Comparative expression analysis in three Brassicaceae species revealed compensatory changes of the underlying gene regulatory network

Jessica Pietsch[1†], Anna Deneer[2†], Christian Fleck[3*]
and Martin Hülskamp[1*]

[1]Botanical Institute, Biocenter, Cologne University, Cologne, Germany, [2]Biometris, Department of
Mathematical and Statistical Methods, Wageningen University, Wageningen, Netherlands, [3]Spatial
Systems Biology Group, Center for Data Analysis and Modeling, University of Freiburg,
Freiburg, Germany

Trichomes are regularly distributed on the leaves of Arabidopsis thaliana. The
gene regulatory network underlying trichome patterning involves more than 15
genes. However, it is possible to explain patterning with only five components.
This raises the questions about the function of the additional components and
the identification of the core network. In this study, we compare the relative
expression of all patterning genes in A. thaliana, A. alpina and C. hirsuta by qPCR
analysis and use mathematical modelling to determine the relative importance
of patterning genes. As the involved proteins exhibit evolutionary conserved
differential complex formation, we reasoned that the genes belonging to the
core network should exhibit similar expression ratios in different species.
However, we find several striking differences of the relative expression levels.
Our analysis of how the network can cope with such differences revealed
relevant parameters that we use to predict the relevant molecular adaptations
in the three species.

KEYWORDS

evolution, trichomes, genetic analysis, patterning, brassicaceae, *arabis alpina*,
cardamine hirsuta

1 Introduction

Evolutionary differences and adaptive strategies within plants are driven by the
structure and function of the underlying gene regulatory networks (GRNs) (Alvarez-
Buylla et al. (2007); Mejia-Guerra et al. (2012); Jones and Vandepoele (2020)). Even
minute changes in a GRN can result in striking differences between species (Guo and
Amir (2021); Schember and Halfon (2022)). In evolutionary developmental approaches,

such differences are studied in order to gain insight into the genetic basis of phenotypic diversity (Purugganan (1998); Simpson (2002); Fernandez-Mazuecos and Glover (2017)). A system that is well-suited for such an approach is trichome patterning in *Arabidopsis thaliana* and other Brassicaceae species (Hülskamp (2004); Doroshkov et al. (2019); Chopra et al. (2019)). Genetic analysis of trichome patterning in *A. thaliana* has revealed a complex GRN that controls the regular distribution of trichomes in the leaf epidermis (Marks et al. (1991); Hülskamp et al. (1994); Pattanaik et al. (2014)). Most of the genes found in *A. thaliana* are present in *Arabis alpina* and appear to have the same function in regulating trichome patterning (Chopra et al. (2014)). This suggests that the core of the GRN found in *A. thaliana* might be operating in other Brassicaceae as well. Therefore, the evolutionary analysis of trichome patterning in different Brassicaceae species may enable the identification of subtle changes of the underlying GRN.

In *A. thaliana*, trichomes are initiated in a regular pattern early in leaf development. Genetic analysis identified mutants in which regular pattern formation was disturbed and subsequent molecular analysis revealed the relevant genes (Hülskamp et al. (1994)). One group of genes promotes trichome formation and a second group inhibits trichome formation (Pesch and Hülskamp (2009)). The core of the network is a group of three genes, the R2R3MYB protein encoding gene *GLABRA1 (GL1)* (Oppenheimer et al. (1991); Zhang et al. (2003); Kirik et al. (2005)), the bHLH protein encoding gene *GLABRA3 (GL3)* (Payne et al. (2000); Zhang et al. (2003); Feller et al. (2011)) and the WD40 protein encoding gene *TRANSPARANT TESTA GLABRA1 (TTG1)* (van Nocker and Ludwig (2003); Zhao et al. (2008); Zhang and Schrader (2017)). The respective proteins form a complex in which GL1 and TTG1 both bind to GL3 (Payne et al. (2000); Pesch et al. (2015)). This so-called MBW (**M**YB, **b**HLH, **W**D40) complex promotes trichome development (Payne et al. (2000)). In addition, *MYB23* and *EGL3* were found to act redundantly with *GL1* and *GL3*, respectively (Kirik et al., 2001; Kirik et al. (2005)); Zhang et al. (2003)). A second group of genes act as inhibitors of trichome formation. These are all encoded by small R3MYB transcription factors including *TRIPTYCHON (TRY)* (Schellmann et al. (2002); Zimmermann et al. (2004); Wang et al. (2008); Pesch and Hülskamp (2011)), *CAPRICE (CPC)* (Wada et al. (1997); Schellmann et al. (2002)), *ENHANCER OF TRY and CPC1, 2 and 3 (ETC1, ETC2, ETC3)* (Kirik et al. (2004); Kirik et al., 2005); Tominaga et al. (2008); Wester et al. (2009)) and *TRICHOMELESS1* and *2 (TCL1* and *TCL2)* (Wang et al. (2007); Gan et al. (2011)). TRY and CPC seem to be the major players as the single mutants exhibit clear phenotypes which is enhanced in combinations with the others suggesting redundant action (Schellmann et al. (2002)). These inhibitors repress the function of the MBW complex by competing with GL1 for binding to GL3/EGL3. The detailed analysis of the function of

these genes led to two principles that can explain the generation of trichome spacing patterns without pre-existing information (*de novo* patterning). In short, the first principle is an activator inhibitor model (Meinhardt and Gierer (2000)): the three MBW proteins activate the expression of the inhibitors, that can move within the tissue and repress the MBW function (Digiuni et al. (2008); Pesch et al. (2015)). The second principle is an activator depletion model (Meinhardt and Gierer (2000)). Here, the activator TTG1 is mobile and captured by GL3 in trichome precursors, which in turn leads to a depletion of TTG1 in the neighbouring cells and thereby inhibition of trichome formation (Bouyer et al. (2008); Balkunde et al. (2011)). It is likely that both principles act in parallel (Balkunde et al. (2020)). Mathematical models have been developed to study the behaviour of these principles in more detail (Bouyer et al. (2008); Digiuni et al. (2008); Benitez et al. (2008); Balkunde et al. (2020)). These principles can explain how a trichome cell is selected. In the selected trichome cell, the homeobox transcription factor gene *GLABRA2 (GL2)* is turned on and considered to initiate the differentiation into a trichome cell (Szymanski et al. (1998); Morohashi et al. (2007); Wang and Chen (2008)).

The systematic forward genetic screen for trichome mutants in *A. alpina* has enabled the identification and functional characterization of trichome patterning genes in this species (Stephan et al. (2019); Chopra et al. (2019)). *A. alpina* diverged from *A. thaliana* between 26 and 40 million years ago (Koch et al. (2006); Beilstein et al. (2010); Willing et al. (2015)). At this evolutionary distance it was possible to identify the gene orthologs to those in *A. thaliana* by synteny on the chromosomes (Chopra et al. (2019)). It was therefore possible to unambiguously recognize not only the homologous genes, but also that two of the seven inhibitor genes, *TCL1* and *ETC2*, are missing. The genetic analysis revealed two interesting changes in the GRN. First, the *GL3* gene in *A. alpina* does not appear to act redundantly with *EGL3*. While in *A. alpina* the *gl3* single mutant is completely devoid of trichomes, it requires the *gl3 egl3* double mutant in *A. thaliana* to express the full phenotype (Chopra et al. (2019)). Other than that, the structure of the GRN in *A. alpina* seems to the same as in *A. thaliana*. It is, however, noteworthy, that the response of the network to overexpression of *GL3* is very different such that this causes the production of more trichomes in *A. thaliana* and less in *A. alpina*. This difference in the behaviour can be explained by different relative expression levels of two key genes, *GL1* and *TRY*, in the two species (Chopra et al. (2019)). Modelling revealed that this change in the parameters is sufficient to explain the different responses to *GL3* overexpression.

In this work, we compared the relative expression levels of trichome patterning genes in *A. thaliana*, *A. alpina* and *Cardamine hirsuta*. The additional species *C. hirsuta* is estimated to have diverged between 13 and 43 million years ago from *A. thaliana* (Koch et al. (2001); Beilstein et al. (2008)). For comparison between the species, the expression levels of the

trichome patterning genes were considered relative to *GL3/EGL3* as all patterning proteins bind to GL3 thereby regulating its activity. We observed striking differences raising the question whether and how the GRN established in *A. thaliana* has adapted to this. We analysed the differences by mathematical modelling and determined which parameters (i.e. interactions and regulations in the GRN) could explain the observed differences in the relative expression levels.

2 Methods

2.1 Primer establishment and validation

qPCR primers must meet particular requirements. Preferably intron-spanning primers were designed using GenScript Real-time PCR Primer Design (www.genscript.com) with an optimal melting temperature of 60±2°C and sequence specific amplicons of ideally 150-200 bp. They exhibit one single band of the expected size in agarose gel electrophoresis and a single peak in the melting curve. Amplification efficiency and correlation were determined based on serial cDNA dilution steps (1:10, 1:20, 1:40,1:80, 1:160, 1:320). Cq and log10 values of the dilution series were used to calculate the slope Δ by

$$\Delta = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{N}(x_i - \bar{x})^2} \qquad (1)$$

where N is the number of dilution steps. The slope served to calculate the primer efficiency E by

$$E = 100 \cdot \left(10^{\frac{-1}{\Delta}} - 1\right) \qquad (2)$$

The R^2 correlation of the Cq and the log10 values was calculated using

$$\rho_{x,y} = \frac{Cov(X,Y)}{\sigma_x \sigma_y} \qquad (3)$$

Amplification efficiencies of 100% ± 20 for genes of interest and 100% ±10 for reference genes as well as a linear standard curve with a correlation of ≥0.99 were accepted. Sequences were taken from TAIR (www.arabidopsis.org), from Genomic resources for *Arabis alpina* (www.arabis-alpina.org) and from *Cardamine hirsuta* genetic and genomic resource (http://chi.mpipz.mpg.de).

2.2 Plant material and sample preparation

For this analysis we used *Arabidopsis thaliana* Col-0, *Arabis alpina* Pajares and *Cardamine hirsuta* Ox. Cotyledons as well as juvenile leaves (leaf one and two for *Arabidopsis thaliana* and *A. alpina*, additionally leaf three for *C. hirsuta*) of plant seedlings were removed to gather 200-400 *μm* sized leaves with on-going

trichome patterning machinery. *A. thaliana* plants were 10 days old, *A. alpina* 14 days and *C. hirsuta* 7 days. Material of up to 45 plants was collected per biological replicate, frozen in liquid nitrogen and stored at −80 C until further processing. RNA extraction was performed using the Tri-Reagent method including DNaseI treatment and quality control was ensured *via* bleach gel and photometry. cDNA synthesis was carried out according to the manufacturer's protocol (RevertAid First Strand cDNA Synthesis Kit; Thermo Fisher Scientific) using 1.5 *μg* RNA per sample because pre-tests had revealed that 1 *μL* undiluted cDNA based on 1 *μg* RNA were required to obtain Cq values<30. qPCR protocols were standardized using three biological as well as three technical replicates, master mixes and always both reference genes on each plate.

2.3 Analysis of qPCR data

A two-sided Grubbs test (α=0.05) was performed to identify outliers. Normalization of the data was conducted according to the geNorm manual (Vandesompele et al. (2002)), describing gene expressions relatively to each other. Special considerations are given to normalization factors and the individual primer efficiencies ∈. Thereby not a generalized gene duplication per cycle $(1 + 1)$ is assumed, but the individual amplification rate $(1 + \in)$ is used for further calculations. The expression data of each species was normalized by two different reference genes. Using the variability of the reference genes and not the Cq values, allows interspecies comparison even with different reference genes for each organism.

2.4 Compiling GL1 synteny

Arabidopsis thaliana was used as reference to elaborate the synteny of *GL1* comparing it with *A. alpina* and *C. hirsuta*. The *AtGL1* sequence was used to perform a BLAST search against the *C. hirsuta* CDS database (http://bioinfo.mpipz.mpg.de/blast/cgi-bin/public_blast_cs.cgi). More than a dozen of loci spanning the first three highest ranked genes were blasted against *Arabidopsis thaliana*. The *AaGL1* ortholog as well as its adjoining genes were identified using the 1x1 orthologs table from the Arabis alpina website (http://www.arabis-alpina.org/data/ArabisAlpina/data/Aa_At_ortho_1x1.txt)

2.5 Mathematical modelling

The model consists of 8 components, which are modelled in the form of a system of coupled ordinary differential equations. These components include the proteins TTG1, GL1, GL3, TRY, CPC and ETC. Note that the species designated by GL1 and ETC are assumed to be the combined behavior of GL1, MYB23 and

ETC1, ETC2 and ETC3, respectively. Additionally, the complex formation between GL3 and TTG1 and GL3 and GL1 is explicitly modelled, whereas the binding between GL3 and the inhibitors TRY, CPC and ETC is implicitly modelled since these do not feed back into the system. This model consists of 31 parameters that describe processes such as degradation, binding, activation and transport. This model is based on previously published versions and is extended by the inclusion of ETC (Digiuni et al. (2008); Bouyer et al. (2008); Chopra et al. (2019); Balkunde et al. (2020)). The system of equations is

$$\partial_t [TTG1]_j = \theta_1 - [TTG1]_j(\theta_2 + \theta_3 [GL3]_j) + \theta_2 \theta_4 \hat{L} [TTG1]_j \quad (4)$$

$$\partial_t [GL1]_j = \theta_5 + \theta_6 [AC2]_j - [GL1]_j(\theta_7 + \theta_8 [GL3]_j)$$
$$+ \theta_7 \theta_{30} \hat{L} [GL1]_j \quad (5)$$

$$\partial_t [GL3]_j = \theta_9 + \frac{\theta_{10} \theta_{11} [AC1]_j^2}{\theta_{11} + [AC1]_j^2} + \frac{\theta_{12} \theta_{13} [AC2]_j^2}{\theta_{13} + [AC2]_j^2} -$$
$$[GL3]_j(\theta_{14} + \theta_3 [TTG1]_j + \theta_8 [GL1]_j + \theta_{15} [TRY]_j + \quad (6)$$

$$\theta_{16} [CPC]_j + \theta_{17} [ETC]_j) + \theta_{14} \theta_{31} \hat{L} [GL3]_j \quad (7)$$

$$\partial_t [TRY]_j = \theta_{18} [AC1]_j^2 - [TRY]_j(\theta_{19} + \theta_{15} [GL3]_j) +$$
$$\theta_{19} \theta_{20} \hat{L} [TRY]_j \quad (8)$$

$$\partial_t [CPC]_j = \theta_{21} [AC2]_j^2 - [CPC]_j(\theta_{22} + \theta_{16} [GL3]_j) +$$
$$\theta_{22} \theta_{23} \hat{L} [CPC]_j \quad (9)$$

$$\partial_t [ETC]_j = \theta_{24} [AC1]_j^2 + \theta_{25} [AC2]_j^2 - [ETC]_j(\theta_{26} - \theta_{17} [GL3]_j) +$$
$$\theta_{26} \theta_{27} \hat{L} [ETC]_j \quad (10)$$

$$\partial_t [AC1]_j = \theta_3 [GL3]_j [TTG1]_j - \theta_{28} [AC1]_j \quad (11)$$

$$\partial_t [AC2]_j = \theta_8 [GL3]_j [GL1]_j - \theta_{29} [AC2]_j \quad (12)$$

where $\hat{L}$ indicates the coupling equation between cells, given by

$$\hat{L}[\chi]_{x,y} = [\chi]_{y-1,x} + [\chi]_{y+1,x} + [\chi]_{y,x-1} + [\chi]_{y,x+1}$$
$$+ [\chi]_{y+1,x-1} + [\chi]_{y-1,x+1} - 6[\chi]_{y,x}. \quad (13)$$

for any species χ and cell at coordinates (x, y). Patterns were simulated on a grid of 20-by-20 cells with hexagonal connectivity on a domain with zero-flux boundary conditions. The initial conditions are given by the steady state of a single-cell model (i.e. $\hat{L} = 0$) plus small inhomogeneous perturbations, sampled from the standard uniform distribution. The trichomes on the grid are identified by cells that have

relatively high amounts of active complex (AC1 + AC2), specifically, cells that have more than the half-maximum of total AC are designated as trichomes. The cluster density of trichomes was averaged over 10 simulations, each with randomized initial conditions. Parameter sets that produced less than 10% clusters are used for further analysis.

Given that we are only interested in parameter sets that form patterns, we apply linear stability anaylsis to identify these sets (Murray (2001)). In the domain of interest, a diffusion-driven instability (Turing instability) occurs (Turing (1952)), resulting in an inhomogeneous patterning state. In linear stability analysis, the stability of a uniform steady state is verified by determining whether effects of small perturbations to the ODE system decay over time. Turing instability was tested by the following criteria: starting from a uniform steady state (i) the steady state in the absence of diffusion is stable and (ii) the steady state in the presence of diffusion is unstable (Murray (2001)). For criterion i this means this means that all eigenvalues of the Jacobian of the system in (4) - (12) evaluated at steady state must be negative. To perform the same test for criterion ii we decoupled the system by Fourier transformation and analysed the eigenvalues (Bouyer et al. (2008); Digiuni et al. (2008); Balkunde et al. (2020)), where the real part of at least one of the eigenvalues must be positive.

2.6 Parameter estimation

The previous section describes the rationale and rules for the parameter sets used for the model. To use the model for our qPCR data set we estimated the parameter sets through an optimization routine where the qPCR data are used in a cost function. The goal is to arrive at a distribution of values for these parameters for each of the species. Note that this problem suffers from non-identifiability (Kreutz et al. (2013)), i.e., no unique value or bounded confidence interval can be determined for the parameters; for this, additional data would be required that is simply not available. Nonetheless, through a multi-start optimization routine that ensures multiple optimal solutions, it is possible to deal with the uncertainty in the system and arrive at predictions about possible genetic adaptations on a regulatory level that differentiates the three species from each other.

The analysis used here requires solving a constrained multivariable minimization problem (Boese et al., 1994). Specifically, we aim to find the minimum of the problem specified by

$$\min_{\theta} f(\theta) \quad \text{such that} \quad \begin{cases} c(\theta) \geq 0 \\ \text{lb} \leq \theta \leq \text{ub} \end{cases} \quad (14)$$

where $c(\theta)$ is a non-linear constraint function, $f(\theta)$ is a scalar cost-function and lb and ub are the lower- and upper-bounds of the parameter vector θ. The cost function f is a normalized sum-of-squares given by

$$f(\theta) = \sum_{i=1}^{N} \frac{(\bar{y}_i(\theta) - y_i)^2}{y_i^2} \qquad (15)$$

where $\bar{y}_i(\theta)$ is the expression level of the i-th gene out of N total genes predicted by the model and y_i is the corresponding datapoint. Given that the model simulates the concentration in a tissue, $\bar{y}_i(\theta)$ is the average of gene i across the tissue, relative to the sum of GL3 and EGL3, similar as the data.

The constraint function $c(\theta)$ is chosen such that the parameters θ fall in the Turing Space, i.e. are capable of patterning. To achieve this, we make use of linear stability analysis as described above and determine the eigenvalues of the Jacobian of the system of equations. $c(\theta)$ is given by

$$c(\theta) = \mathrm{Re}(\lambda_{\max}) \qquad (16)$$

where $\lambda_{\max}$ is the largest eigenvalue of the Jacobian. By determining whether the real part of the largest eigenvalue is positive (i.e. $c(\theta) \geq 0$), we learn that the parameter set θ can form a pattern, which constrains the allowable range of parameters. Note that this range is also constrained by the choice of bounds (lb and ub) of the optimization problem. In this case, we set the interval for the each of the parameters in θ to [0.01,100], to allow a range of multiple orders of magnitude. One further constraint, which is applied in post-processing, is that the pattern produced by θ must not show any clusters of trichomes, as is the case in all the patterns formed by the three species. As such, we limit the range of parameter values to those that simulate a realistic pattern and produce the best possible fit to the data according to $f(\theta)$. Finally, the optimization problem is started from multiple, randomly generated initial points. This set of initial points is generated by a Sobol sequence to ensure a good coverage of the parameter space (Sobol (2001)). Starting from these randomly generated initial points, the optimization converges to a local minimum that satisfies the constraints, leading to a distribution of optimal parameter sets $\hat{\theta}$ that correspond to the local minima. This procedure is followed until 100 optimized parameter sets are obtained for each of the species.

These distributions are then used in a statistical analysis to determine which of the parameter distributions are statistically different between the species. Towards this end, we use the two-sample Kolmogorov-Smirnov (KS) goodness-of-fit hypothesis test to determine if two empirical distributions are drawn from the same (unknown) underlying population cumulative distribution functions (Stephens (1974)).

2.7 Sensitivity analysis

The sensitivity of the trichome density to the each of the individual parameters is determined, using a variation on the elementary effects (EE) test (Campolongo et al., 2007; Saltelli, 2008). The EE is a one-at-a-time screening method, i.e.,only one

parameter is varied at a time and the variation in the output is measured (Campolongo et al. (2007); Saltelli (2008)). For a model with N parameters, each parameter $\theta_i, i = 1,\ldots N$, is assumed to vary across p selected levels in the parameter space. The region of experimentation Ω is an N-dimensional p-level grid. In standard sensitivity practices, parameters are assumed to be uniformly distributed in [0,1] and then transformed from the unit hypercubeto their known distributions (Saltelli (2008)). In this case, we adapt this region to ensure that the parameters fall within the Turing Space. The lower limit of Ω is by default chosen to be 10^{-1} and the upper limit 10,where every point Δ in the grid is the perturbation applied to θ_i for which the EE_i is determined. In the case that either limit would shift θ outside of the Turing space, then the lower limit is adjusted to the smallest value between $[10^{-1},1]$ that according to linear stability analysis falls within the Turing space, and the upper limit is the largest value between [1,10] that falls within the Turing space. This means that the p-level grid in Ω can have different upper and lower limits, depending on the allowable range according to linear stability analysis, but always consists of the same number of grid-points. Furthermore, these grid-points are chosen such that they are logarithmically spaced.

For the trichome patterning model, we have $N = 31$ and choose $p = 10$. We perform the EE sensitivity analysis for the top 10 best-fitting parameter sets $\hat{\theta}$ resulting from the optimization routine. For a given set $\hat{\theta}^k$, $k = 1,\ldots,10$ the EE of the i-th parameter is defined as:

$$\mathrm{EE}_i\left(\hat{\theta}^k\right) = \frac{Y(\theta_1,\ldots,\theta_{i-1},\theta_i,\theta_{i+1},\ldots,\theta_N) - Y\left(\hat{\theta}^k\right)}{\theta_i \cdot \Delta - \theta_i} \qquad (17)$$

where Δ is a value in the p-level grid with the limits chosen as described above. Then, the absolute values of the EE_i, computed at p different grid points, are averaged to get

$$\mathrm{E\bar{E}}_i = \frac{1}{p}\sum_{j=1}^{p} | \mathrm{EE}_i^j |. \qquad (18)$$

Finally, we average over all $\mathrm{E\bar{E}}_i$ for every $\hat{\theta}^k$.

3 Results

3.1 Comparison of trichome gene expression in different species

All three species considered here, *A. thaliana*, *A. alpina* and *C. hirsuta* produce regularly spaced trichomes on leaves (Greese et al. (2012); Chopra et al. (2019)). The trichome density differs such that *C. hirsuta* has a lower density, and *A. alpina* a higher density as compared to *A. thaliana*. A meaningful quantitative comparison of trichome density appears not to be possible as leaf sizes, growth dynamics and the juvenile-to-adult transition differ

making it arbitrary to choose the proper mature leaves for comparison. We therefore focused on qualitative and ratiometric comparisons in this study.

A direct comparison of the expression levels of trichome genes between species by qPCR is not possible for various reasons. In particular, the primers are different for the same genes and normalization was done with difference reference genes. We therefore compared the expression levels between species by normalizing the expression to the bHLH genes. The bHLH protein is the central component of the MBW complex to which the activators *TTG1* and R2R3 MYB proteins bind (Payne et al. (2000)) and on which the R3 MYB negative regulators exert their repressive effect by competitive binding with the R2R3 MYBs (Digiuni et al. (2008)). It is conceivable that the outcome of this GRN depends on the concentrations of the other patterning proteins relative to the bHLH. We therefore considered the bHLH expression levels a good reference to judge and compare to the relative changes of all other patterning genes. We combined *GL3* and *EGL3* for the comparison between the species because the two genes act redundantly in *A. thaliana* and have similar molecular roles in trichome patterning (Zhang et al. (2003); Bernhardt et al. (2003); Morohashi et al. (2007)).

Another aspect to enable a comparison between the species is the choice of plant material. For our qPCR experiments we used young leaves at developmental stages in which trichome patterning was still ongoing as recognized by the presence of incipient and young stages of trichome development at the base of leaves. These stages could be unambiguously identified in all three species.

3.2 Relative trichome gene expression differs in *A. thaliana*, *A. alpina* and *C. hirsuta*

In a first step, we identified the bonafide orthologs of the *Arabidopsis* trichome patterning genes in *A. alpina* and *C. hirsuta* by sequence similarity and synteny (Supplementary Figure S1). Primers were designed to meet the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines (Bustin et al. (2005)).

Plants were grown on soil and young leaves were harvested at stages at which incipient developing trichomes were seen. The first two leaves and the cotyledons were removed. Quantitative Real-Time PCR experiments were done with three biological replicas and normalized to a set of species-specific reference genes. To enable a comparison between species we normalized all expressions with *GL3/EGL3*. Figure 1 shows the relative expression levels of patterning genes normalized to the combined transcript levels of *GL3* and *EGL3* which was set to one (see also Table S1). In *A. thaliana*, *GL3/EGL3* appear to be the limiting factors among the other trichome activating genes
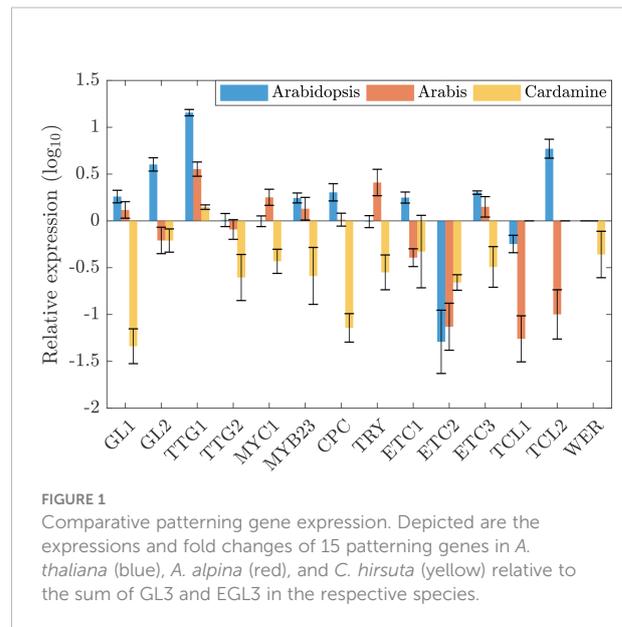

FIGURE 1
Comparative patterning gene expression. Depicted are the expressions and fold changes of 15 patterning genes in *A. thaliana* (blue), *A. alpina* (red), and *C. hirsuta* (yellow) relative to the sum of GL3 and EGL3 in the respective species.

(Figure 1). *AtTTG1* expression is about 14-fold higher and the expression of *AtGL1* and *AtMYB23* both show an about 2-fold higher expression. Consistent with the previous finding that *AtGL1* and *AtMYB23* act redundantly in *Arabidopsis* (Kirik et al., 2001), they show a similar expression level and we combined their expression for the modelling approach to reduce the complexity (see below). The relative expression levels of the inhibitors were different with *AtCPC*, *AtETC1*, *AtETC3* and *AtTCL2* being higher and *AtETC2* and *AtTCL1* lower than *AtGL3/AtEGL3*.

The expression profile in *A. alpina* is fairly similar to that in *A. thaliana*. In *C. hirsuta* we found a strikingly different pattern of the relative expression of trichome patterning genes. Here, most of the patterning genes exhibit lower expressions as compared to *ChGL3/ChEGL3*. In particular, the expression of *ChGL1* and *ChCPC* were drastically lower as compared to the other two species.

3.3 GL1, MYB23 and WER expression differs in *Arabidopsis thaliana*, and *C. hirsuta*

The very low relative and also absolute expression levels of *ChGL1* in *C. hirsuta* raised the question whether the function of *ChGL1* is redundantly provided by *ChMYB23* (Kirik et al. (2001)) or even *ChWER* (Lee and Schiefelbein (1999)). To study this in more detail, we compared the expression of the three genes in five different tissues between *A. thaliana* and *C. hirusta* (Figure 2). To facilitate a comparison in the context of trichome patterning, we normalized the expression levels with respect to the *GL1* expression in young leaves. As expected, *AtGL1* and *AtMYB23* are expressed in most aerial tissues in
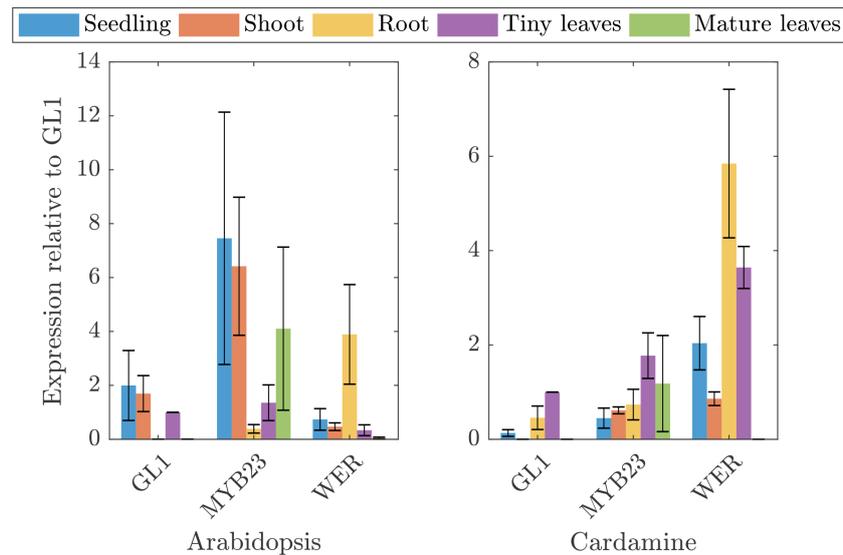
FIGURE 2
Quantitative expression analysis of three MYB homologs in *A. thaliana* and *C. hirsuta* in different tissues. Depicted is the expression of *GL1,* *MYB23*, and *WER* in *Arabidopsis thaliana* and *Cardamine hirsuta* in seedlings (blue), shoots (red), roots (yellow), tiny leaves (purple), and mature leaves (green), relative to GL1 expression in tiny leaves of the respective species.

Arabidopsis thaliana but not in the root, whereas *AtWER* expression was detected strongly in the root. In *C. hirsuta*, *ChGL1* and *ChMYB23* expression was absent or low in all tissues. Surprisingly, *ChWER* expression was not only high in roots, but also in young leaves. Here, *ChWER* expression was 2.6 fold higher than that of *ChGL1*. These findings suggest that the tissue specific functions of *GL1, MYB23* and *WER* might be different in the species. Given that *AtGL1* and *AtWER* proteins have equivalent function during trichome initiation in *A. thaliana* (Lee and Schiefelbein (2001); Kirik et al. (2005)), it is conceivable that the higher expression of *ChWER* can substitute the low levels of *ChGL1*.
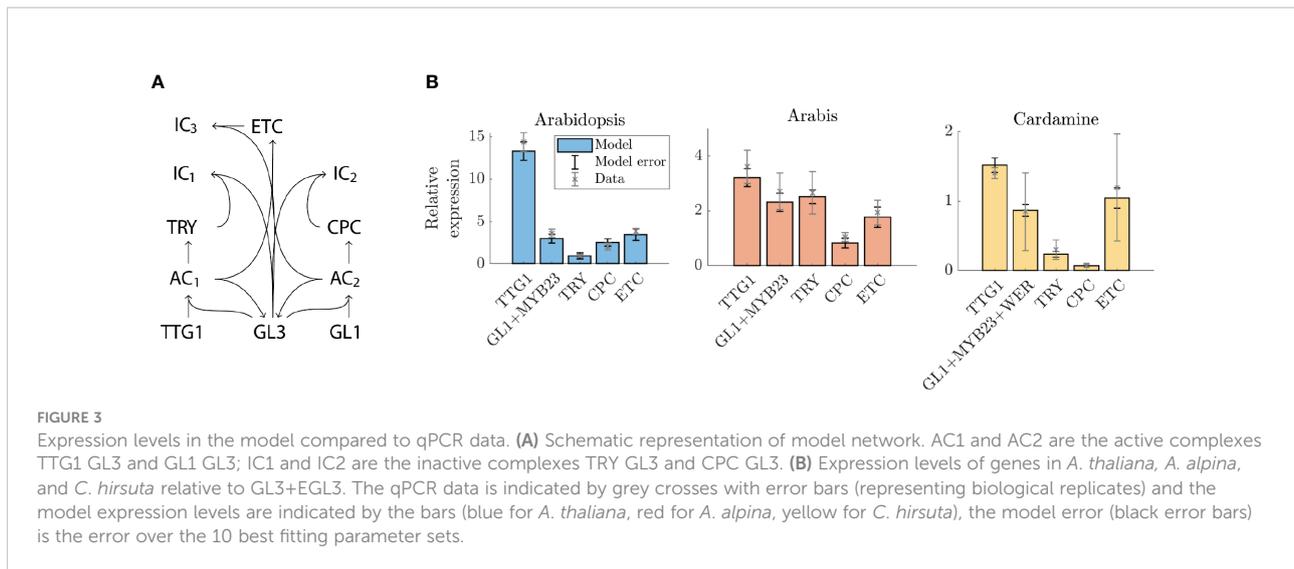
3.4 Modelling to predict the molecular adaptations to relative expression differences between the three species

The functional comparison of the regulation between *A. thaliana* and *A. alpina* suggests that the core of the underlying regulatory network of trichome patterning is conserved (Chopra et al. (2019)). Consistent with this, all relevant orthologs of the relevant *A. thaliana* genes are also found in *C. hirsuta*. The qPCR data show striking differences in the relative expression levels. Given that trichome initiation is driven by the activity of the MBW complex, in which the components undergo competitive binding, it is surprising that the GRN can tolerate such differences (Digiuni et al. (2008); Pesch et al. (2015)). We used mathematical modelling to explore whether the

Arabidopsis-based GRN is capable of coping with such differing relative expression patterns. And if so, which parameters can compensate for changes in the relative expression levels and to which of these is the pattern most sensitive? Towards this end, we developed a model based on previous versions (Bouyer et al. (2008); Digiuni et al. (2008); Balkunde et al. (2020)) and consisting of *TTG1, TRY, CPC, ETC1/ETC2/ETC3, GL3/EGL3* and *GL1/MYB23/WER* (Figure 3A). Note that the GRN that is modelled is the same for each species (Figure 3A) and that the difference between the species comes from differences in the underlying parameters.

We define two criteria for the model to fulfil. First, the parameter set has to reproduce the same relative differences as found for the patterning genes. Second, it has to simulate realistic trichome patterns, i.e., patterns must not show any clustering of trichomes (Greese et al., 2012). We varied all 31 parameters using a non-linear optimization routine such that the model most accurately reproduces the expression data (Figure 3). This is surprisingly well possible for the expression data sets of all three species with many different parameter sets.

The identification of a large number of parameter combinations for each species enabled us to compare the distributions of the parameters between the three species in search for striking differences. Towards this end, we used a Kolmogorov-Smirnov test (Stephens (1974)) to identify parameter distributions that are significantly different between the species. Out of the 31 parameters only 13 fulfilled this criterion and were considered parameters that are relevant for compensating different expression ratios in all three species. The

FIGURE 3
Expression levels in the model compared to qPCR data. **(A)** Schematic representation of model network. AC1 and AC2 are the active complexes TTG1 GL3 and GL1 GL3; IC1 and IC2 are the inactive complexes TRY GL3 and CPC GL3. **(B)** Expression levels of genes in *A. thaliana*, *A. alpina*, and *C. hirsuta* relative to GL3+EGL3. The qPCR data is indicated by grey crosses with error bars (representing biological replicates) and the model expression levels are indicated by the bars (blue for *A. thaliana*, red for *A. alpina*, yellow for *C. hirsuta*), the model error (black error bars) is the error over the 10 best fitting parameter sets.

distributions are shown in Figure 4. The 13 parameters have significantly different distributions for at least two of the species, indicating that the *Arabidopsis*-based model can cope with different relative expression levels by compensatory changes of different parameters. For the other 18 parameters we did not find a significant difference (Supplementary Figure S2).
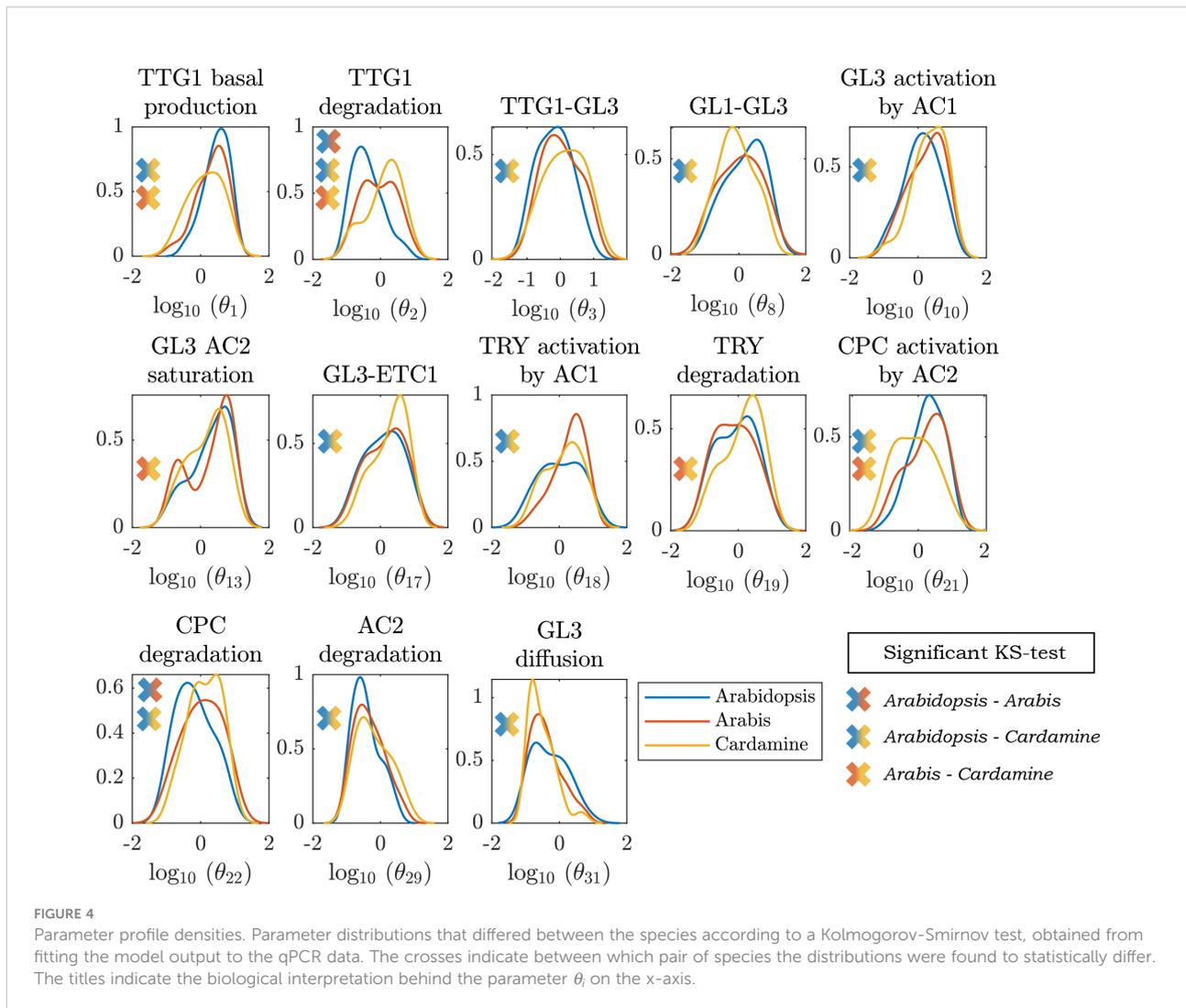
This analysis revealed three interesting differences. First, all parameters regulating the activity of TTG1 in *A. thaliana* differ from the distributions of *A. alpina* and *C. hirsuta* (except for the diffusion rate of TTG1). It is conceivable that this is due to the relatively high expression of *AtTTG1* in *A. thaliana* that could be compensated by parameters changes reducing its activity such as the basal production rate and degradation rate. Second, the parameters regulating *TRY* activity differ between *A. alpina* and the other two species. The model predicts a higher activation rate of *AaTRY* by the complex between GL3 and TTG1 in *A. alpina* that would explain the relatively high levels of *AaTRY* in this species. Third, the regulation of *ChCPC* in *C. hirsuta* differs from the other two species. We found significant differences for the activation of *ChCPC* by the complex between GL3 and GL1 and its degradation rate. Both would compensate for the relative low amount of *ChCPC* expression in *C. hirsuta*. These three cases exemplify the versatility of the trichome patterning network and show how, despite the varying underlying differences between regulatory mechanisms, the same core network is capable of robustly producing a realistic trichome pattern in all three species.

The comparison of parameter distributions provides insight into the adaptability of the network to the different expression levels but does not immediately provide information on the effects on trichome patterning. This is possible by determining the sensitivity of trichome density to changes in each of the parameters in all three species (Figure 5). This allows predicitions on which parameter is most influential in patterning and whether this varies between species.

Our sensitivity analysis predicts that for all three species the stability of the TTG1-GL3 complexes (θ_{28}) is one of the most sensitive parameters. In *A. thaliana* and *A. alpina* the stability of the GL1-GL3 complex (θ_{29}) and the degradation rate of GL1 (θ_{7}) are among the most sensitive parameters. In *C. hirsuta*, the basal production of GL3/EGL3 is relevant (θ_{9}) and the degradation of TTG1 (θ_{2}). Taken together, the sensitivity analysis predicts that the amount of the active complexes most strongly influences the trichome density and that this is a common feature in all three species. *C. hirsuta* differs in that the role of GL3-TTG1 is more relevant than that of GL3-GL1. For an overview of the biological interpretation of all other parameters in Figure 5 see Supplementary Table 2.

4 Discussion

In this study, we have compared the expression levels of trichome patterning genes in the three closely related Brassicaceae species *A. thaliana*, *A. alpina* and *C. hirsuta*. Given that we selected one specific ecotype of each species, our analysis is only a snapshot of possible variation in each of the three species. However, within this limitation, the analysis and comparison of the GRN properties is possible. We aimed to use the variation of the relative expression levels to understand the potential of the GRN underlying trichome patterning. For our mathematical modelling approach, we used a complex model that considers the genetic interactions and simulates concentrations on the protein level to consider differential complex formations (Pesch et al., 2015). The details of transcription and translation are not explicitly modelled as this would add another layer of complexity and thereby more parameters without gaining an extra value in the absence of additional data. This approach enabled us to evaluate the

FIGURE 4
Parameter profile densities. Parameter distributions that differed between the species according to a Kolmogorov-Smirnov test, obtained from fitting the model output to the qPCR data. The crosses indicate between which pair of species the distributions were found to statistically differ. The titles indicate the biological interpretation behind the parameter θ_i on the x-axis.

parameter changes with respect to many different aspects of the patterning process including transcriptional regulation, differential complex formation and stability of the transcript/protein. The possible downside is that we have to consider a 31-dimensional parameter space making it necessary to use statistical approaches to monitor the effect of different parameters and their combinations.

What did we learn? First of all, the structure of the GRN network established in *A. thaliana* is sufficient to generate a trichome pattern even if the relative expression levels show an order of magnitude difference. Second, the GRN compensates for differences in the relative expression patterns by changing other parameters. Not all, but only a subset of the parameters is important for this. Third, one type of parameter – the stability of the MBW complexes – is among the most important in all three species. These predictions might be instrumental for future experiments as they help to focus on aspects of the GRN network that have not been studied so far.

A second unexpected result followed from our analysis of the relative transcript levels of the three R2R3MYBs: *GL1, MYB23* and *WER*. The three *Arabidopsis thaliana* genes cluster together in a subgroup of R2R3-MYB and are characterized by a unique amino acid motif near their C-termini, which is not within the MYB domain (Stracke et al. (2001)) and are likely to be the result of gene duplications. *GL1* and *MYB23* act redundantly in the regulation of trichome patterning but have distinct functions in the regulation of trichome branching (Kirik et al. (2005)). Both are only important for trichome formation but not involved in root hair patterning, which is specifically regulated by *WER*. This trait specificity is due to differences in the transcriptional regulation as WER and GL1 are equivalent proteins (Lee and Schiefelbein (2001)). Also, overexpression of *MYB23* can rescue the *wer* mutant phenotype indicating that the protein can substitute WER in this context (Tominaga-Wada et al. (2012)). In *Arabidopsis* we found *AtWER* expression in leaves was detectable, but very low
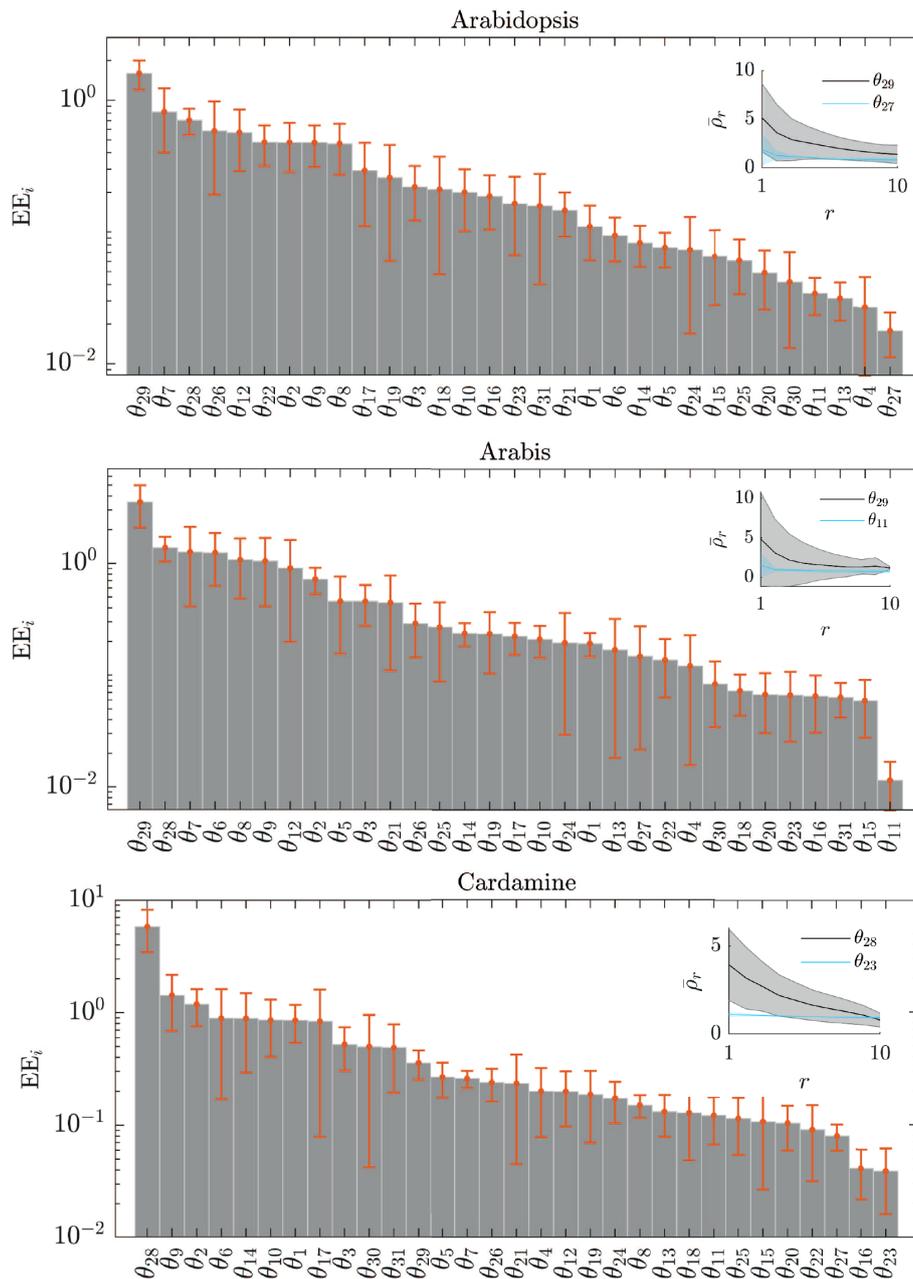
FIGURE 5

Sensitivity of parameters to trichome density. The elementary effects (EEs) of each of the parameters θ_i in the model indicates the sensitivity of the trichome density to changes in θ_i, sorted by the respective EE value. The error bars indicate the standard deviation in the EE for the ten best-fitting parameter sets. The inset shows the average and the spread (shaded region) of the trichome density ($\bar\rho_r$) for the ten best-fitting parameter sets at each of the ten grid points r of the EE test, for the most sensitive parameter (grey) and the least sensitive parameter (blue).

compared to *AtMYB23* and *AtGL1*. When considering that the expression of the three genes is important for their function and that the proteins are functionally similar, it is conceivable that *WER* might have a function in trichome development in *Cardamine*. In fact, by this argument, *WER* would be most important in root and leaf epidermal patterning as it is also

most prominently expressed in roots. We therefore hypothesize that our findings reflect the evolutionary sub-functionalization of the three homologous *MYB* genes in trichome and root hair regulation. Functional assays, ideally including mutant analysis in *Cardamine* will be required to test this hypothesis.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Author contributions

JP planned and did the experimental work and analysis. AD developed the mathematical models and performed the analysis. CF and MH contributed to the design and conception of the study. JP, AD, CF and MH wrote the manuscript. All authors contributed to the article and approved the submitted version.

Funding

Acknowledgments

Conflict of 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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1086004/full#supplementary-material

References

Alvarez-Buylla, E. R., Benitéz, M., Dávila, E. B., Chaos, A., Espinosa-Soto, C., and Padilla-Longoria, P. (2007). Gene regulatory network models for plant development. *Curr. Opin. Plant Biol.* 10, 83–91. doi: 10.1016/j.pbi.2006.11.008

Balkunde, R., Bouyer, D., and Hülskamp, M. (2011). Nuclear trapping by GL3 controls intercellular transport and redistribution of TTG1 protein in arabidopsis. *Development* 138, 5039–5048. doi: 10.1242/dev.072454

Balkunde, R., Deneer, A., Bechtel, H., Zhang, B., Herberth, S., Pesch, M., et al. (2020). Identification of the trichome patterning core network using data from weak ttg1 alleles to constrain the model space. *Cell Rep.* 33. doi: 10.1016/j.celrep.2020.108497

Beilstein, M. A., Al-Shehbaz, I. A., Mathews, S., and Kellogg, E. A. (2008). Brassicaceae phylogeny inferred from phytochrome a and ndhF sequence data: tribes and trichomes revisited. *Am. J. Bot.* 95, 1307–1327. doi: 10.3732/ajb.0800065

Beilstein, M. A., Nagalingum, N. S., Clements, M. D., Manchester, S. R., and Mathews, S. (2010). Dated molecular phylogenies indicate a Miocene origin for arabidopsis thaliana. *Proc. Natl. Acad. Sci. U. S. A.* 107, 18724–18728. doi: 10.1073/pnas.0909766107

Benitez, M., Espinosa-Soto, C., Padilla-Longoria, P., and Alvarez-Buylla, E. R. (2008). Interlinked nonlinear subnetworks underlie the formation of robust cellular patterns in arabidopsis epidermis: a dynamic spatial model. *BMC Syst. Biol.* 2, 1–16. doi: 10.1186/1752-0509-2-98

Bernhardt, C., Lee, M. M., Gonzalez, A., Zhang, F., Lloyd, A., and Schiefelbein, J. (2003). The bHLH genes GLABRA3 (GL3) and ENHANCER OF GLABRA3 (EGL3) specify epidermal cell fate in the arabidopsis root. *Development* 130, 6431–6439. doi: 10.1242/dev.00880

Boese, K. D., Kahng, A. B., and Muddu, S. (1994). A new adaptive multi-start technique for combinatorial global optimizations. *Operations Res. Lett.* 16, 101–113. doi: 10.1016/0167-6377(94)90065-5

Bouyer, D., Geier, F., Kragler, F., Schnittger, A., Pesch, M., Wester, K., et al. (2008). Two-dimensional patterning by a trapping / depletion mechanism : The role of TTG1 and GL3 in arabidopsis trichome formation. *PloS Biol.* 6. doi: 10.1371/journal.pbio.0060141

Bustin, S. A., Benes, V., Nolan, T., and Pfaffl, M. W. (2005). Quantitative real-time RT–PCR–a perspective. *J. Mol. Endocrinol.* 34, 597–601. doi: 10.1677/jme.1.01755

Campolongo, F., Cariboni, J., and Saltelli, A. (2007). An effective screening design for sensitivity analysis of large models. *Environ. Model. software* 22, 1509–1518. doi: 10.1016/j.envsoft.2006.10.004

Chopra, D., Mapar, M., Stephan, L., Albani, M. C., Deneer, A., Coupland, G., et al. (2019). Genetic and molecular analysis of trichome development in arabis alpina. *Proc. Natl. Acad. Sci. United States America* 116, 12078–12083. doi: 10.1073/pnas.1819440116

Chopra, D., Wolff, H., Span, J., Schellmann, S., Coupland, G., Albani, M. C., et al. (2014). Analysis of TTG1 function in arabis alpina. *BMC Plant Biol.* 14, 16. doi: 10.1186/1471-2229-14-16

Digiuni, S., Schellmann, S., Geier, F., Greese, B., Pesch, M., Wester, K., et al. (2008). A competitive complex formation mechanism underlies trichome patterning on arabidopsis leaves. *Mol. Syst. Biol.* 1, 1–11. doi: 10.1038/msb.2008.54

Doroshkov, A. V., Konstantinov, D. K., Afonnikov, D. A., and Gunbin, K. V. (2019). The evolution of gene regulatory networks controlling arabidopsis thaliana l. trichome development. *BMC Plant Biol.* 19, 71–85. doi: 10.1186/s12870-019-1640-2

Feller, A., Machemer, K., Braun, E. L., and Grotewold, E. (2011). Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J.* 66, 94–116. doi: 10.1111/j.1365-313X.2010.04459.x

Fernandez-Mazuecos, M., and Glover, B. J. (2017). The evo-devo of plant speciation. *Nat. Ecol. Evol.* 1. doi: 10.1038/s41559-017-0110

Gan, L., Xia, K., Chen, J.-G., and Wang, S. (2011). Functional characterization of trichomeless2, a new single-repeat r3 myb transcription factor in the regulation of trichome patterning in arabidopsis. *BMC Plant Biol.* 11, 1–12. doi: 10.1186/1471-2229-11-176

Greese, B., Wester, K., Bensch, R., Ronneberger, O., Timmer, J., Lskamp, M. H., et al. (2012). Influence of cell-to-cell variability on spatial pattern formation. *IET Syst. Biol.* 6, 143–153. doi: 10.1049/iet-syb.2011.0050

Guo, Y., and Amir, A. (2021). Exploring the effect of network topology, mRNA and protein dynamics on gene regulatory network stability. *Nat. Commun.* 12, 1–10.

Hülskamp, M. (2004). Plant trichomes: a model for cell differentiation. *Nat. Rev. Mol. Cell Biol.* 5, 471–480. doi: 10.1038/nrm1404

Hülskamp, M., Miséra, S., and Jürgens, G. (1994). Genetic dissection of trichome cell development in arabidopsis. *Cell* 76, 555–566. doi: 10.1016/0092-8674(94)90118-X

Jones, D. M., and Vandepoele, K. (2020). Identification and evolution of gene regulatory networks: insights from comparative studies in plants. *Curr. Opin. Plant Biol.* 54, 42–48. doi: 10.1016/j.pbi.2019.12.008

Kirik, V., Lee, M. M., Wester, K., Herrmann, U., Zheng, Z., Oppenheimer, D., et al. (2005). Functional diversification of MYB23 and GL1 genes in trichome morphogenesis and initiation. doi: 10.1242/dev.01708

Kirik, V., Schnittger, A., Radchuk, V., Adler, K., Hülskamp, M., and Bäumlein, H. (2001). Ectopic expression of the arabidopsis AtMYB23 gene induces differentiation of trichome cells. *Dev. Biol.* 235, 366–377. doi: 10.1006/dbio.2001.0287

Kirik, V., Simon, M., Huelskamp, M., and Schiefelbein, J. (2004). The ENHANCER OF TRY AND CPC1 gene acts redundantly with TRIPTYCHON and CAPRICE in trichome and root hair cell patterning in arabidopsis. *Dev. Biol.* 268, 506–513. doi: 10.1016/j.ydbio.2003.12.037

Koch, M., Haubold, B., and Mitchell-Olds, T. (2001). Molecular systematics of the brassicaceae: evidence from coding plastidic matK and nuclear chs sequences. *Am. J. Bot.* 88, 534–544. doi: 10.2307/2657117

Koch, M. A., Kiefer, C., Ehrich, D., Vogel, J., Brochmann, C., and Mummenhoff, K. (2006). Three times out of Asia minor: the phylogeography of arabis alpina l. (Brassicaceae). *Mol. Ecol.* 15, 825–839. doi: 10.1111/j.1365-294X.2005.02848.x

Kreutz, C., Raue, A., Kaschek, D., and Timmer, J. (2013). Profile likelihood in systems biology. *FEBS J.* 280, 2564–2571. doi: 10.1111/febs.12276

Lee, M. M., and Schiefelbein, J. (1999). WEREWOLF, a MYB-related protein in arabidopsis, is a position-dependent regulator of epidermal cell patterning. *Cell* 99, 473–483. doi: 10.1016/S0092-8674(00)81536-6

Lee, M. M., and Schiefelbein, J. (2001). Developmentally distinct myb genes encode functionally equivalent proteins in arabidopsis. *Development* 128, 1539–1546. doi: 10.1242/dev.128.9.1539

Marks, M. D., Esch, J., Herman, P., Sivakumaran, S., and Oppenheimer, D. (1991). A model for cell-type determination and differentiation in plants. *Symp. Soc. Exp. Biol.* 45, 77–87.

Meinhardt, H., and Gierer, A. (2000). Pattern formation by local self-activation and lateral inhibition. *Bioessays* 22, 753–760. doi: 10.1002/1521-1878(200008)22:8<753::AID-BIES9>3.0.CO;2-Z

Mejia-Guerra, M. K., Pomeranz, M., Morohashi, K., and Grotewold, E. (2012). From plant gene regulatory grids to network dynamics. *Biochim. Biophys. Acta (BBA)-Gene Regul. Mech.* 1819, 454–465.

Morohashi, K., Zhao, M., Yang, M., Read, B., Lloyd, A., Lamb, R., et al. (2007). Participation of the arabidopsis bHLH factor GL3 in trichome initiation regulatory events. *Plant Physiol.* 145, 736–746. doi: 10.1104/pp.107.104521

Murray, J. D. (2001). "I. an introduction," in *Mathematical biology*, vol. 17. (Springer Science & Business Media).

Oppenheimer, D. G., Herman, P. L., Sivakumaran, S., Esch, J., and Marks, M. D. (1991). A myb gene required for leaf trichome differentiation in arabidopsis is expressed in stipules. *Cell* 67, 483–493. doi: 10.1016/0092-8674(91)90523-2

Pattanaik, S., Patra, B., Singh, S. K., and Yuan, L. (2014). An overview of the gene regulatory network controlling trichome development in the model plant, arabidopsis. *Front. Plant Sci.* 5. doi: 10.3389/fpls.2014.00259

Payne, C. T., Zhang, F., and Lloyd, A. M. (2000). GL3 encodes a bHLH protein that regulates trichome development in arabidopsis through interaction with GL1 and TTG1. *Genetics* 156, 1349–1362. doi: 10.1093/genetics/156.3.1349

Pesch, M., and Hülskamp, M. (2009). One, two, three… models for trichome patterning in arabidopsis? *Curr. Opin. Plant Biol.* 12, 587–592. doi: 10.1016/j.pbi.2009.07.015

Pesch, M., and Hülskamp, M. (2011). Role of TRIPTYCHON in trichome patterning in arabidopsis. *BMC Plant Biol.* 11, 1–14. doi: 10.1186/1471-2229-11-130

Pesch, M., Schultheiß, I., Klopffleisch, K., Uhrig, J. F., Koegl, M., Clemen, C. S., et al. (2015). TRANSPARENT TESTA GLABRA1 and GLABRA1 compete for binding to GLABRA3 in arabidopsis. *Plant Physiol.* 168, 584–597. doi: 10.1104/pp.15.00328

Purugganan, M. D. (1998). The molecular evolution of development. *Bioessays* 20, 700–711. doi: 10.1002/(Sici)1521-1878(199809)20:9<700

Saltelli, A. (2008). "Global sensitivity analysis," in *The primer* (Chichester, England ; Hoboken, NJ: John Wiley).

Schellmann, S., Schnittger, A., Kirik, V., Wada, T., Okada, K., Beermann, A., et al. (2002). TRIPTYCHON and CAPRICE mediate lateral inhibition during trichome and root hair patterning in arabidopsis. *EMBO J.* 21, 5036–5046. doi: 10.1093/emboj/cdf524

Schember, I., and Halfon, M. S. (2022). Common themes and future challenges in understanding gene regulatory network evolution. *Cells* 11, 510. doi: 10.3390/cells11030510

Simpson, P. (2002). Evolution of development in closely related species of flies and worms. *Nat. Rev. Genet.* 3, 907–917. doi: 10.1038/nrg947

Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics Comput. simulation* 55, 271–280. doi: 10.1016/S0378-4754(00)00270-6

Stephan, L., Tilmes, V., and Hulskamp, M. (2019). Selection and validation of reference genes for quantitative real-time PCR in arabis alpina. *PloS One* 14, e0211172. doi: 10.1371/journal.pone.0211172

Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *J. Am. Stat. Assoc.* 69, 730–737. doi: 10.1080/01621459.1974.10480196

Stracke, R., Werber, M., and Weisshaar, B. (2001). The r2r3-myb gene family in arabidopsis thaliana. *Curr. Opin. Plant Biol.* 4, 447–456. doi: 10.1016/S1369-5266(00)00199-0

Szymanski, D. B., Jilk, R. A., Pollock, S. M., and Marks, M. D. (1998). Control of GL2 expression in arabidopsis leaves and trichomes. *Development* 125, 1161–1171. doi: 10.1242/dev.125.7.1161

Tominaga, R., Iwata, M., Sano, R., Inoue, K., Okada, K., and Wada, T. (2008). Arabidopsis caprice-like myb 3 (cpl3) controls endoreduplication and flowering development in addition to trichome and root hair formation. doi: 10.1242/dev.017947

Tominaga-Wada, R., Nukumizu, Y., Sato, S., Kato, T., Tabata, S., and Wada, T. (2012). Functional divergence of myb-related genes, werewolf and atmyb23 in arabidopsis. *Biosci. biotechnol Biochem.* 76, 883–887. doi: 10.1271/bbb.110811

Turing, A. M. (1952). The chemical basis of morphogenesis. *Bull. Math. Biol.* 52, 153–197. doi: 10.1016/S0092-8240(05)80008-4

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., et al. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3, 1–12. doi: 10.1186/gb-2002-3-7-research0034

van Nocker, S., and Ludwig, P. (2003). The WD-repeat protein superfamily in arabidopsis: Conservation and divergence in structure and function. *BMC Genomics* 4, 1–11. doi: 10.1186/1471-2164-4-50

Wada, T., Tachibana, T., Shimura, Y., and Okada, K. (1997). Epidermal cell differentiation in arabidopsis determined by a myb homolog, CPC. *Science* 277, 1113–1116. doi: 10.1126/science.277.5329.1113

Wang, S., and Chen, J.-G. (2008). Arabidopsis transient expression analysis reveals that activation of GLABRA2 may require concurrent binding of GLABRA1 and GLABRA3 to the promoter of GLABRA2. *Plant Cell Physiol.* 49, 1792–1804. doi: 10.1093/pcp/pcn159

Wang, S., Hubbard, L., Chang, Y., Guo, J., Schiefelbein, J., and Chen, J.-G. (2008). Comprehensive analysis of single-repeat R3 MYB proteins in epidermal cell patterning and their transcriptional regulation in arabidopsis. *BMC Plant Biol.* 8, 1–13. doi: 10.1186/1471-2229-8-81

Wang, S., Kwak, S.-H., Zeng, Q., Ellis, B. E., Chen, X.-Y., Schiefelbein, J., et al. (2007). Trichomeless1 regulates trichome patterning by suppressing glabra1 in arabidopsis. doi: 10.1242/dev.009597

Wester, K., Digiuni, S., Geier, F., Timmer, J., Fleck, C., and Hulskamp, M. (2009). Functional diversity of R3 single-repeat genes in trichome development. *Development* 136, 1487–1496. doi: 10.1242/dev.021733

Willing, E. M., Rawat, V., Mandakova, T., Maumus, F., James, G. V., Nordstrom, K. J., et al. (2015). Genome expansion of arabis alpina linked with retrotransposition and reduced symmetric DNA methylation. *Nat. Plants* 1, 14023. doi: 10.1038/nplants2014.23

Zhang, F., Gonzalez, A., Zhao, M., Payne, C. T., and Lloyd, A. (2003). A network of redundant bHLH proteins functions in all TTG1-dependent pathways of arabidopsis. *Development* 130, 4859–4869. doi: 10.1242/dev.00681

Zhang, B., and Schrader, A. (2017). TRANSPARENT TESTA GLABRA 1-dependent regulation of flavonoid biosynthesis. *Plants* 6, 1–30. doi: 10.3390/plants6040065

Zhao, M., Morohashi, K., Hatlestad, G., Grotewold, E., and Lloyd, A. (2008). The TTG1-bHLH-MYB complex controls trichome cell fate and patterning through direct targeting of regulatory loci. *Development* 135, 1991–1999. doi: 10.1242/dev.016873

Zimmermann, I. M., Heim, M. A., Weisshaar, B., and Uhrig, J. F. (2004). Comprehensive identification of arabidopsis thaliana MYB transcription factors interacting with R/B-like BHLH proteins. *Plant J.* 40, 22–34 8. doi: 10.1111/j.1365-313X.2004.02183.x