, especially due to its association with Christmas time in North America, Europe, and Asia, with annual sales reaching nearly 150 million dollars in the USA (USDA, 2018). Its ornamental value is based on its intensely coloured bracts, which can be red, white, pink, or yellow or even have dual, scattered, or marbled colourations. Nonetheless, poinsettia breeding still focuses on red- and white-coloured varieties due to higher acceptance by consumers. In 2018, in Germany, approximately 80% of the poinsettias grown were red, 11% were white, and 9% were pink or had dual/scattered colouration (van der Ploeg, 2018).

Ionizing radiation is an important tool in mutation breeding for new colour variations in poinsettia, with nearly all pink and white varieties being obtained after gamma or X-ray mutagenesis of shoots of red varieties. Poinsettia mutation breeding is usually performed on cuttings that are irradiated with moderate dosages (~ 20 Gy), and mutants are selected on side shoots of the originally irradiated shoots. Flowering induction in the species occurs under short-day conditions and is accompanied by the development and

colouration of bracts. Therefore, green leaves and red bracts occur concomitantly and accumulate different groups of pigments, i.e., chlorophylls and anthocyanins (Pomar and Ros Barceló, 2007; Moustaka et al., 2018). Several anthocyanin types have been identified in poinsettia bracts and are responsible for its colouration range (Asen, 1958; Slatnar et al., 2013; Nitarska et al., 2018); however, molecular information is still limited for the species (Gu et al., 2018; Vilperte et al., 2019). Nonetheless, genes responsible for the biosynthesis of the anthocyanin pathway have been intensively characterized in a range of species, with its regulation being highly dependent on R2R3-MYB regulatory genes and MYB-bHLH-WD40 (MBW) regulatory complexes (Dubos et al., 2010; Petroni and Tonelli, 2011; Zhao and Tao, 2015).

Once synthesized on the cytoplasmic surface of the endoplasmic reticulum (ER), anthocyanin molecules need to be stored in the vacuole to prevent oxidation and loss of colour (Bassham, 2015). Two main models of anthocyanin transport have been proposed: i) a vesicle trafficking-mediated model, where vesicle-like structures filled with anthocyanins are imported into the central vacuole via vesicle fusion (Poustka et al., 2007; Grotewold and Davies, 2008; Gomez et al., 2011); and ii) a transporter-mediated model, where anthocyanins are carried across the vacuolar membrane by transport proteins (e.g., ABC and MATE transporters) with the help of glutathione S-transferase (GST) enzymes (Grotewold, 2006; Zhao and Dixon, 2010; Saito et al., 2013). GSTs can bind to anthocyanin molecules to form a complex, thus escorting them from the ER to the vacuole, preventing oxidation (Mueller et al., 2000; Zhao and Dixon, 2010; Sun et al., 2012). Anthocyanin-related GSTs play major roles in anthocyanin transport, since loss of function of these proteins leads to phenotypes with a lack of pigmentation, such as *bz2* (*Bronze-2*) in maize, *an9* (*Anthocyanin 9*) in petunia, *tt19* (*Transparent Testa 19*) in Arabidopsis, *fl3* (*Flavonoid3*) in carnation, *riant* (*regulator involved in anthocyanin transport*) in peach, and *rap* (*reduced anthocyanin in petioles*) in strawberry (Alfenito et al.; Marrs et al., 1995; Larsen et al., 2003; Kitamura et al., 2004; Cheng et al., 2015; Luo et al., 2018). In our previous study, an anthocyanin-related GST-like gene showed higher expression in a red poinsettia variety than in the white counterpart, thus making it a promising candidate responsible for the so-called '*white paradox*' (Vilperte et al., 2019).

In our current study, we identified an anthocyanin-related *GST* as the most likely target of the radiation-induced mutation of red poinsettias in white bract sports. Using different approaches, this study demonstrates the functionality of the poinsettia *GST* as an anthocyanin transporter. Most importantly, we show that a short repeat motif within the coding region of the gene is highly unstable upon mutation treatment, which leads to the high frequency of anthocyanin mutations observed in commercial mutation breeding. In addition to facilitating mutation breeding for bract colours, these results may be a starting point for analysing the genetic instability of short repeat sequences in plants.

## Results

### Identification and characterization of Bract1

In a previous study (Vilperte et al., 2019), we observed higher expression of an anthocyanin-related *GST-like* gene in the red poinsettia variety ‘Christmas Feelings’ than in its white counterpart ‘Christmas Feelings Pearl’. To investigate whether a similar phenomenon is observed in other red and white poinsettia pairs, we performed RT-qPCRs for six pairs of red-bracted poinsettia varieties and their independently generated white mutants. NRQ values were calculated relative to one of the biological replicates of the ‘Chr. Glory’ variety according to the Pfaffl method and equations (Pfaffl, 2001). The levels of *Bract1* expression varied among all varieties, with the varieties ‘Christmas Feelings’, ‘Titan’ and ‘SK130’ showing the highest relative expression. Although no lack of expression was observed in any of the white varieties, all red varieties showed significantly higher expression of *Bract1* than their white counterparts (Fig. 1).

To further characterize the anthocyanin-related *GST* in poinsettia, termed *Bract1* hereafter, we sequenced the complete coding and intronic regions of the gene for the ‘Vintage’ variety. The final full-length *Bract1* sequence contains 2,314 bp, with three exons (147 bp, 48 bp and 450 bp) and two introns (455 bp and 1,214 bp) (Fig. 2a). The coding sequences (CDSs) of *Bract1* from 14 red- and white-bracted poinsettia varieties (‘Noel’, ‘Valentino’, ‘Christmas Feelings’, ‘Christmas Feelings Pearl’, ‘Christmas Glory’, ‘Christmas Glory White’, ‘Joy’, ‘Joy White’, ‘Titan’, ‘Titan White’, ‘Bravo’, ‘Bravo White’, ‘SK130’ and ‘SK130 White’) were further sequenced. The sequence alignment showed high similarity in the CDS for all varieties, except for six single-nucleotide polymorphisms (SNPs) that were identified in both the ‘Chr. Glory’ and ‘Bravo’ varieties (Table 1). This result shows the presence of at least two allelic forms of the *Bract1* gene. Additionally, a 4 bp deletion located 8 bp upstream of the first exon-intron junction was observed in all white varieties (Fig. 2b). The deletion is located in a short repeat locus, resembling a short simple sequence repeat (SSR), with a tetranucleotide motif ((CTTC)<sub>3</sub>) composition. The exact location of the (CTTC)<sub>3</sub> motif is shown in Fig. 1a. The full-length gene sequence and CDS are available in Additional File S1.

Table 1  
List of SNPs identified in the sequenced varieties in comparison to the Vintage variety.

Position	Original	Alternative	Varieties
90	T	A	Chr. Glory/Bravo
120	T	C	Chr. Glory/Bravo
150	C	A	Chr. Glory/Bravo
525	C	T	Chr. Glory/Bravo
578	A	G	Chr. Glory/Bravo
604	C	A	Chr. Glory/Bravo

*Bract1* encodes a putative functional protein of 214 amino acids (aa) and a mass of 24.6 kDa, with distinctive GST components: a conserved GSH-binding site (G-site) located in the N-terminal domain and

a C-terminal substrate-binding domain (H-site) (Dixon et al., 2002). The predicted protein from the CDS containing the 4 bp deletion is a putative truncated protein with an early stop codon at position aa52 due to a frameshift in the mRNA. The full-length amino acid sequence and the truncated version are available in Additional File S1.

### **Bract colouration associated with a deletion in the *Bract1* gene**

The colour range in poinsettia varieties is obtained either through classic breeding (crossing) or mutagenic breeding (radiation), thus generating a spectrum of bract colours, such as pink, marble, orange and white/creamy. The white varieties are often obtained through radiation mutagenesis of the red varieties, followed by shoot development and trait selection. Therefore, red and white poinsettias from the same variety are referred to as 'pairs' due to their highly similar genetic background. However, not all red varieties can produce white sports through radiation. Therefore, red poinsettia varieties are distinguished into 'heterozygous' and 'homozygous' for the colouration locus according to their ability to generate white sports and according to the segregation of red and white phenotypes in progeny of crosses with white genotypes.

Since the 4 bp *indel* in the SSR locus of *Bract1* had shown indications of polymorphism among the different poinsettia varieties—and a correlation with bract colouration—we used a genotyping approach based on the fluorescent labelling of PCR fragments. We genotyped 22 different poinsettia varieties bearing red and white bracts (Fig. 3a). All the red heterozygous varieties showed two distinct copies of the allele (with and without the 4 bp deletion), while their white counterparts showed only the copy with the deletion. On the other hand, homozygous red varieties (i.e., those unable to generate white sports) showed only the copy without the deletion.

We further genotyped a segregating population with 190 progeny from 'Joy' (Rr) x 'Joy White' (rr) containing 36 white and 154 red plants (Fig. 3b). Contrary to expectation, we observed a deviation in the segregation ratio, which was approximately 4:1 (red:white), instead of the expected 1:1 ratio for this crossing. This may be explained by the fact that seeds from white varieties are less vital than those from red varieties (von Tubeuf, Selecta One, pers. comm.). In addition, white varieties also exhibit lower pollen fertility, thus increasing the chances of self-pollination when red varieties are used as a female parent (von Tubeuf, Selecta One, pers. comm.). In fact, 17 red progeny showed only the wild-type copy of the allele (data not shown), which can be attributed only to self-pollination. Nonetheless, all the white progeny showed only the allele copy containing the deletion, thus reinforcing our hypothesis that the presence of the allele containing the deletion in a homozygous recessive state is correlated with the white phenotype.

### ***Bract1* is the anthocyanin-related GST orthologue in poinsettia**

As GST genes occur in large gene families, we wanted to analyse whether the poinsettia GST gene was related to other GST genes involved in anthocyanin transport to the vacuole. Therefore, we computed a phylogenetic tree from the deduced amino acid sequences of 95 GST family members from our

previously assembled poinsettia transcriptome (Vilperte et al., 2019), as well as the *Bract1* and anthocyanin-related GSTs from other species (*CkmGST3*, *LcGST4*, *VvGST4*, *PhAN9*, *PpRiant1*, *PpRiant2*, *AtGSTF11* and *AtTT19*). Nine GST classes were identified among the poinsettia GSTs: Tau, Theta, Lambda, Zeta, Phi, tetrachlorohydroquinone dehalogenase (TCHQD), glutathionyl hydroquinone reductase (GHR), dehydroascorbate reductase (DHAR) and eukaryotic translation elongation factor 1B-γ (Ef1By). Except Tau and Ef1By, all other GST classes showed a single cluster (Fig. 4). All anthocyanin-related GSTs belong to the Phi class and clustered together in the phylogenetic tree, with *Bract1* showing high similarity with these GSTs.

By aligning the *Bract1* nucleotide CDSs with those of anthocyanin-related GSTs from other species, an overall nucleotide similarity of 61.9% was observed (Additional File S3). Protein alignment of BRACT1 with the other anthocyanin-related GSTs resulted in an overall similarity of 58.3%, with the peach Riant1 protein showing the highest similarity (66.5%) (Additional File S3). Interestingly, we identified seven amino acid residues, previously reported as specific to anthocyanin-related GSTs (Kitamura et al., 2012; Hu et al., 2016; Kou et al., 2019), that are conserved in the protein alignment, except in ATGSTF11: 2Val, 11Ala, 13Cys, 62Phe, 90Leu, 91Glu and 152Ser (Fig. 5). In summary, these results indicate that *Bract1* is the anthocyanin-related GST orthologue in poinsettia.

### **Bract1 functionally complements the Arabidopsis *tt19* mutant phenotype**

To examine the *in vivo* function of *Bract1* as an anthocyanin transporter, we tested the ability of *Bract1* cDNA to functionally complement the Arabidopsis GST mutant *tt19*, which is defective in the expression of anthocyanins in aboveground organs and seeds. Two constructs containing the *Bract1* cDNA (with and without the 4 bp deletion) under the cauliflower mosaic virus (CaMV) 35S promoter were introduced into the *tt19* mutant by the floral-dip method (Clough and Bent, 1998; Zhang et al., 2006). Although the constructs contained a GFP marker for the selection of transgenic events, we genotyped 10 independent biological replicates from the T<sub>2</sub> progeny of *tt19/35S::Bract1* and *tt19/35S::Bract1\_mut* transgenic plants. All progeny contained the correct allele from the *Bract1* gene, thus confirming the correct integration of the transgenic construct (Fig. 6a).

Upon stimulation of anthocyanin accumulation in seedlings by irradiation with red/blue LEDs, the *tt19/35S::Bract1* transgenic lines displayed a purple hypocotyl phenotype at the seedling stage, similar to the Columbia (*Col-0*) line but not the *tt19* mutant (Fig. 6b). On the other hand, *tt19/35S::Bract1* transgenic lines did not show complementation of the anthocyanin phenotype. The *ProAtUbi::GFP* construct, used as a control for infiltration, did not result in any phenotypic changes. Moreover, transgenic plants harbouring *Bract1* did not complement the seed colour of *tt19*, as the seed colour at the ripening stage remained the same as that of the mutant *tt19* in transgenic plants (Fig. 6c). This finding suggests that *Bract1* may have distinct functions from *TT19* during seed coat pigmentation. Taken together, these results not only emphasize the role of *Bract1* in anthocyanin transport in poinsettia but also demonstrate that a deletion in its coding region leads to a colourless phenotype.

## De novo mutations occur with high frequency and include deletion of the 4 bp repeat

To study the stability of the 4 bp repeat within the first exon of the *Bract1* gene, we analysed DNA samples from mutation experiments conducted over the last 4 years at Selecta One. In brief, 10 cuttings from the varieties 'Aurora', 'SK159 Dark Pink', 'Aurora Jingle' and 'SK183' were irradiated with 20 Gy (30 Gy for 'SK183'), and subsequently, side shoots from those cuttings were further propagated. DNA was extracted and analysed as previously described from 377, 191, 188 and 186 of the propagated side shoots. Table 2 shows the results indicating that out of 942 samples, 9 mutations could be detected. Three mutated progeny were identified in both 'SK159 Dark Pink' and 'Aurora Jingle' individuals and two in the 'SK183' individuals, and only one mutated individual was identified in the 'Aurora' progeny. Unfortunately, as this was part of a commercial breeding programme, individual shoots were not labelled in a way that would allow tracing them back to one of the original shoots that were irradiated. However, even if all the mutations detected in each of the separate mutation treatments were redundant and originated from one original mutational event, the frequency was extraordinarily high.

Table 2

**Fragment analysis of progeny from three X-ray-irradiated poinsettia varieties.** Two methods were used for the fragment analysis: polyacrylamide gel electrophoresis (PAGE) and fragment length analysis (FLA) by capillary electrophoresis.

Variety/ year of irradiation	Number of progeny			Type of analysis
	Total	Homozygous (RR)	Heterozygous (Rr)	
Aurora/2016	377	376	1	PAGE
SK159 Dark Pink/2018	191	187	3	FLA
Aurora Jingle/2018	188	185	3	FLA
SK183/2018	186	1	184	FLA

## Microsatellite repeats are not an anthocyanin-related feature

The microsatellite repeat present in the *Bract1* gene shows signs of instability upon irradiation treatment. To identify whether such repeats are a common feature for anthocyanin-related GSTs in Euphorbiaceae or related taxa or are a family-specific feature, we first computed a phylogenetic tree from the CDSs of *Bract1*, known anthocyanin-related GSTs (*CkmGST3*, *LcGST4*, *VvGST4*, *PhAN9*, *PpRiant1*, *PpRiant2*, and *AtTT19*) and GST-like orthologues from Euphorbiaceae species (*Euphorbia esula*, *Euphorbia peginensis*, *Ricinus communis*, *Jatropha curcas*, *Hevea brasiliensis* and *Manihot esculenta*). Figure 7 shows that *Bract1* shared high similarity with the GSTs from the two *Euphorbia* species (*E. esula* and *E. peginensis*) but also closely clustered with the GSTs from the other Euphorbiaceae species. Although none of the

GST-like genes from other Euphorbiaceae have been investigated as putative anthocyanin transporters, they may perform a similar function due to their homology with *Bract1* and the other anthocyanin-related GSTs.

Furthermore, we wanted to assess the distribution of the microsatellite repeat and possibly investigate its origin. The sequence alignment shows that none of the GSTs analysed contain the same (CTTC)<sub>3</sub> SSR motif observed in *Bract1* (Fig. 8). When analysing the GSTs from Euphorbiaceae species, two CTTC repeat units were present in *E. pekinensis*, with a single nucleotide substitution responsible for the loss of the third repeat. *J. curcas*, *H. brasiliensis* and *M. esculenta* showed two CTTC repeat units, with two nucleotide substitutions related to the loss of the third repeat. Last, *E. esula* and *R. communis* showed a single CTTC unit and three nucleotide substitutions in the microsatellite region. The anthocyanin-related GSTs showed greater distinction in the microsatellite region. Three of the genes contained a single CTTC repeat unit (*CkmGST3*, *AtTT19* and *LcGST4*), while the others did not contain any repeat unit (*VvGST4*, *PpRiant1*, *PpRiant2* and *PhAN9*), which was due to the higher number of nucleotide substitutions (ranging from 4 to 7 substitutions). In conclusion, even though some of the known anthocyanin-related GSTs contain a CTTC sequence, CTTC repeats seem to be a common feature of Euphorbiaceae GSTs, and the three repeats from the *E. pulcherrima* GST (*Bract1*) behave in an unstable manner upon being subjected to ionizing irradiation.

## Discussion

Anthocyanins, a class of flavonoid secondary metabolite compounds (Liu et al., 2018), are responsible for providing orange to blue colours in plant tissues, and their biosynthetic and regulatory mechanisms have been widely characterized (Zhao and Tao, 2015). However, there is still debate on the mechanism of anthocyanin transport from the ER to the vacuole (Poustka et al., 2007; Saito et al., 2013). Strong evidence for the involvement of transport proteins (Goodman et al., 2004; Francisco et al., 2013), with a special role of GST enzymes, has been reported in several plant species (Alfenito et al.; Larsen et al., 2003; Kitamura et al., 2004). In the present study, we demonstrated that the *Bract1* gene functions as an anthocyanin transporter in poinsettia and that a highly mutable repeat in its coding region leads to frequent deletions and therefore to a colour-deficient phenotype.

### **Bract1 is a functional GST gene related to anthocyanin transport in poinsettia**

GSTs are a large and diverse group of enzymes with multifunctional roles, especially in the detoxification of xenobiotics as well as in responses to biotic and abiotic stresses (Agrawal et al., 2002; Dixon et al., 2010). The classification of GSTs is based on sequence conservation, genomic organization, and physiochemical properties, among other features (Edwards et al., 2000; Lallement et al., 2014; Islam et al., 2019). Based on our previous study (Vilperte et al., 2019), we identified 95 *GST* genes in poinsettia and phylogenetically classified them into nine different classes based on their similarity with known *Arabidopsis* GSTs (Fig. 4). To date, 14 GST classes have been identified in plants: tau (U), phi (F), lambda

(L), DHAR, theta (T), zeta (Z), EF1B $\gamma$ , TCHQD, microsomal prostaglandin E-synthase type 2 (mPGES-2), GHR, metaxin, Ure2p, hemerythrin (H) and iota (I) (reviewed by (Lallement et al., 2014)).

A large number of GSTs have been identified in plant species, such as 49 in *Capsella rubella* (He et al., 2016), 55 in *Arabidopsis* (Dixon and Edwards, 2010), 61 in Citrus (Licciardello et al., 2014) and 139 in *L. chinensis* (Hu et al., 2016). *Bract1* clusters with high bootstrap support with anthocyanin-related GSTs from other species (e.g., *AtTT19*, *PhAN9* and *VvGST4*), with all of these GSTs belonging to the phi class. Known anthocyanin-related GSTs belong almost exclusively to the phi class, except for *Bronze-2* from maize, which belongs to the tau class (Marrs et al., 1995). Further support for *Bract1* being a member of the phi-type plant GST genes is provided by the presence of two introns as a characteristic of this group of genes, such as *AN9* from petunia and *TT19* from *Arabidopsis* (Alfenito et al.; Mueller et al., 2000).

Complementation studies using *Arabidopsis tt19* mutants have been widely applied as proof of concept for the function of GSTs as anthocyanin transporters (Alfenito et al.; Mueller et al., 2000; Kitamura et al., 2012; Hu et al., 2016; Pérez-Díaz et al., 2016; Jiang et al., 2019; Kou et al., 2019; Liu et al., 2019). Due to the high amino acid conservation of GST enzymes involved in flavonoid accumulation among species (Zhao, 2015), they can complement each other's anthocyanin-deficient mutants (Alfenito et al.; Larsen et al., 2003). However, similar to our observation for *Bract1*, not all of these genes complemented both the shoot and seed phenotypes (Luo et al., 2018; Jiang et al., 2019). A direct complementation of poinsettia white mutants with the functional *Bract1* would ultimately prove its function in bract colouration. However, neither *Agrobacterium*-mediated infiltration nor biolistic particle delivery system (a.k.a. gene gun) were successful for transient expression studies (data not shown). Stable transformation in poinsettia have been done using electrophoresis-based methods (Vik et al., 2001; Clarke et al., 2006), but no stable transgenic poinsettia was obtained. Successful stable transformation via *Agrobacterium*-mediated infiltration has been previously achieved, but the process is time-consuming (Clarke et al., 2008). Attempts to perform stable transformation of poinsettia with *Bract1* alleles will bridge the current knowledge gap but are out of the scope of the present study.

### **A loss-of-function mutation in *Bract1* is the cause of the “white paradox” in poinsettia**

Based on our results, we hypothesize that deletion of one unit of the repeat in the *Bract1* gene is responsible for most of the white genotypes in poinsettia. This hypothesis is strongly supported by the evidence that the *tt19/35S::Bract1\_mut* mutant was not able to complement the anthocyanin phenotype in the *Arabidopsis tt19* mutant, unlike the *tt19/35S::Bract1* mutant. Mutations in GSTs leading to colourless phenotypes have been previously reported. A mutation in the *fl3* gene in carnation leads to a light pink phenotype, but a brighter phenotype is observed upon complementation by petunia *AN9* and maize *Bz2* (Larsen et al., 2003). In peach, four alleles of a GST gene (*Riant*) were identified, with two of them containing frameshift mutations and unable to complement the *Arabidopsis tt19* phenotype. Varieties containing copies of the mutated alleles in a homozygous state showed flowers with white variegated phenotypes (Cheng et al., 2015). Last, a single-nucleotide polymorphism (SNP) in the strawberry *RAP* gene, leading to a premature stop codon, results in a mutant with green petioles and

leaves. The non-functional *rap* gene was not able to complement Arabidopsis *tt19*, while wild-type *RAP* was successful (Luo et al., 2018).

In our analyses, all six independently generated white mutants of red varieties displayed the same deletion of a 4 bp repeat in *Bract1*, whereas the original varieties all contained a fully functional copy of the gene. In addition, co-segregation of the deletion with the white phenotype was observed in a segregated population of 190 progeny. Furthermore, a novel mutation leading to a homozygous recessive allele of *Bract1* among 184 samples obtained from irradiated cuttings of the heterozygous line SK183 led to a white phenotype (Table 3). Altogether, the results of this study present strong evidence that the four-base deletion in *Bract1* is the cause of the red-to-white shift in the poinsettia varieties analysed here. However, as anthocyanin biosynthesis involves several steps, other regulatory and structural genes might give rise to white mutants as well, as has been shown in numerous other examples (Koseki et al., 2005; Morita et al., 2012; Ben-Simhon et al., 2015; Luo et al., 2016). We did not detect these genes in our current plant material perhaps due to the much higher mutation rate of the *Bract1* gene than of less mutable genes.

Table 3  
List of poinsettia varieties used in the present study.

Variety ID	Variety name	Bract colour	Observation
1	Christmas Feelings	Red	
2	Christmas Glory	Red	
3	Joy	Red	
4	Bravo	Red	
5	Titan	Red	
6	SK130	Red	
7	Christmas Feelings Pearl	White	Mutation from Chr. Feelings
8	Christmas Glory White	White	Mutation from Chr. Glory
9	Joy White	White	Mutation from Joy
10	Bravo White	White	Mutation from Bravo
11	Titan White	White	Mutation from Titan
12	SK130 White	White	Mutation from SK130
13	Vintage	Red	
14	Christmas Aurora	Red	
15	Happy Day	Red	
16	Tabalunga	Red	
17	Christmas Day	Red	
18	Christmas Eve	Red	
19	Noel	Red	
20	Valentino	Red	
21	Prestige Red	Red	
22	Christmas Cracker	Red	

### Bract1 contains a short highly mutable four-base repeat

Upon X-ray treatment, red poinsettia plants produce progeny bearing white phenotypes with high frequencies, often based on only 10 irradiated cuttings (von Tubeuf, Selecta One, pers. comm., Selecta One). This phenomenon is associated with a deletion in a short repeat in the *Bract1* gene of white mutants in a homozygous state. The mutations in all six independent mutant pairs that we detected are

exactly identical, which indicates that the X-ray treatment did not directly cause the mutation but rather led to changes indirectly by stimulating the DNA repair mechanisms via replication errors, by increasing recombination or by the other mechanisms discussed for mutations in repeat sequences (Pearson et al., 2005). The possible involvement of replication-based errors is supported by our observation that upon amplification of the repeat via standard PCR from cloned *Bract1* wild-type or mutant allele, a low level of variants carrying four-base indels can always be detected (data not shown).

Radiation is frequently used as a tool for mutagenic breeding in poinsettia. In contrast to ethyl-methanesulphonate (EMS)-based chemical mutagenesis, which produces point mutations with high frequency (Greene et al., 2003), ionizing radiation (e.g., X-rays and  $\gamma$ -rays) induces DNA oxidative damage, such as double-strand breaks (DSBs), base substitutions, deletions and chromosomal alterations, at a lower frequency, frequently resulting in loss of gene function (Morita et al., 2009; Kazama et al., 2011; Jo and Kim, 2019). SSRs are among the most variable types of repetitive sequences in the genome (Ellegren, 2004). Studies have shown that SSR instability increases with plant development (Golubov et al., 2010) and abiotic stress (Yao and Kovalchuk, 2011). This might be another explanation for the frequent observation of repeat changes in the *Bract1* gene after X-ray irradiation, although the small number of repeats (i.e., three) of four base pairs each does not fit the most widely applied criteria used to define SSRs, which usually focus on sequences with a larger number of repeats.

However, little information about the genetics and dynamics is available for short repeats. A majority of studies compared historical events for mostly shorter SSRs (2 and 3 bp repeats with larger repeat numbers) in present-day populations or the dynamic repeats responsible for human diseases (mostly trinucleotide repeats), which usually display effects beyond those of large numbers of repeats (> 30 repeats (Pearson et al., 2005))

Our observation that a large number of mutation events could be observed in the side shoots of ten irradiated plants indicates an unusually high mutation rate, which is in contrast to the few reports in which exact mutation rates have been reported for vegetatively propagated crops (Schum and Preil, 1998). In one example, the woody ornamental plant *Tibouchina urvelliana* was irradiated three independent times with a 45 Gy dose, resulting in 0.06% dwarf mutants each time (Schum and Preil, 1998). However, several authors reported that the radiosensitivity of vegetative tissues varies greatly among species and tissues (Esnault et al., 2010), so exact comparative estimations of mutation frequencies have a very limited accuracy among species and conditions. However, experiments with transgenic *Arabidopsis* lines harbouring constructs designed to analyse restoration of GUS open reading frames by either recombination or by restoring in-frame translation by mutations in SSRs demonstrate the occurrence of easily detectable numbers of somatic mutation events (Golubov et al., 2010; Yao and Kovalchuk, 2011). Together with the careful selection of side shoots after X-ray irradiation of poinsettia, this finding may explain the high rate of recessive mutations detected here.

In this study, we showed that the poinsettia *Bract1* gene is an active GST gene involved in the expression of anthocyanins in poinsettia bracts. Furthermore, a 4 bp deletion in a short repeat within the coding

region of *Bract1* is the most likely cause of many mutations that lead to a white bract colour. This mutation occurs with an unusually high frequency and is presumably an indirect effect of X-ray mutagenesis. Future analyses using mutagenesis in transgenic *Arabidopsis* lines harbouring *Bract1* might help elucidate the causes of the high instability of this repeat. Moreover, this result might also serve as a reference for the study of other repeat-containing structural genes as potential mutational hot spots in plant genomes.

## Methods

### Plant material

A range of red- and white-bracted varieties of poinsettia was used in this study for the different analyses (Table 3). Among the varieties, we used so-called pairs of red varieties and white mutants generated from these varieties by  $\gamma$ -ray or X-ray mutagenesis by the company Selecta One (Stuttgart, Germany). In addition, a segregating population containing 190 progeny from the “Joy x Joy white” cross, one such pair, was also used for analysis. The plant material used are exclusively cultivated varieties of the species *Euphorbia pulcherrima* which were provided by the company Selecta One. Identification of the material was conducted by Dr. Robert Boehm and Mr. Guido von Tubeuf. There were no vouchers taken and deposited.

Leaf and bract samples for DNA and RNA isolation were harvested at Selecta One, immediately frozen in liquid nitrogen, shipped on dry ice and stored at  $-80\text{ }^{\circ}\text{C}$  at the Institute for Plant Genetics of the Leibniz Universität Hannover (Hannover, Germany). Moreover, rooted cuttings of three red (‘Aurora’, ‘Aurora Jingle’ and ‘SK183’) and one pink (‘SK159 Dark Pink’) variety were irradiated with  $\gamma$ -rays (20 Gy) and further developed, and new cuttings were generated by the company Selecta One. The progeny were used for fragment analysis (please refer to section “Poinsettia genotyping and fragment analysis”).

Seeds of the *Arabidopsis thaliana* Columbia (*col-0*) genotype were available at the Institute for Plant Genetics of the Leibniz Universität Hannover; these seeds were originally obtained from the Arabidopsis information service in Frankfurt, Germany (Debener et al., 1991), and subsequently propagated in isolated greenhouse compartments. Seeds of the mutant line *tt19-8* (stock number: CS2105587) were obtained from the Arabidopsis Biological Resource Center (ABRC). Seeds were sown in Einheitserde P substrate, and seedlings were placed in long-day conditions (16 h light/8 h dark,  $22\text{ }^{\circ}\text{C}$ ) for 2 weeks to induce flowering. Seedlings of the wild-type and mutants analysed for the expression of anthocyanins were grown under supplemental red/blue LED light (GP LED production DR/B 120 LB, Philips, Germany) to stimulate anthocyanin development.

### DNA and RNA isolation

For the poinsettia samples, DNA was isolated from approximately 100 mg of leaf tissue using the NucleoSpin® Plant II Kit (Macherey–Nagel GmbH & Co., KG, Düren, Germany) according to the manufacturer’s instructions. Total RNA was isolated from approximately 100 mg of bract tissue using the

mirPremier™ miRNA Isolation Kit (Sigma-Aldrich, St. Louis, USA). For Arabidopsis samples, total RNA was isolated from approximately 50 mg of leaf tissue using the Quick-RNA Plant Kit (Zymo Research, Irvine, USA). cDNA synthesis was performed using the FastGene Scriptase Basic cDNA Kit (Nippon Genetics Europe GmbH, Düren, Germany) according to the manufacturer's recommendations. The DNA and total RNA concentrations and quality were assessed using a NanoDrop™ 2000 (Thermo Fisher Scientific, Wilmington, USA) and gel electrophoresis.

## **GST expression by RT-qPCR**

Two endogenous reference genes (Translation elongation factor 1 beta (*EF1B*) and Translation elongation factor 1 alpha (*EF1A*)) were used to normalize the *Bract1* expression data. Primer sequences are available in Additional File S4. The amplification efficiency for all primers was obtained from relative standard curves. Three independent biological replicates were used for each of the varieties. RT-qPCRs were performed using the qPCRBIO SyGreen Mix Lo-ROX Kit (Nippon Genetics Europe GmbH) according to the manufacturer's recommendations. Briefly, reactions were carried out in technical triplicates in a volume of 10 µL containing 5 µL of qPCRBIO SyGreen Mix Lo-ROX, 10 µmol of gene-specific forward and reverse primers, and 4 µL of a 1:50 cDNA dilution. RT-qPCRs were performed using a StepOne™ Real-Time PCR System (Applied Biosystems, Singapore, Singapore). The normalized relative quantity (NRQ) was calculated according to the Pfaffl equations (Pfaffl, 2001). The real-time data for this study are provided according to the Minimum Information for Publication of Quantitative Real-Time PCR Experiments guidelines (Bustin et al., 2009). Pairwise statistical analysis between each red variety and its white counterpart was performed using the Relative Expression Software Tool (REST) v2.0.13 (Pfaffl et al., 2002).

### **Bract1 gene sequencing**

The poinsettia variety 'Vintage' was used for full-length sequencing of the GST-like gene (hereafter named *Bract1*). PCRs were performed in a 50 µL reaction containing 50 ng of DNA template, 1X PrimeSTAR® Buffer (Mg<sup>2+</sup> plus), 0.2 mM each dNTP, 0.25 µM forward and reverse primers and 1.25 U of PrimeSTAR® HS DNA Polymerase (Takara Bio Inc, Kusatsu, Japan). The cycling conditions were 95 °C for 3 min; 30 cycles of 95 °C for 30 sec, 60 °C for 30 sec and 72 °C for 2 min; and a final extension of 10 min at 72 °C. The PCR products were resolved in a 1% (w/v) agarose gel by horizontal electrophoresis for 90 min at 100 V. The correct bands were excised from the gel and purified using the NucleoSpin® Gel and PCR Clean-up Kit (Macherey–Nagel) following the manufacturer's recommendations. Finally, the purified PCR fragments were sent to Eurofins Genomics (Ebersberg, Germany) for Sanger sequencing. The generated sequences were aligned using the ClustalW function in the BioEdit Sequence Alignment Editor v7.2.5 (Hall et al., 1999), and a final full-length gene sequence for *Bract1* was generated.

PCRs were performed in a 50 µL reaction containing 1 µL of undiluted cDNA, 1X PrimeSTAR® Buffer (Mg<sup>2+</sup> plus), 0.2 mM each dNTP, 0.25 µM forward and reverse primers and 1.25 U of PrimeSTAR® HS DNA Polymerase (Takara). The cycling conditions were 95 °C for 3 min; 30 cycles of 95 °C for 30 sec, 60 °C for 30 sec and 72 °C for 45 sec; and a final extension of 10 min at 72 °C. The sequencing strategy was the

same as that used for the full-length sequencing analysis. The generated sequences were aligned using the ClustalW function in the BioEdit Sequence Alignment Editor v7.2.5. Primer sequences are available in Additional File S4.

## Poinsettia genotyping and fragment analysis

To detect changes in the repeat structure of the *Bract1* gene, a genotyping approach based on the fluorescent labelling of PCR fragments (Schuelke, 2000) was applied. DNA samples were PCR amplified in a 20 µL reaction containing 50 ng of DNA template, 1X Williams buffer, 0.15 mM each dNTP, 0.0125 µM forward primer, 0.07 µM universal FAM-labelled M13 primer, 0.25 µM reverse primer and 1 U of DCSPol DNA Polymerase (DNA Cloning Service, Hamburg, Germany). The cycling conditions were 94 °C for 3 min; 24 cycles of 94 °C for 45 sec, 59 °C for 1 min and 72 °C for 1 min; 6 cycles of 94 °C for 30 sec, 52 °C for 45 sec and 72 °C for 1 min; and a final extension of 10 min at 72 °C. Fifty microliters of formamide loading dye (98% formamide, 10 mM EDTA, 0.05% pararosanilin) was added to each reaction, which was then incubated at 95 °C for 5 min. The PCR products were resolved in a 6% (w/v) acrylamide gel via vertical electrophoresis using a LI-COR Gene Reader 4200 DNA Analyser (LI-COR Biosciences, Nebraska, USA). The varieties from Table 3 and the progeny of the irradiated variety 'Aurora' were genotyped using fluorescent labelling of PCR fragments. The progeny of the irradiated varieties 'Aurora Jingle' and 'SK159 Dark Pink' were analysed by capillary electrophoresis on an ABI 3730 XL system at Microsynth AG (Balgach, Switzerland). Primer sequences are available in Additional File S4

## Phylogenetic analysis

Protein sequences for *Bract1* and 95 different poinsettia GSTs, retrieved from our previous study (Vilperte et al., 2019), were predicted with TransDecoder (Haas et al., 2013) and used for the construction of a phylogenetic tree. Moreover, protein sequences for known anthocyanin-related GSTs from other species were included in the analysis: *CkmGST3* (*Cyclamen persicum* x *Cyclamen purpurascens*, AB682678.1), *LcGST4* (*Litchi chinensis*, KT946768.1), *VvGST4* (*Vitis vinifera*, AY971515.1), *PhAN9* (*Petunia hybrida*, Y07721.1), *PpRiant1* (*Prunus persica*, KT312847.1), *PpRiant2* (*P. persica*, KT312848.1), *AtGSTF11* (*Arabidopsis thaliana*, NM\_111189.3) and *AtTT19* (*A. thaliana*, NM\_121728.4). The putative protein sequences of all the GSTs are available in Additional File S5.

Sequence alignment was performed using MUSCLE (Edgar, 2004), and the phylogenetic tree was constructed with MEGA X v10.0.5 (Kumar et al., 2018) using the maximum likelihood (ML) method with the Whelan and Goldman matrix-based model using a discrete gamma distribution (WAG + G) (Whelan and Goldman, 2001). The best model was estimated using MEGAX. The tree topology was tested via a bootstrap analysis with 1000 replicates. For better visualization of the phylogenetic tree, Tree Of Life (iTOL) software, version 4.2.3 (Letunic and Bork, 2016) (<https://itol.embl.de/>), was used.

## Plasmid construction and Agrobacterium-mediated infiltration

The coding sequence of *Bract1* was amplified from the poinsettia varieties 'Vintage' and 'Christmas Feelings Pearl' to capture both wild-type and mutated alleles. The primers used for amplification are

available in Additional File S4. The PCR fragments were inserted in the sense orientation into the *BamHI-HindIII* site of the C757pGFPU10-35 s-ocs-LH (Additional File S6) binary vector (DNA Cloning Service, Hamburg, Germany) using the In-Fusion® HD Cloning Kit (Takara Bio Inc., Kusatsu, Japan). The vector contains a GFP gene under the control of the Arabidopsis ubiquitin promoter (*ProAtUbi::GFP*) and a 35S promoter upstream of the multiple cloning site into which inserts were cloned. The final expression vectors contained either the wild-type allele (*35S::Bract1*) or the mutated allele (*35S::Bract1\_mut*). The vector containing only GFP was also used for transformation as a negative control (only *ProAtUbi::GFP*).

The expression vectors were introduced into *Agrobacterium tumefaciens* strain GV3101 via electroporation. Transformation of the *A. thaliana* mutant line *tt19-8* was performed using the floral dip method (Clough and Bent, 1998; Zhang et al., 2006). For transgenic plant selection, T<sub>0</sub> seeds were sown in soil, and GFP-expressing seedlings were selected to produce T<sub>1</sub> and subsequently T<sub>2</sub> progeny to achieve *GST* homozygosity. T<sub>2</sub> seedlings of *tt19/35S::Bract1* and *tt19/35S::Bract1\_mut* transgenic plants were used for phenotypic analysis. Non-transformed *tt19-8*, *Col-0* and *tt19/ProAtUbi::GFP* seedlings were used as controls. Seedlings were placed under red light to stimulate anthocyanin biosynthesis. To confirm correct *GST* integration, 10 independent biological replicates, each representing an independently selected transgenic line from the T<sub>2</sub> progeny of *tt19/35S::Bract1* and *tt19/35S::Bract1\_mut* transgenic plants, were used for *GST* genotyping. The protocol was the same as that used in the section “*Poinsettia* genotyping and fragment analysis”.

## Analysis of the GST repeat

To understand the origin of the microsatellite-like repeat in the *Bract1* gene, orthologous *GST* genes from Euphorbiaceae species were retrieved by BLASTN against the Euphorbiaceae (taxid: 3977) nucleotide database. *GST*-like genes from the Euphorbiaceae species *Ricinus communis* (XM\_002532928.3), *Manihot esculenta* (XM\_021748071.1), *Jatropha curcas* (XM\_012219312.2), *Hevea brasiliensis* (XM\_021787187.1), *Euphorbia esula* (PJAE01736713.1) and *Euphorbia pекinensis* (Cao et al., 2017; Vilperte et al., 2019), as well the anthocyanin-related *GSTs* *CkmGST3* (*C. persicum* x *C. purpurascens*, AB682678.1), *LcGST4* (*L. chinensis*, KT946768.1), *VvGST4* (*V. vinifera*, AY971515.1), *PhAN9* (*P. hybrida*, Y07721.1), *PpRiant1* (*P. persica*, KT312847.1), *PpRiant2* (*P. persica*, KT312848.1) and *AtTT19* (*A. thaliana*, NM\_121728.4) were used to construct a phylogenetic tree.

Sequence alignment was performed using MUSCLE (Edgar, 2004), and the phylogenetic tree was constructed with MEGA X v10.0.5 (Kumar et al., 2018) using the ML method with the Tamura 3-parameter matrix-based model (Tamura, 1992) using a discrete gamma distribution with invariant sites (T92 + G + I). The best model was estimated using MEGAX. The tree topology was tested via a bootstrap analysis with 1000 replicates. For better visualization of the phylogenetic tree, Tree Of Life (iTOL) software, version 4.2.3 (Letunic and Bork, 2016) (<https://itol.embl.de/>), was used.

## Abbreviations

ABC

ATP binding cassette transporter  
CaMV  
Cauliflower mosaic virus  
cDNA  
Complementary DNA  
CDS  
Coding sequence  
DNA  
Deoxyribonucleic acid  
DSB  
Double-strand break  
EMS  
ethyl-methanesulphonate  
ER  
Endoplasmic reticulum  
GFP  
Green fluorescent protein  
GST  
Glutathione S-transferase  
GUS  
 $\beta$ -glucuronidase protein  
Gy  
Gray (unit of ionizing radiation)  
Indel  
Insertion and deletion variations  
kDA  
Kilo Dalton  
MATE  
Multi-antimicrobial extrusion protein  
MBW complex  
MYB-bHLH-WD40 complex  
mRNA  
Messenger RNA  
NRQ  
Normalized relative quantity  
PCR  
Polymerase chain reaction  
RNA  
Ribonucleic acid  
RT-qPCR

Quantitative reverse transcription PCR  
SNP  
Single nucleotide polymorphism  
SSR  
Single sequence repeat  
tt19  
Transparent testa 19 Arabidopsis mutant

## **Declarations**

### **Ethics approval and consent to participate**

Not applicable.

### **Consent for publication**

Not applicable.

### **Availability of data and material**

**All data generated or analysed during this study are included in this published article and its supplementary information files.**

### **Competing interest**

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

### **Acknowledgments**

We would like to thank Guido von Tubeuf for providing the poinsettia varieties.

### **Author contribution**

RB, TD and VV conceived the research. VV and TD designed the experiments. VV conducted the laboratory and bioinformatics analyses. RB provide the capillary electrophoresis results. VV and TD wrote the manuscript. All authors revised and approved the final manuscript.

### **Funding**

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 675657 Flower Power. This funding body had no role in the design of the study, collection, analysis, or interpretation of data, or in writing the manuscript. The publication of this article was funded by the Open Access fund of Leibniz Universität Hannover.

## References

1. Agrawal, G. K., Jwa, N.-S., and Rakwal, R. (2002). A pathogen-induced novel rice (*Oryza sativa* L.) gene encodes a putative protein homologous to type II glutathione S-transferases. *Plant Sci.* 163, 1153–1160. doi:10.1016/S0168-9452(02)00331-X.
2. Alfenito, M. R., Souer, E., Goodman, C. D., Buell, R., Mol, J., Koes, R., et al. Functional Complementation of Anthocyanin Sequestration in the Vacuole by Widely Divergent Glutathione S-Transferases. 16.
3. Asen, S. (1958). Anthocyanins in Bracts of *Euphorbia Pulcherrima* as Revealed by Paper Chromatographic and Spectrophotometric Methods. *Plant Physiol.* 33, 14–7. doi:10.1104/pp.33.1.14.
4. Bassham, D. C. (2015). Pigments on the move. *Nature* 526, 644–645. doi:10.1038/526644a.
5. Ben-Simhon, Z., Judeinstein, S., Trainin, T., Harel-Beja, R., Bar-Ya'akov, I., Borochoy-Neori, H., et al. (2015). A “White” Anthocyanin-less Pomegranate (*Punica granatum* L.) Caused by an Insertion in the Coding Region of the Leucoanthocyanidin Dioxygenase (LDOX; ANS) Gene. *PLOS ONE* 10, e0142777. doi:10.1371/journal.pone.0142777.
6. Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., et al. (2009). The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin. Chem.* 55, 611–622. doi:10.1373/clinchem.2008.112797.
7. Cao, X., Zhang, F., Yuan, B., Meng, L., Yang, X., and Jiang, J. (2017). De novo transcriptome sequencing and analysis of *Euphorbia pekinensis* Rupr. and identification of genes involved in diterpenoid biosynthesis. *Plant Gene* 12, 33–42. doi:10.1016/j.plgene.2017.07.001.
8. Cheng, J., Liao, L., Zhou, H., Gu, C., Wang, L., and Han, Y. (2015). A small indel mutation in an anthocyanin transporter causes variegated colouration of peach flowers. *J. Exp. Bot.* 66, 7227–7239. doi:10.1093/jxb/erv419.
9. Clarke, J. L., Klemsdal, S. S., Floistad, E., Hvoslef-Eide, A. K., Haugslie, S., Moe, R., et al. (2006). Genetic Engineering of Poinsettia with the Aim of Enhancing Its Resistance to Poinsettia Mosaic Virus. *Acta Hort.*, 321–325. doi:10.17660/ActaHortic.2006.722.40.
10. Clarke, J. L., Spetz, C., Haugslie, S., Xing, S., Dees, M. W., Moe, R., et al. (2008). Agrobacterium tumefaciens-mediated transformation of poinsettia, *Euphorbia pulcherrima*, with virus-derived hairpin RNA constructs confers resistance to Poinsettia mosaic virus. *Plant Cell Rep.* 27, 1027–38. doi:10.1007/s00299-008-0526-9.
11. Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*: Floral dip transformation of *Arabidopsis*. *Plant J.* 16, 735–743. doi:10.1046/j.1365-313x.1998.00343.x.
12. Debener, T., Lehnackers, H., Arnold, M., and Dangl, J. L. (1991). Identification and molecular mapping of a single *Arabidopsis thaliana* locus determining resistance to a phytopathogenic *Pseudomonas syringae* isolate. *Plant J. Cell Mol. Biol.* 1, 289–302. doi:10.1046/j.1365-313X.1991.t01-7-00999.x.

13. Dixon, D. P., and Edwards, R. (2010). Glutathione Transferases. *Arab. Book* 8, e0131. doi:10.1199/tab.0131.
14. Dixon, D. P., Laphorn, A., and Edwards, R. (2002). Plant glutathione transferases. *Genome Biol.* 3, reviews3004.1. doi:10.1186/gb-2002-3-3-reviews3004.
15. Dixon, D. P., Skipsey, M., and Edwards, R. (2010). Roles for glutathione transferases in plant secondary metabolism. *Phytochemistry* 71, 338–350. doi:10.1016/j.phytochem.2009.12.012.
16. Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in Arabidopsis. *Trends Plant Sci.* 15, 573–581. doi:10.1016/j.tplants.2010.06.005.
17. Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi:10.1093/nar/gkh340.
18. Edwards, R., Dixon, D. P., and Walbot, V. (2000). Plant glutathione S-transferases: enzymes with multiple functions in sickness and in health. *Trends Plant Sci.* 5, 193–198. doi:10.1016/s1360-1385(00)01601-0.
19. Ellegren, H. (2004). Microsatellites: simple sequences with complex evolution. *Nat. Rev. Genet.* 5, 435–445. doi:10.1038/nrg1348.
20. Esnault, M.-A., Legue, F., and Chenal, C. (2010). Ionizing radiation: Advances in plant response. *Environ. Exp. Bot.* 68, 231–237. doi:10.1016/j.envexpbot.2010.01.007.
21. Francisco, R. M., Regalado, A., Ageorges, A., Burla, B. J., Bassin, B., Eisenach, C., et al. (2013). ABCC1, an ATP binding cassette protein from grape berry, transports anthocyanidin 3-O-Glucosides. *Plant Cell* 25, 1840–1854. doi:10.1105/tpc.112.102152.
22. Golubov, A., Yao, Y., Maheshwari, P., Bilichak, A., Boyko, A., Belzile, F., et al. (2010). Microsatellite Instability in Arabidopsis Increases with Plant Development1[W][OA]. *Plant Physiol.* 154, 1415–1427. doi:10.1104/pp.110.162933.
23. Gomez, C., Conejero, G., Torregrosa, L., Cheynier, V., Terrier, N., and Ageorges, A. (2011). In vivo grapevine anthocyanin transport involves vesicle-mediated trafficking and the contribution of anthoMATE transporters and GST: Anthocyanin trafficking in grapevine. *Plant J.* 67, 960–970. doi:10.1111/j.1365-313X.2011.04648.x.
24. Goodman, C. D., Casati, P., and Walbot, V. (2004). A Multidrug Resistance–Associated Protein Involved in Anthocyanin Transport in *Zea mays*. *Plant Cell* 16, 1812–1826. doi:10.1105/tpc.022574.
25. Greene, E. A., Codomo, C. A., Taylor, N. E., Henikoff, J. G., Till, B. J., Reynolds, S. H., et al. (2003). Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. *Genetics* 164, 731–740.
26. Grotewold, E. ed. (2006). *The science of flavonoids*. New York: Springer.
27. Grotewold, E., and Davies, K. (2008). Trafficking and Sequestration of Anthocyanins. *Nat. Prod. Commun.* 3, 1934578X0800300. doi:10.1177/1934578X0800300806.

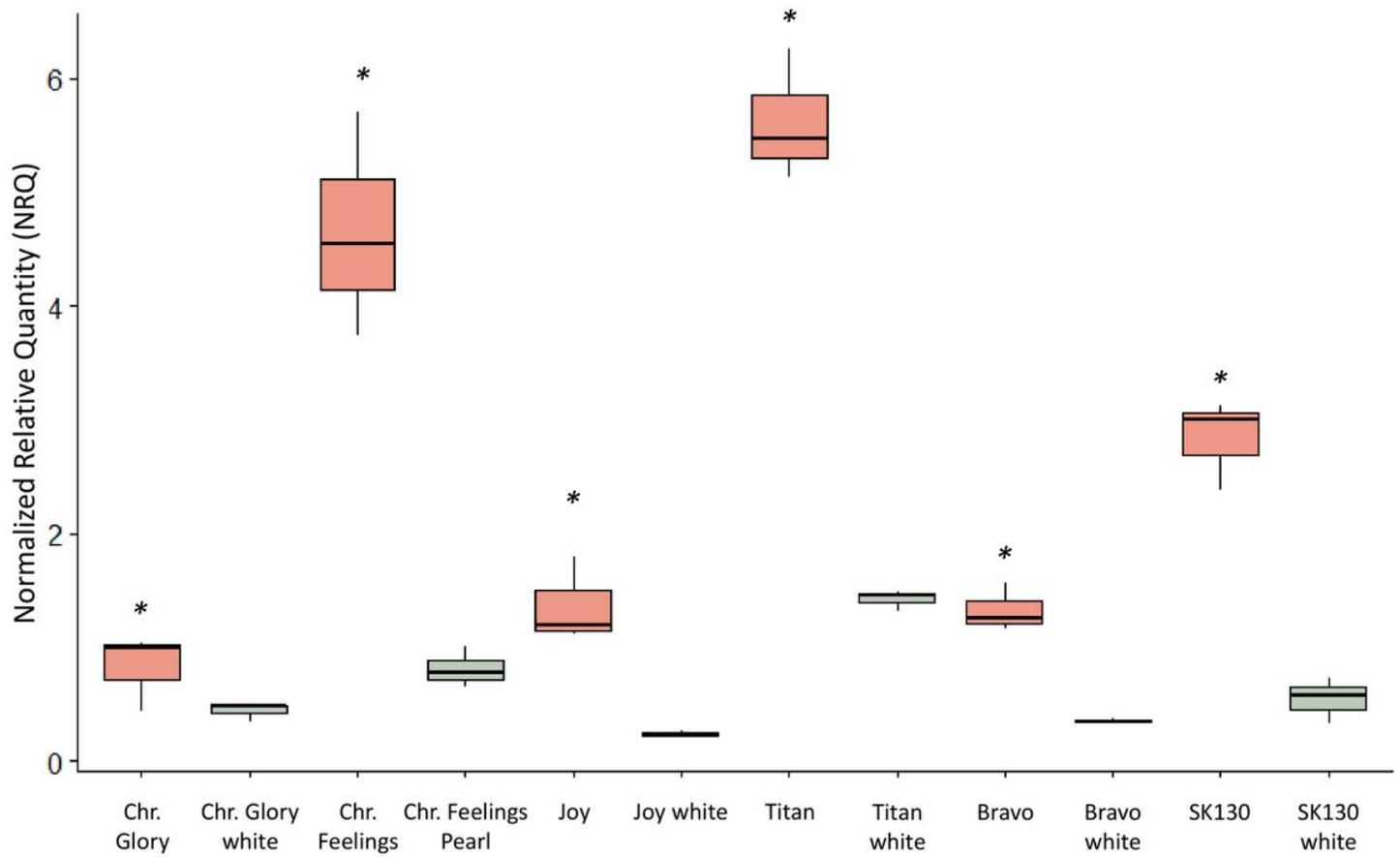
28. Gu, Z., Chen, H., Yang, R., and Ran, M. (2018). Identification of DFR as a promoter of anthocyanin accumulation in poinsettia (*Euphorbia pulcherrima*, Willd. ex Klotzsch) bracts under short-day conditions. *Sci. Hortic.* 236, 158–165. doi:10.1016/j.scienta.2018.03.032.
29. Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512. doi:10.1038/nprot.2013.084.
30. Hall, T. A., Hall, T., Hall, T. A., Hall, T., Hall, T. A., Hall, T. A., et al. (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. doi:10.14601/Phytopathol\_Mediterr-14998u1.29.
31. He, G., Guan, C.-N., Chen, Q.-X., Gou, X.-J., Liu, W., Zeng, Q.-Y., et al. (2016). Genome-Wide Analysis of the Glutathione S-Transferase Gene Family in *Capsella rubella*: Identification, Expression, and Biochemical Functions. *Front. Plant Sci.* 7. doi:10.3389/fpls.2016.01325.
32. Hu, B., Zhao, J., Lai, B., Qin, Y., Wang, H., and Hu, G. (2016). LcGST4 is an anthocyanin-related glutathione S-transferase gene in *Litchi chinensis* Sonn. *Plant Cell Rep.* 35, 831–843. doi:10.1007/s00299-015-1924-4.
33. Islam, S., Sajib, S. D., Jui, Z. S., Arabia, S., Islam, T., and Ghosh, A. (2019). Genome-wide identification of glutathione S-transferase gene family in pepper, its classification, and expression profiling under different anatomical and environmental conditions. *Sci. Rep.* 9, 9101. doi:10.1038/s41598-019-45320-x.
34. Jiang, S., Chen, M., He, N., Chen, X., Wang, N., Sun, Q., et al. (2019). MdGSTF6, activated by MdMYB1, plays an essential role in anthocyanin accumulation in apple. *Hortic. Res.* 6, 40. doi:10.1038/s41438-019-0118-6.
35. Jo, Y. D., and Kim, J.-B. (2019). Frequency and Spectrum of Radiation-Induced Mutations Revealed by Whole-Genome Sequencing Analyses of Plants. *Quantum Beam Sci.* 3, 7. doi:10.3390/qubs3020007.
36. Kazama, Y., Hirano, T., Saito, H., Liu, Y., Ohbu, S., Hayashi, Y., et al. (2011). Characterization of highly efficient heavy-ion mutagenesis in *Arabidopsis thaliana*. *BMC Plant Biol.* 11, 161. doi:10.1186/1471-2229-11-161.
37. Kitamura, S., Akita, Y., Ishizaka, H., Narumi, I., and Tanaka, A. (2012). Molecular characterization of an anthocyanin-related glutathione S-transferase gene in cyclamen. *J. Plant Physiol.* 169, 636–642. doi:10.1016/j.jplph.2011.12.011.
38. Kitamura, S., Shikazono, N., and Tanaka, A. (2004). *TRANSPARENT TESTA 19* is involved in the accumulation of both anthocyanins and proanthocyanidins in *Arabidopsis*. *Plant J.* 37, 104–114. doi:10.1046/j.1365-313X.2003.01943.x.
39. Koseki, M., Goto, K., Masuta, C., and Kanazawa, A. (2005). The Star-type Color Pattern in *Petunia hybrida* 'Red Star' Flowers is Induced by Sequence-Specific Degradation of Chalcone Synthase RNA. *Plant Cell Physiol.* 46, 1879–1883. doi:10.1093/pcp/pci192.
40. Kou, M., Liu, Y., Li, Z., Zhang, Y., Tang, W., Yan, H., et al. (2019). A novel glutathione S-transferase gene from sweetpotato, *IbGSTF4*, is involved in anthocyanin sequestration. *Plant Physiol. Biochem.* 135,

- 395–403. doi:10.1016/j.plaphy.2018.12.028.
41. Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi:10.1093/molbev/msy096.
42. Lallement, P.-A., Brouwer, B., Keech, O., Hecker, A., and Rouhier, N. (2014). The still mysterious roles of cysteine-containing glutathione transferases in plants. *Front. Pharmacol.* 5. doi:10.3389/fphar.2014.00192.
43. Larsen, E. S., Alfenito, M. R., Briggs, W. R., and Walbot, V. (2003). A carnation anthocyanin mutant is complemented by the glutathione S-transferases encoded by maize Bz2 and petunia An9. *Plant Cell Rep.* 21, 900–904. doi:10.1007/s00299-002-0545-x.
44. Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44, W242-245. doi:10.1093/nar/gkw290.
45. Licciardello, C., D'Agostino, N., Traini, A., Recupero, G., Frusciante, L., and Chiusano, M. (2014). Characterization of the glutathione S-transferase gene family through ESTs and expression analyses within common and pigmented cultivars of *Citrus sinensis* (L.) Osbeck. *BMC Plant Biol.* 14, 39. doi:10.1186/1471-2229-14-39.
46. Liu, Y., Qi, Y., Zhang, A., Wu, H., Liu, Z., and Ren, X. (2019). Molecular cloning and functional characterization of AcGST1, an anthocyanin-related glutathione S-transferase gene in kiwifruit (*Actinidia chinensis*). *Plant Mol. Biol.* 100, 451–465. doi:10.1007/s11103-019-00870-6.
47. Liu, Y., Tikunov, Y., Schouten, R. E., Marcelis, L. F. M., Visser, R. G. F., and Bovy, A. (2018). Anthocyanin Biosynthesis and Degradation Mechanisms in Solanaceous Vegetables: A Review. *Front. Chem.* 6, 52. doi:10.3389/fchem.2018.00052.
48. Luo, H., Dai, C., Li, Y., Feng, J., Liu, Z., and Kang, C. (2018). Reduced Anthocyanins in Petioles codes for a GST anthocyanin transporter that is essential for the foliage and fruit coloration in strawberry. *J. Exp. Bot.* 69, 2595–2608. doi:10.1093/jxb/ery096.
49. Luo, P., Ning, G., Wang, Z., Shen, Y., Jin, H., Li, P., et al. (2016). Disequilibrium of Flavonol Synthase and Dihydroflavonol-4-Reductase Expression Associated Tightly to White vs. Red Color Flower Formation in Plants. *Front. Plant Sci.* 6. doi:10.3389/fpls.2015.01257.
50. Marrs, K. A., Alfenito, M. R., Lloyd, A. M., and Walbot, V. (1995). A glutathione S-transferase involved in vacuolar transfer encoded by the maize gene Bronze-2. *Nature* 375, 397–400. doi:10.1038/375397a0.
51. Morita, R., Kusaba, M., Iida, S., Yamaguchi, H., Nishio, T., and Nishimura, M. (2009). Molecular characterization of mutations induced by gamma irradiation in rice. *Genes Genet. Syst.* 84, 361–370. doi:10.1266/ggs.84.361.
52. Morita, Y., Saito, R., Ban, Y., Tanikawa, N., Kuchitsu, K., Ando, T., et al. (2012). Tandemly arranged chalcone synthase A genes contribute to the spatially regulated expression of siRNA and the natural

- bicolor floral phenotype in *Petunia hybrida*: Endogenous siRNA and bicolored flowers of *Petunia*. *Plant J.* 70, 739–749. doi:10.1111/j.1365-313X.2012.04908.x.
53. Moustaka, J., Panteris, E., Adamakis, I.-D. S., Tanou, G., Giannakoula, A., Eleftheriou, E. P., et al. (2018). High anthocyanin accumulation in poinsettia leaves is accompanied by thylakoid membrane unstacking, acting as a photoprotective mechanism, to prevent ROS formation. *Environ. Exp. Bot.* 154, 44–55. doi:10.1016/j.envexpbot.2018.01.006.
54. Mueller, L. A., Goodman, C. D., Silady, R. A., and Walbot, V. (2000). AN9, a *Petunia* Glutathione *S*-Transferase Required for Anthocyanin Sequestration, Is a Flavonoid-Binding Protein. *Plant Physiol.* 123, 1561–1570. doi:10.1104/pp.123.4.1561.
55. Nitarska, D., Stefanini, C., Haselmair-Gosch, C., Miosic, S., Walliser, B., Mikulic-Petkovsek, M., et al. (2018). The rare orange-red colored *Euphorbia pulcherrima* cultivar “Harvest Orange” shows a nonsense mutation in a flavonoid 3'-hydroxylase allele expressed in the bracts. *BMC Plant Biol.* 18, 216. doi:10.1186/s12870-018-1424-0.
56. Pearson, C. E., Edamura, K. N., and Cleary, J. D. (2005). Repeat instability: mechanisms of dynamic mutations. *Nat. Rev. Genet.* 6, 729–742. doi:10.1038/nrg1689.
57. Pérez-Díaz, R., Madrid-Espinoza, J., Salinas-Cornejo, J., González-Villanueva, E., and Ruiz-Lara, S. (2016). Differential Roles for VviGST1, VviGST3, and VviGST4 in Proanthocyanidin and Anthocyanin Transport in *Vitis vinifera*. *Front. Plant Sci.* 7. doi:10.3389/fpls.2016.01166.
58. Petroni, K., and Tonelli, C. (2011). Recent advances on the regulation of anthocyanin synthesis in reproductive organs. *Plant Sci.* 181, 219–229. doi:10.1016/j.plantsci.2011.05.009.
59. Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29, 45e–445. doi:10.1093/nar/29.9.e45.
60. Pfaffl, M. W., Horgan, G. W., and Dempfle, L. (2002). Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res.* 30, 10.
61. Pomar, F., and Ros Barceló, A. (2007). Are red leaves photosynthetically active? *Biol. Plant.* 51, 799–800. doi:10.1007/s10535-007-0164-z.
62. Poustka, F., Irani, N. G., Feller, A., Lu, Y., Pourcel, L., Frame, K., et al. (2007). A Trafficking Pathway for Anthocyanins Overlaps with the Endoplasmic Reticulum-to-Vacuole Protein-Sorting Route in *Arabidopsis* and Contributes to the Formation of Vacuolar Inclusions. *Plant Physiol.* 145, 1323–1335. doi:10.1104/pp.107.105064.
63. Saito, K., Yonekura-Sakakibara, K., Nakabayashi, R., Higashi, Y., Yamazaki, M., Tohge, T., et al. (2013). The flavonoid biosynthetic pathway in *Arabidopsis*: Structural and genetic diversity. *Plant Physiol. Biochem.* 72, 21–34. doi:10.1016/j.plaphy.2013.02.001.
64. Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nat. Biotechnol.* 18, 233–234. doi:10.1038/72708.
65. Schum, A., and Preil, W. (1998). “Induced Mutations in Ornamental Plants,” in *Somaclonal Variation and Induced Mutations in Crop Improvement* Current Plant Science and Biotechnology in Agriculture,

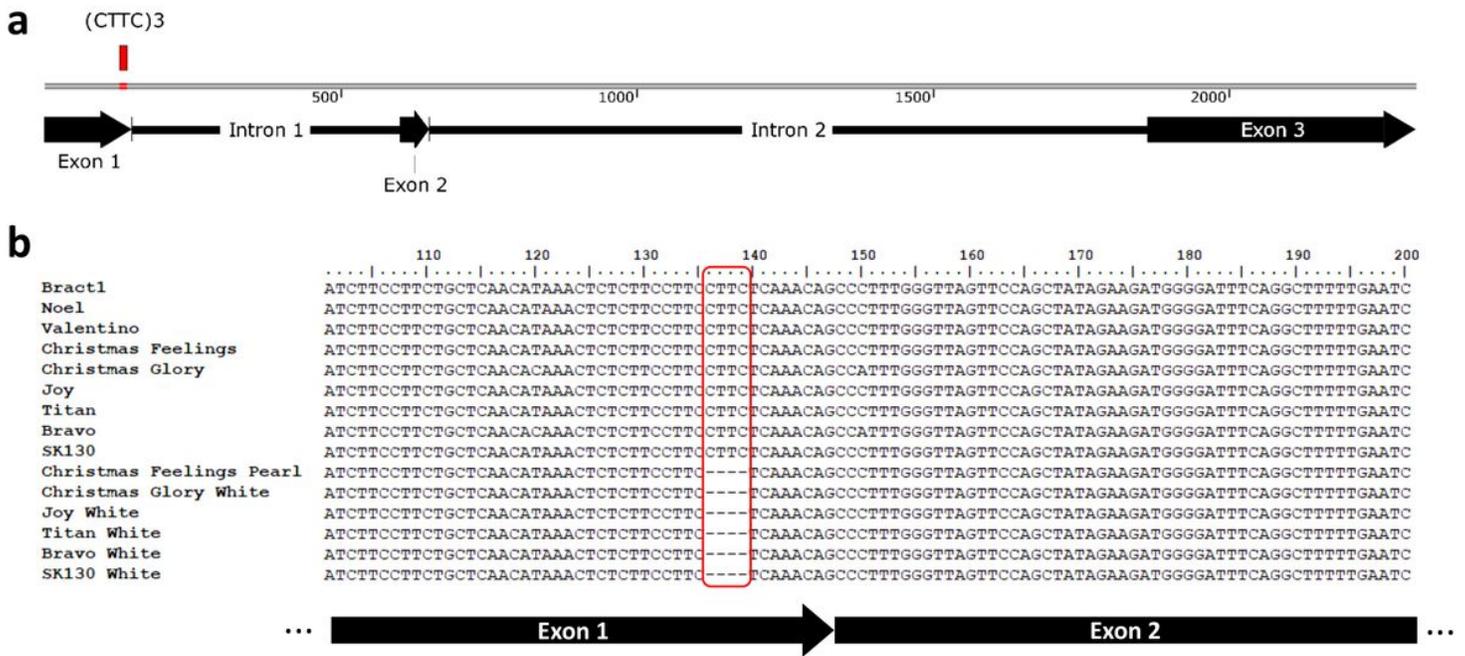
- eds. S. M. Jain, D. S. Brar, and B. S. Ahloowalia (Dordrecht: Springer Netherlands), 333–366. doi:10.1007/978-94-015-9125-6\_17.
66. Slatnar, A., Mikulic-Petkovsek, M., Veberic, R., Stampar, F., and Schmitzer, V. (2013). Anthocyanin and chlorophyll content during poinsettia bract development. *Sci. Hortic.* 150, 142–145. doi:10.1016/j.scienta.2012.10.014.
67. Sun, Y., Li, H., and Huang, J.-R. (2012). Arabidopsis TT19 Functions as a Carrier to Transport Anthocyanin from the Cytosol to Tonoplasts. *Mol. Plant* 5, 387–400. doi:10.1093/mp/ssr110.
68. Tamura, K. (1992). Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. *Mol. Biol. Evol.* 9, 678–687. doi:10.1093/oxfordjournals.molbev.a040752.
69. USDA (2018). Floriculture Crops 2018 Summary. National Agricultural Statistics Service. Available at: [https://www.nass.usda.gov/Publications/Todays\\_Reports/reports/floran19.pdf](https://www.nass.usda.gov/Publications/Todays_Reports/reports/floran19.pdf).
70. van der Ploeg, R. (2018). FloraCulture International. *Int. Assoc. Hortic. Prod. IAHP*. Available at: <https://www.floraculture.eu/digital-magazine/fci-november-december-2018/> [Accessed March 18, 2020].
71. Vik, N. I., Hvoslef-Eide, A. K., Gjerde, H., and Bakke, K. (2001). Stable transformation of poinsettia via electrophoresis. *Acta Hortic.*, 101–103. doi:10.17660/ActaHortic.2001.560.12.
72. Vilperte, V., Lucaciu, C. R., Halbwirth, H., Boehm, R., Rattei, T., and Debener, T. (2019). Hybrid de novo transcriptome assembly of poinsettia (*Euphorbia pulcherrima* Willd. Ex Klotsch) bracts. *BMC Genomics* 20, 900. doi:10.1186/s12864-019-6247-3.
73. Whelan, S., and Goldman, N. (2001). A General Empirical Model of Protein Evolution Derived from Multiple Protein Families Using a Maximum-Likelihood Approach. *Mol. Biol. Evol.* 18, 691–699. doi:10.1093/oxfordjournals.molbev.a003851.
74. Yao, Y., and Kovalchuk, I. (2011). Abiotic stress leads to somatic and heritable changes in homologous recombination frequency, point mutation frequency and microsatellite stability in *Arabidopsis* plants. *Mutat. Res.* 707, 61–66. doi:10.1016/j.mrfmmm.2010.12.013.
75. Zhang, X., Henriques, R., Lin, S.-S., Niu, Q.-W., and Chua, N.-H. (2006). Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nat. Protoc.* 1, 641–646. doi:10.1038/nprot.2006.97.
76. Zhao, D., and Tao, J. (2015). Recent advances on the development and regulation of flower color in ornamental plants. *Front. Plant Sci.* 6. doi:10.3389/fpls.2015.00261.
77. Zhao, J. (2015). Flavonoid transport mechanisms: how to go, and with whom. *Trends Plant Sci.* 20, 576–585. doi:10.1016/j.tplants.2015.06.007.
78. Zhao, J., and Dixon, R. A. (2010). The ‘ins’ and ‘outs’ of flavonoid transport. *Trends Plant Sci.* 15, 72–80. doi:10.1016/j.tplants.2009.11.006.

## Figures



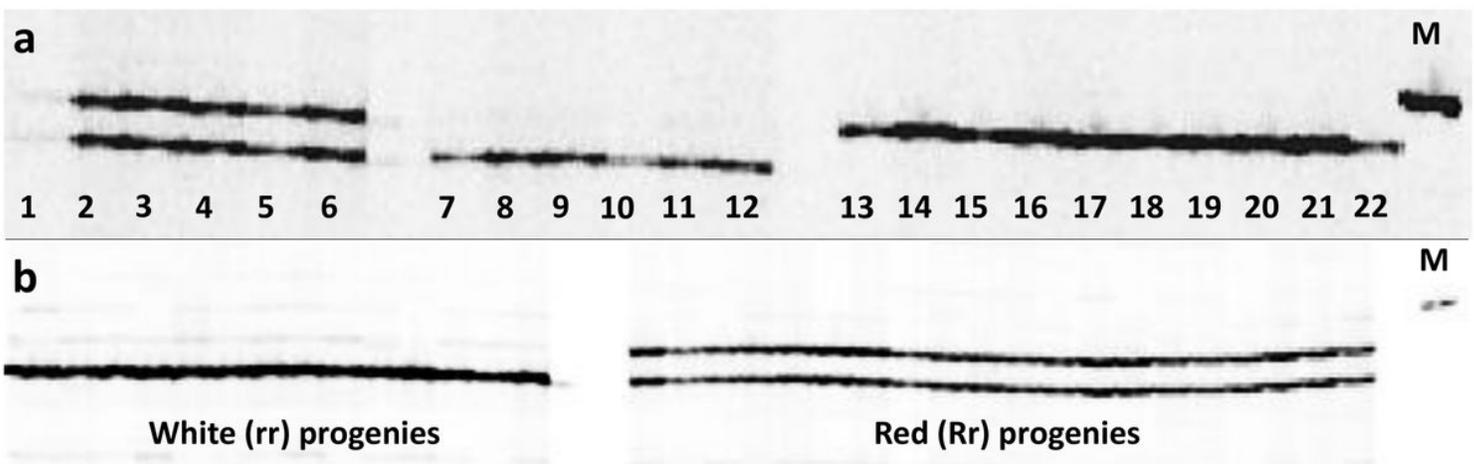
**Figure 1**

RT-qPCR of Bract1 for six pairs of red-bracted poinsettia varieties and their independently generated white mutants. The normalized relative quantity (NRQ) was calculated according to the Pfaffl equations (Pfaffl, 2001) and using the 'Chr. Glory' variety as a reference sample. The '\*' symbol indicates significant differences calculated with REST software between red and white pairs at  $p \leq 0.05$ .



**Figure 2**

Characterization of the anthocyanin-related GST gene (Bract1) in *Euphorbia pulcherrima*. A) Schematic representation of the full-length sequence (2,314 bp) of Bract1 in the 'Vintage' variety. Black arrows represent the exonic regions. Black lines represent the intronic regions. The red square represents the location of the tetranucleotide motif SSR locus (CTTC<sup>3</sup>). B) Nucleotide alignment of the Bract1 CDS for 14 red- and white-bracted poinsettia varieties. The figure shows a 100 bp region of the CDS in which a 4 bp deletion (red box) is observed only in the white varieties. Black arrows below the sequences show the location of the sequences in each exon. The first sequence corresponds to Bract1 from the 'Vintage' variety and was used as a reference for the alignment.

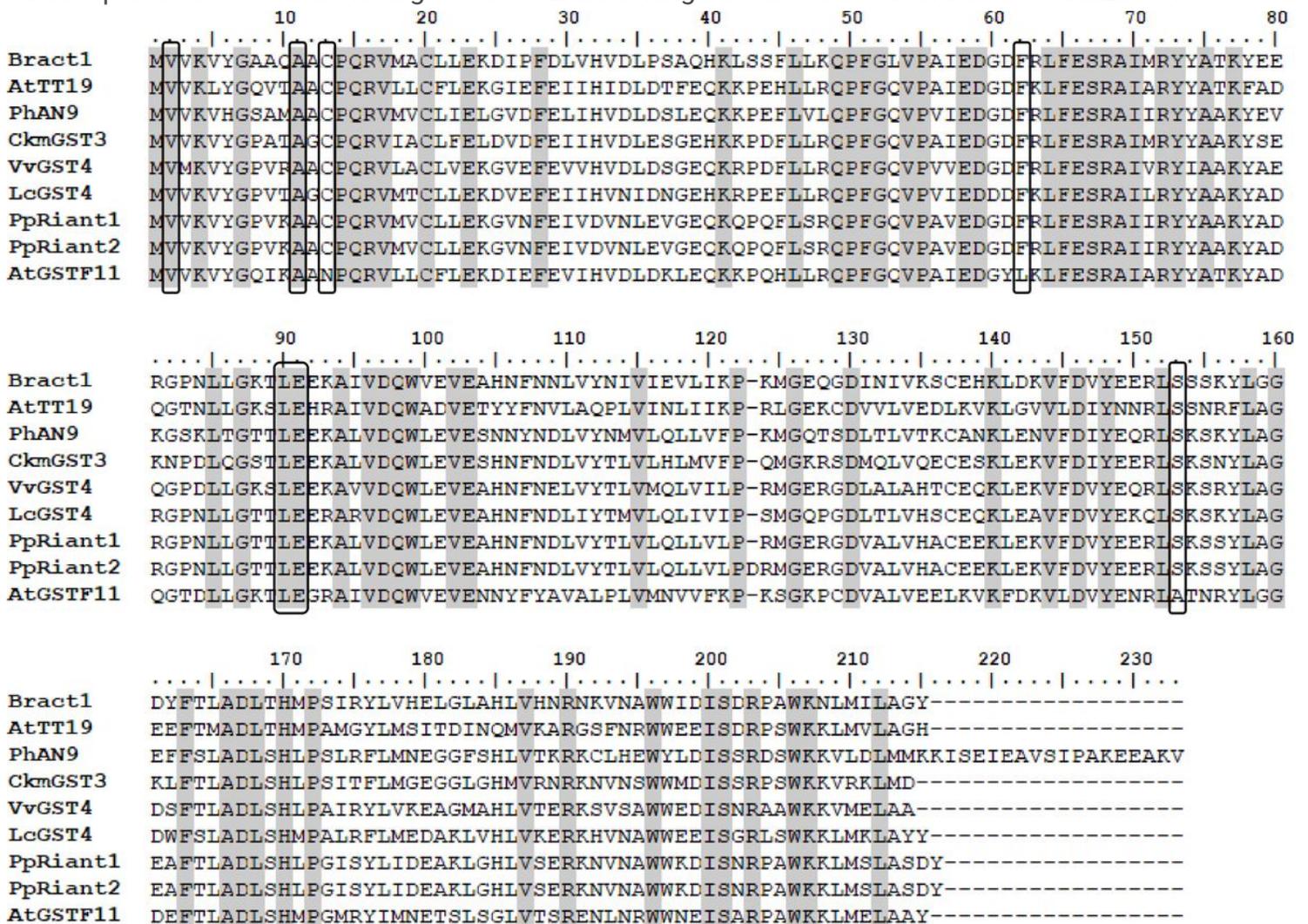


**Figure 3**

PCR amplification of the tetranucleotide motif SSR locus (CTTC)<sup>3</sup> from the Bract1 gene. A) Band patterns from the amplified PCR fragments for Bract1 in 22 red- and white-bracted poinsettia varieties. Samples 1-

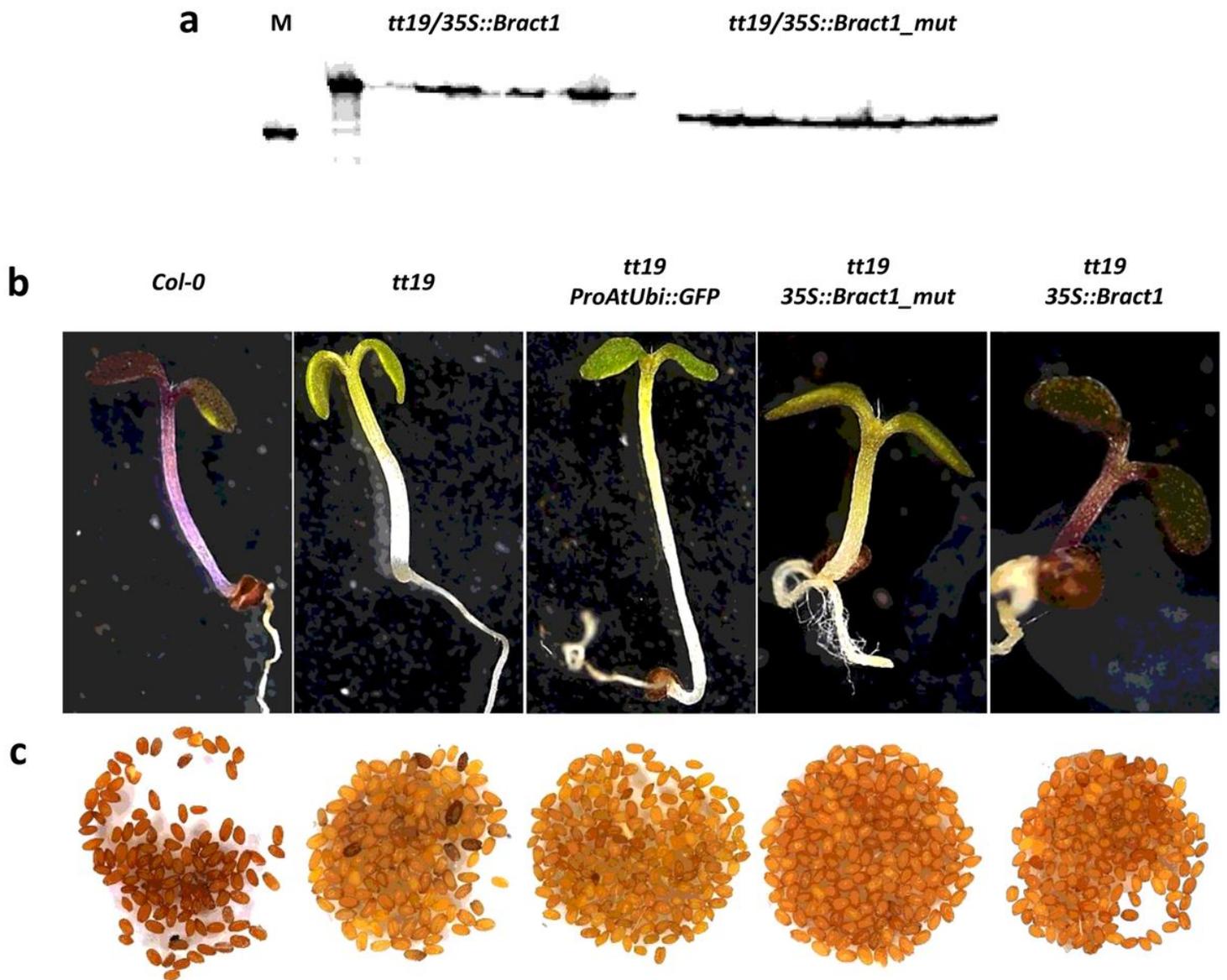


matrix-based model was used to calculate the phylogenetic tree. Phylogenetic testing was performed using the bootstrap method with 1000 replicates, which are depicted as triangles, where the smallest value represents 1.3% and the largest 100%. Branch lengths were omitted for better visualization.



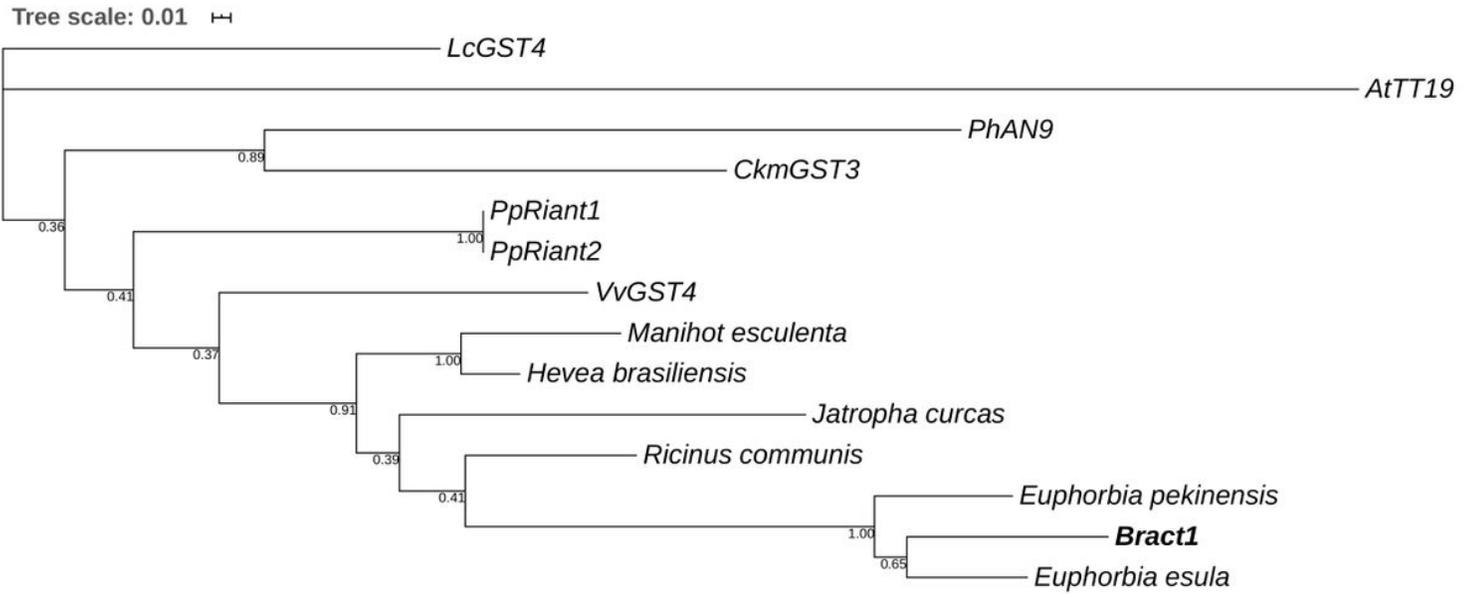
**Figure 5**

Protein sequence alignment of BRACT1 and anthocyanin-related GSTs from other plant species. The numbers in the alignments indicate the amino acid positions, and black boxes show amino acids that are known to be conserved in anthocyanin-related GSTs (Kitamura et al., 2012; Hu et al., 2016; Kou et al., 2019). Sequences were aligned using the ClustalW function in the BioEdit Sequence Alignment Editor v7.2.5.



**Figure 6**

Functional complementation of the Arabidopsis *tt19* mutant with the *Bract1* gene. A) Genotyping of 10 independent biological replicates from the T2 progeny of *tt19/35S::Bract1* and *tt19/35S::Bract1\_mut* transgenic plants. Figure was cropped for better visualization. Full length figure is available in Additional File S2. B) Phenotypes of seedlings (14 days old) and C) mature seeds of *Col-0* and *tt19* and the transgenic lines *tt19/ProAtUbi::GFP*, *tt19/35S::Bract1* and *tt19/35S::Bract1\_mut* in the *tt19* background.



**Figure 7**

Phylogenetic tree of anthocyanin-related GSTs and GST-like genes from Euphorbiaceae species. CDS nucleotide sequences were aligned using MUSCLE. The maximum likelihood (ML) method based on the T92 matrix-based model was used to calculate the phylogenetic tree. Phylogeny testing was performed using the bootstrap method with 1000 replicates, which are depicted next to the branches.

<i>Bract1</i>	106	CCT	TCT	GCT	CAA	CAT	AAA	CTC	<b>TCT</b>	<b>TCC</b>	<b>TTC</b>	<b>CTT</b>	<b>CTC</b>	AAA	CAG	CCC	TTT	GGG	TTA	GTT	CCA	165
	36	Pro	Ser	Ala	Gln	His	Lys	Leu	Ser	Ser	Phe	Leu	Leu	Lys	Gln	Pro	Phe	Gly	Leu	Val	Pro	55
<i>Euphorbia esula</i>	106	CAT	TCT	GCT	CAA	CAT	AAA	CTC	<b>CCT</b>	<b>TCT</b>	<b>TAC</b>	<b>CTT</b>	<b>TGC</b>	AAA	CAG	CCC	TTT	GGG	TTA	GTT	CCA	165
	36	His	Ser	Ala	Gln	His	Lys	Leu	Pro	Ser	Tyr	Leu	Cys	Lys	Gln	Pro	Phe	Gly	Leu	Val	Pro	55
<i>Euphorbia pekinensis</i>	106	CAC	TCT	TCT	CAC	CAC	AAA	CTC	<b>CCT</b>	<b>TCC</b>	<b>TTT</b>	<b>CTT</b>	<b>CTC</b>	AAA	CAG	CCT	TTT	GGG	CTC	GTT	CCG	165
	36	His	Ser	Ser	His	His	Lys	Leu	Pro	Ser	Phe	Leu	Leu	Lys	Gln	Pro	Phe	Gly	Leu	Val	Pro	55
<i>Ricinus communis</i>	106	GAT	TCT	GGA	GAG	CAT	AAG	CGA	<b>TCT</b>	<b>GAA</b>	<b>TTC</b>	<b>CTT</b>	<b>CTC</b>	AAA	CAG	CCA	TTT	GGG	CAA	GTC	CCG	165
	36	Asp	Ser	Gly	Glu	His	Lys	Arg	Ser	Glu	Phe	Leu	Leu	Lys	Gln	Pro	Phe	Gly	Gln	Val	Pro	55
<i>Jatropha curcas</i>	106	AAA	TCC	GGG	GAC	CAT	AAG	CAT	<b>CCT</b>	<b>GAC</b>	<b>TTC</b>	<b>CTT</b>	<b>CTC</b>	AAA	CAG	CCG	TTT	GGG	CAA	GTT	CCG	165
	36	Lys	Ser	Gly	Asp	His	Lys	His	Pro	Asp	Phe	Leu	Leu	Lys	Gln	Pro	Phe	Gly	Gln	Val	Pro	55
<i>Hevea brasiliensis</i>	106	GCT	TCT	GGA	GAG	CAC	AAA	CGA	<b>CCT</b>	<b>GAC</b>	<b>TTC</b>	<b>CTT</b>	<b>CTC</b>	AAA	CAG	CCA	TTT	GGG	CAA	GTT	CCT	165
	36	Ala	Ser	Gly	Glu	His	Lys	Arg	Pro	Asp	Phe	Leu	Leu	Lys	Gln	Pro	Phe	Gly	Gln	Val	Pro	55
<i>Manihot esculenta</i>	106	GAT	TCC	GGC	GAG	CAA	AAG	CAA	<b>CCT</b>	<b>CAC</b>	<b>TTC</b>	<b>CTT</b>	<b>CTC</b>	AAA	CAG	CCA	TTT	GGC	CAA	GTT	CCA	165
	36	Asp	Ser	Gly	Glu	Gln	Lys	Gln	Pro	His	Phe	Leu	Leu	Lys	Gln	Pro	Phe	Gly	Gln	Val	Pro	55
<i>VvGST4</i>	106	GAC	TCT	GGC	GAG	CAA	AAA	CGG	<b>CCT</b>	<b>GAT</b>	<b>TTC</b>	<b>CTC</b>	<b>CTT</b>	CGA	CAG	CCT	TTT	GGG	CAA	GTT	CCA	165
	36	Asp	Ser	Gly	Glu	Gln	Lys	Arg	Pro	Asp	Phe	Leu	Leu	Arg	Gln	Pro	Phe	Gly	Gln	Val	Pro	55
<i>PpRiant1</i>	106	GAG	GTG	GGA	GAG	CAA	AAG	CAA	<b>CCT</b>	<b>CAG</b>	<b>TTC</b>	<b>CTC</b>	<b>TCC</b>	CGT	CAG	CCG	TTT	GGT	CAA	GTT	CCA	165
	36	Glu	Val	Gly	Glu	Gln	Lys	Gln	Pro	Gln	Phe	Leu	Ser	Arg	Gln	Pro	Phe	Gly	Gln	Val	Pro	55
<i>PpRiant2</i>	106	GAG	GTG	GGA	GAG	CAA	AAG	CAA	<b>CCT</b>	<b>CAG</b>	<b>TTC</b>	<b>CTC</b>	<b>TCC</b>	CGT	CAG	CCG	TTT	GGT	CAA	GTT	CCA	165
	36	Glu	Val	Gly	Glu	Gln	Lys	Gln	Pro	Gln	Phe	Leu	Ser	Arg	Gln	Pro	Phe	Gly	Gln	Val	Pro	55
<i>CkmGST3</i>	106	GAA	TCC	GGG	GAG	CAT	AAG	AAG	<b>CCC</b>	<b>GAT</b>	<b>TTT</b>	<b>CTT</b>	<b>CTT</b>	CGC	CAG	CCC	TTT	GGA	CAA	GTC	CCA	165
	36	Glu	Ser	Gly	Glu	His	Lys	Lys	Pro	Asp	Phe	Leu	Leu	Arg	Gln	Pro	Phe	Gly	Gln	Val	Pro	55
<i>PhAN9</i>	106	GAT	TCT	CTC	GAG	CAG	AAA	AAA	<b>CCT</b>	<b>GAG</b>	<b>TTT</b>	<b>CTA</b>	<b>GTT</b>	TTA	CAG	CCA	TTT	GGA	CAA	GTT	CCT	165
	36	Asp	Ser	Leu	Glu	Gln	Lys	Lys	Pro	Glu	Phe	Leu	Val	Leu	Gln	Pro	Phe	Gly	Gln	Val	Pro	55
<i>AtTTP19</i>	106	GAT	ACA	TTT	GAG	CAA	AAA	AAA	<b>CCA</b>	<b>GAA</b>	<b>CAT</b>	<b>CTT</b>	<b>CTT</b>	CGT	CAG	CCA	TTT	GGT	CAA	GTT	CCA	165
	36	Asp	Thr	Phe	Glu	Gln	Lys	Lys	Pro	Glu	His	Leu	Leu	Arg	Gln	Pro	Phe	Gly	Gln	Val	Pro	55
<i>LcGST4</i>	106	GAC	AAT	GGA	GAG	CAT	AAG	CGT	<b>CCT</b>	<b>GAG</b>	<b>TTT</b>	<b>CTT</b>	<b>CTT</b>	CGA	CAG	CCT	TTT	GGG	CAA	GTT	CCA	165
	36	Asp	Asn	Gly	Glu	His	Lys	Arg	Pro	Glu	Phe	Leu	Leu	Arg	Gln	Pro	Phe	Gly	Gln	Val	Pro	55

## Figure 8

Partial sequence alignment of Bract1, anthocyanin-related GSTs and orthologue GSTs from Euphorbiaceae species. The alignment spans a 60 bp region of the CDS containing the (CTTC)<sub>3</sub> SSR motif (represented in red). The numbers in the alignments indicate the nucleotide and amino acid positions in the CDS. Sequences were aligned using the ClustalW function in the BioEdit Sequence Alignment Editor v7.2.5. The complete alignment is available in Additional File S7.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AdditionalFileS7.docx](#)
- [AdditionalFileS6.docx](#)
- [AdditionalFileS5.docx](#)
- [AdditionalFileS4.xlsx](#)
- [AdditionalFileS3.docx](#)
- [AdditionalFileS2.docx](#)

- [AdditionalFileS1.docx](#)