System drift in the evolution of plant meristem development

Pjotr L. van der Jagt [1,2], Steven Oud [1,2], and Renske M.A. Vroomans [1,2]

[1]Sainsbury Laboratory, University of Cambridge, Bateman Street, Cambridge CB2 1LR, UK

[2]Department of Genetics, Downing Site, University of Cambridge, Cambridge CB2 3EJ, UK

Author Summary

A key open question in evolution of development (evo-devo) is the evolvability of complex phenotypes. Developmental system drift (DSD) contributes to evolvability by exploring different genotypes with similar phenotypic outcome, but with mutational neighbourhoods that have different, potentially adaptive, phenotypes. We investigated the potential for DSD in plant development using a computational model of developmental evolution. We found that the regulatory interactions between genes changed extensively, resulting in the continual neutral rewiring of the gene regulatory network underpinning development. Even regulatory interactions that were essential for correct development were replaced over long evolutionary time scales. Using plant genome and gene expression data from two publicly available datasets, we found high turnover of conserved non-coding sequences, which often contain regulatory sequences, occurring at both short and long time scales. This did not correlate consistently with gene expression changes in plant tissue, supporting the prevalence of DSD as predicted by our model.

Abstract

Developmental system drift (DSD) is a process where a phenotypic trait is conserved over evolutionary time, while the genetic basis for the trait changes. DSD has been identified in models with simpler genotype-phenotype maps (GPMs), such as RNA folding, however the extent of DSD in more complex GPMs, such as developmental pattern formation, is debated. To investigate the occurrence of DSD in complex developmental GPMs, we constructed a multi-scale computational model of the evolution of gene regulatory networks (GRNs) governing plant meristem (stem cell niche) development. We found that, during adaptation, some regulatory interactions became essential for the correct expression of stem cell niche genes. These regulatory interactions were subsequently conserved for thousands of generations. Nevertheless, we observed that these deeply conserved regulatory interactions could be lost over the extended period of stabilising evolution. These losses were compensated by changes elsewhere in the GRN, which then became conserved as well. This gain and loss of regulatory interactions resulted in a continual *cis*-regulatory rewiring in which accumulated changes caused changes in the expression of several genes. Using two publicly available datasets we found frequent changes in conserved non-coding sequences across six evolutionarily divergent plant species, and showed that these changes do not correlate with changes in gene expression patterns, demonstrating the occurrence of DSD. These findings align with the results from our computational model, showing that DSD is pervasive in the evolution of complex developmental systems.

Developmental system drift | Computational biology | Cis-regulation | Evolutionary developmental biology

Correspondence: *plv21@cam.ac.uk*

Introduction

Phenotypic traits are often conserved between related species, even when the developmental process that generates them diverged significantly (Cotterell and Sharpe, 2010; Krol et al., 2011; Verster et al., 2014; Brun-Usan et al., 2022). This phenomenon is called developmental system drift (DSD) or phenogenetic drift (True and Haag, 2001; McColgan and DiFrisco, 2024; Weiss and Fullerton, 2000). It can result from compensatory mutations after adaptive change in a pleiotropic gene (Johnson and Porter, 2007; Pavlicev and Wagner, 2012), or from neutral mutations that change the genotype but not the phenotype (Wagner, 2008a; Crombach et al., 2016). The potential for DSD in the latter case depends on the number of genotypes resulting in the same phenotype and how they are mutationally connected – aka how long is the neutral path in genotype-phenotype space (Alberch, 1991; Huynen, 1996; Hogeweg, 2000b). DSD can drive speciation (Schiffman and Ralph, 2022), may accelerate adaptation (Zheng et al., 2019, 2020), and is a possible evolutionary mechanism giving rise to the developmental hourglass (Wotton et al., 2015; Crombach et al., 2016).

However, it remains poorly understood how complex phenotypes are distributed throughout genotype space within highly complex GPMs: do neutral paths continue to percolate through genotype space, or do complex phenotypes occur in isolated genotype islands, limiting the extent of DSD? Theoretical and computational models of high-dimensional GPMs, such as RNA folding (Schuster et al., 1994; Huynen, 1996; Wagner, 2008b; Van Nimwegen et al., 1999), protein folding (Kimura, 1979; Lipman and Wilbur, 1991), gene regulatory networks (GRNs) (Ciliberti et al., 2007; Payne et al., 2014) have revealed extensive neutral paths that allow for significant genotype variation while maintaining the same phenotype. For other GPMs, such as those resulting from multicellular development, it has been suggested that complex phenotypes are sparsely distributed in genotype space, and have low potential for DSD because the number of neutral mutations anti-correlates with phenotypic complexity (Orr, 2000; Hagolani et al., 2021). On the other hand, theoretical and experimental studies in nematodes and fruit flies have shown that DSD is present in a phenotypically complex context (Verster et al., 2014; Crombach et al., 2016; Jaeger, 2018). It therefore remains debated how much DSD actually occurs in species undergoing multicellular development.

DSD in plants has received little attention. One multicellular structure which displays evidence of DSD in plants is the shoot apical meristem (SAM), which are multicellular structures containing a stem cell niche that generate all above-ground plant tissues (Szövényi et al., 2019). Differences in both morphology and gene expression across various vascular plant lineages suggest that SAMs originated independently multiple times before the emergence of the angiosperm clade (Frank et al., 2015; Harrison, 2015; Jill Harrison, 2017; Fouracre and Harrison, 2022). Within angiosperms however, SAMs are generally regarded as homologous structures due to their structural similarity and strong overlap in associated genes. Some key regulators of SAM stem cells, such as the CLAVATA3/Embryo Surrounding Region-Related (CLE) peptide family, are deeply conserved among land plants (Whitewoods et al., 2018). Nevertheless, the precise CLE peptides governing SAM function differ between angiosperm species, and their expression patterns also vary (Somssich et al., 2016), suggesting that DSD plays a role in SAM evolution.

To investigate the potential for DSD in plant SAMs, we developed a computational model of gene regulatory network evolution, extending an evo-devo approach previously applied to animal development (ten Tusscher and Hogeweg, 2011; Vroomans et al., 2016, 2018). In our evolutionary simulations, a small number of regulatory interactions become highly conserved as GRNs evolve to generate a functional tissue pattern. Surprisingly, we found that even these deeply conserved, essential regulatory interactions can diverge over longer evolutionary timescales, resulting in concomitant shifts in gene expression patterns and DSD. To validate these theoretical findings we performed a bio-informatics analysis on two publicly available datasets: one on conserved non-coding sequences (CNSs) in plants (Hendelman et al., 2021), and another on organ-specific RNA expression across plant species at different evolutionary distances (Schuster et al., 2026). By combining the CNS data with the cross-species RNA expression, we found that entirely different sets of CNSs can still have similar gene expression patterns, providing empirical evidence for a many-to-one GP mapping underlying plant gene expression and regulatory rewiring. Altogether these findings highlight the prevalence of DSD and its role in shaping developmental evolution in plants.

Results

Model overview

We developed a computational model of gene expression evolution in shoot apical meristems (SAMs). We modelled a population of SAMs, each undergoing a developmental process encoded by a heritable genome. A genome consists of a string of genetic elements representing genes (each of which encodes a transcription factor (TF)), and transcription factor binding sites (TFBSs) which regulate the expression of the downstream gene (Fig. 1A) (Crombach and Hogeweg, 2008). The genome therefore encodes a gene regulatory network (GRN), which governs gene expression in the cells of a two-dimensional tissue representing a longitudinal cut through the SAM (Fig. 1B-D). A subset of TFs exhibit specific properties, such as the ability to diffuse, form dimers, or mediate cell-cell communication with directly neighbouring cells (Methods, section D). Gene expression and protein production are subject to a small amount of molecular noise, modelled with stochastic differential equations (Methods, section E). Development therefore consists of spatiotemporal changes in protein distribution within the tissue due to gene expression, diffusion and noise (Fig. 1D).

The development of each individual begins with one TF uniformly distributed throughout the tissue, while a second diffusible TF is constitutively expressed in the epidermal layer (L1) (Fig. 1C). From this initial condition, individuals have a fixed amount of time to express other genes based on the interactions encoded by the genome. At the end of this period, each individual is assigned a fitness score based on the protein concentration of two target genes in specific regions of the SAM: one in the central zone (CZ), and one in the organizing center (OC) (Fig. 1C). This fitness score determines the probability that the individual will produce offspring in the next generation (Fig. 1E). During reproduction, the parent's genome is inherited by the offspring with random mutations (Fig. 1F). This cycle of development and reproduction with mutation is repeated for 50,000 generations.

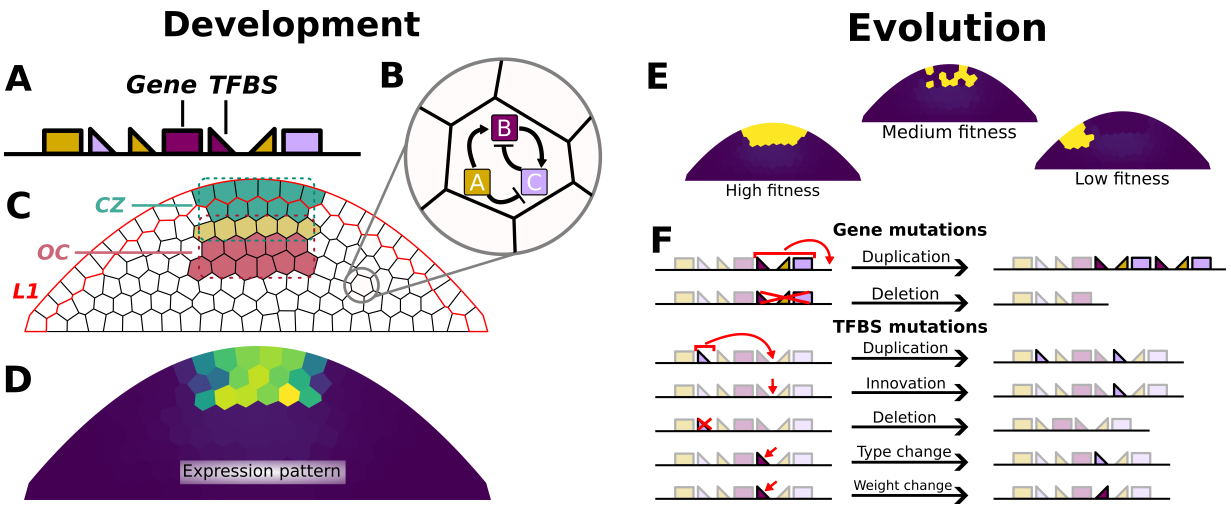

Figure 1. Overview of the computational evo-devo model. **A**. A schematic representation of an *in silico* genome including genes and Transcription Factor Binding Sites (TFBSs). **B**. Such a genome can be translated into a gene regulatory network (GRN) which is used to simulate the development. **C**. The different functional areas in the simulated tissues. The L1 layer is shown as red-lined cells, the CZ is shown in the tip of the tissue, the OC is in the center, and their regions overlap in a single row of cells. A cell being in the CZ/OC area is determined by their centroid being in the respective bounding box (indicated with dotted lines), see Methods F. **D**. An example of the expression pattern of the CZ gene at the end of development in a fit individual. **E**. Selection is based on the expression of fitness genes, shown here are examples of a high fitness, medium fitness and low fitness CZ gene expression pattern (left-to-right). **F**. The different types of structural genome mutations possible during the simulations. In addition to these mutations, mutations are possible in the binding constants of TFBSs, and the maximum transcription rates of genes.

Conservation of regulatory interactions during evolution of developmental programs

We ran 20 simulations, each with a constant population size of 1000 individuals. Out of the 20 populations, 15 evolved to correctly express both the OC and CZ genes, resulting in a fitness (f) score of at least 75 out of 100 (Fig. 2A, examples of high ($\geq$75) and low ($<$75) fitness patterns in B). To investigate how the developmental programs evolved that generated the pattern, we tracked how long each regulatory interaction between TFs was conserved in the ancestral lineage of each high-fitness population. Most interactions persisted for only a few hundred to a few thousand generations (Fig. 2C), in the same pattern as a control simulation (C) run with random reproduction and no

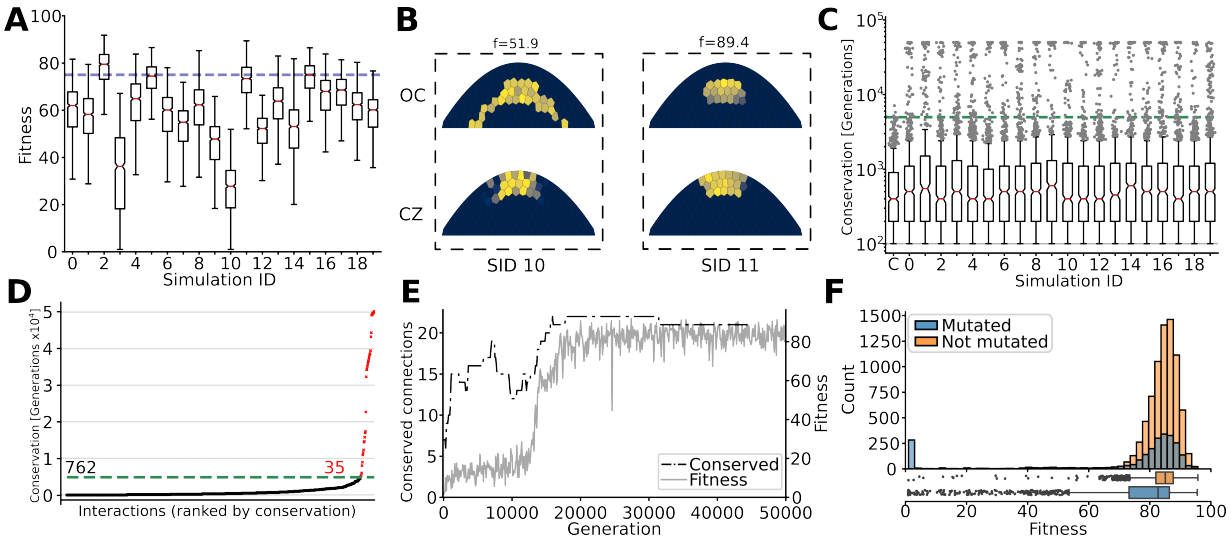

Figure 2. Fitness increase is correlated with GRN conservation. **A**. Fitness of all individuals in the 20 populations at generation 50,000 (outliers not shown). Dotted line denotes the cutoff fitness of 75. **B**. Protein pattern of the two fitness genes for the fittest individuals of simulations 10 and 11, with a fitness of 51.9 and 89.4, respectively. **C**. The number of generations each regulatory interaction was conserved in the ancestor trace of a population. The box plot indicated with C is a control simulation without selection, resulting in the absence of highly conserved interactions. Dotted line indicates the 'conservation cutoff' determined by the maximum conservation time of interactions within the control simulation (7100 generations), which we rounded down to 5000 for the rest of this work. **D**. The sorted conservation times of all regulatory interactions within the ancestry trace of simulation 2. In red are shown all interactions with a conservation time greater than the conservation cutoff of $>$5000 generations. **E**. Number of conserved ($>$5000 generations) interactions and fitness of ancestor trace for the first 50,000 generations of simulation 2. **F**. Fitness distribution of 10,000 randomly generated offspring of the fittest individual from simulation 2 at generation 50,000. Orange: offspring inheriting genomes without any mutations n =7503 IQR=5.69; Blue: offspring with mutation(s) n =2497 IQR=13.24.

selection for a pattern. However, a small but significant subset of interactions, which we call "conserved interactions," remained consistently present for more than 5000 generations, which did not occur in the control simulation (Fig. 2C, D). These conserved interactions emerged as individuals in a population achieved higher fitness (Fig. 2E, S1, control in S2). This suggests that deeply conserved interactions are a consequence of selection.

To assess the potential for neutral evolution and DSD after the target expression pattern evolved, we created 10,000 offspring of the highest-fitness individual of population 2 at generation 50,000 (Fig. 2F). As mutations are probabilistic, the majority of offspring inherit the parental genome without any mutations. Any variation in fitness of these non-mutated offspring results from variation in cell connectivity in the SAM (Methods, section C) and noisy gene expression (Methods, section E). We found that fitness variation in these non-mutated individuals follows a tight distribution around a high mean fitness, showing that the developmental mechanism is robust to these sources of noise (Fig. 2F, orange). In offspring which inherited a genome with mutations, the fitness distribution instead followed a U shape (Fig. 2F, blue). The majority of mutations was (near) neutral, indicating a high degree of mutational robustness and redundancy in the GRN, while a smaller set was very deleterious, resulting in a near-zero fitness. Mutations resulting in an 'in between' fitness were more rare, consistent with previous findings on fitness landscapes (Hogeweg, 2012). Overall this shows that the evolved genotypes are both developmentally and mutationally robust which indicates the existence of neutral areas within the GPM around this complex expression phenotype.

Developmental system drift in evolved gene regulatory networks

DSD occurs when the phenotype remains conserved along the ancestral lineage while the underlying regulatory architecture changes. Our simulations matched the target phenotype by 30,000 generations, after which they entered a fitness plateau where evolution was mostly driven by stabilising selection (Fig. S1). To study DSD at this fitness plateau, we selected the eight populations which reached the highest maximum fitness, created five clones of each, and evolved these for an additional 50,000 generations (Fig. 3A). The median fitness increase remained indistinguishable from background variability during this period (Fig. S3), indicating predominantly neutral evolution (we excluded the clones of 2/8 populations due to more significant fitness increase, Fig. S3, Appendix S1). Next, we investigated whether developmental or mutational robustness increased over this time period, which could explain small fitness gains of the population over longer periods of time. During the fitness plateau, mutational robustness and developmental robustness both fluctuate between generations, without either of their distributions

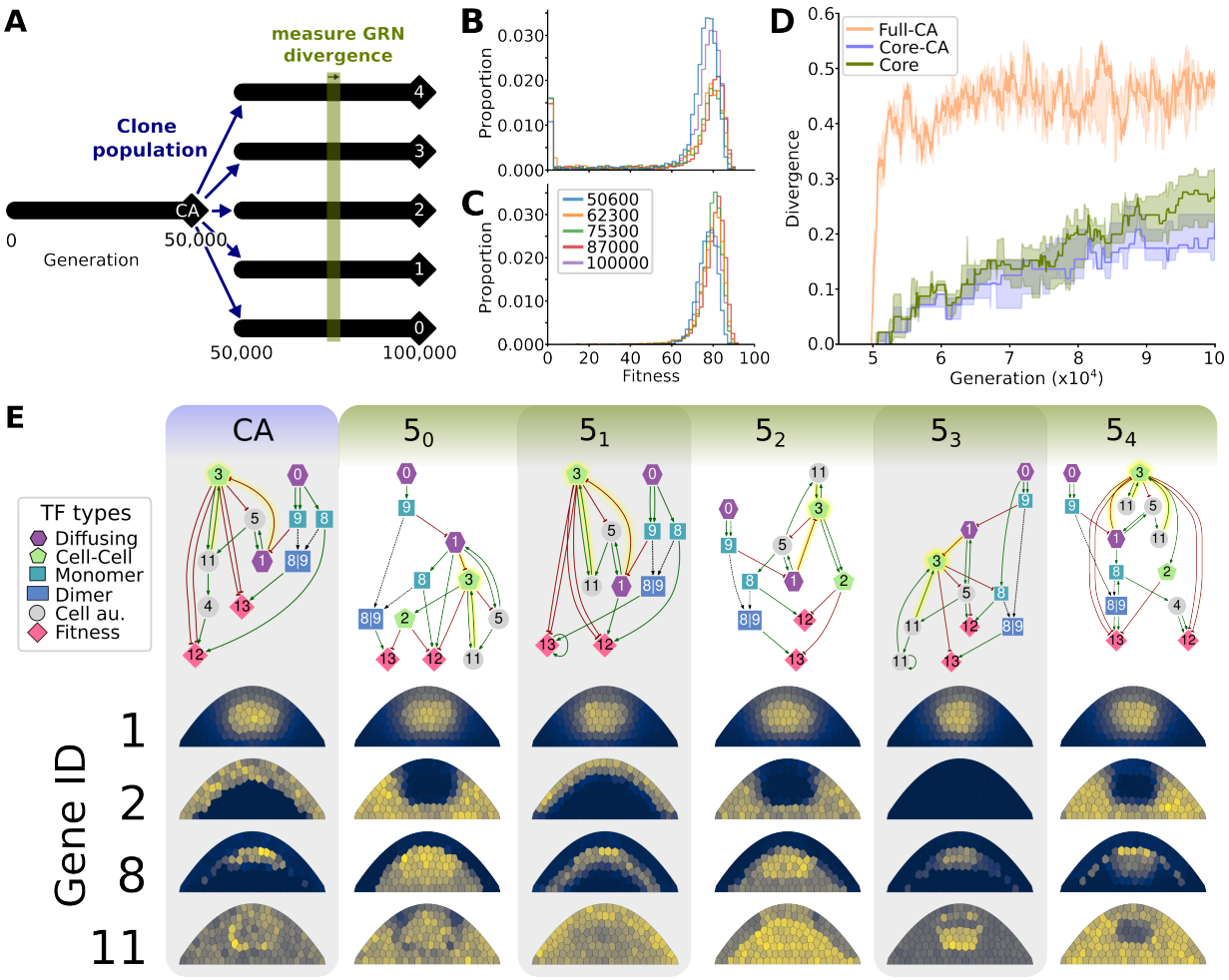

Figure 3. Neutral evolutionary change in regulatory mechanisms. **A**. At generation 50,000, populations that reached a fitness plateau are cloned into 5 separate populations that continue to evolve independently but share a common ancestor before generation 50,000. **B**. Fitness distributions of mutated offspring (mutational robustness) of related ancestors from different generations. **C**. Fitness distributions of clones (developmental robustness) of related ancestors from different generations. **D**. Divergence of interactions between lineages of each clonal population. Divergence is shown for lineages of the full GRNs with that of the common ancestor of all lineages (Full-CA); only the conserved set of interactions with the common ancestor (Core-CA, > 5000 generations); and the conserved set of interactions between the populations (Core). Shown are the medians and IQRs. Divergence between networks is calculated as in Methods J. **E**. Functional networks (Methods H) and expression patterns of the five fittest individuals at generation 100,000 from simulations 5_0 to 5_4, and their common ancestor (CA) at generation 49,800. The top row shows the functional networks, that descend from the CA in different cloned populations. The expression patterns of genes 1,2,8 and 11 are shown to illustrate different levels of phenotypic divergence. For instance, expression of gene 2 diverged significantly between some lineages, whereas expression of gene 1 is very conserved. Different TF types indicated by symbols, see Methods D.

becoming consistently higher or lower compared to ancestral states, indicating drift rather than adaptation of robustness (Fig. 3B,C; Appendix S2).

We quantified genetic divergence of individuals from their ancestor before the cloning at generation 50,000, using a divergence score based on the adjacency matrices of their GRNs (Methods J). We found that the full GRNs diverged rapidly from the common ancestor (CA) (Fig. 3D, S5, orange line), which was likely due to changes in the redundant or non-functional parts in the GRN, which can be mutated without any phenotypic consequence (see Fig. 2F). Therefore, we also measured divergence by including only the conserved interactions in the adjacency matrix (Fig. 3D, S5, blue line). Strikingly, these conserved interactions gradually diverged from the CA, indicating turnover despite their initial conservation. Since conserved interactions emerge during adaptation (Fig. 2E) and are therefore likely functionally important, their turnover may indicate that the regulatory dynamics that generate the target pattern have changed as well. To assess whether conserved interactions follow similar evolutionary trajectories across independent lineages, we performed a pairwise comparison of evolved GRNs between cloned populations (Fig. 3D, S5, green line). We found that conserved regulatory interactions diverged between populations at a

similar rate as each lineage diverged from the CA. This indicates that overall, the different lineages follow different paths of divergence from the CA. To investigate whether this divergence can be explained by indirect selection on developmental robustness, we ran simulations without gene expression noise. In these simulations, GRNs still diverge after a fitness plateau has been reached (Fig. S4). Although this does not exclude developmental robustness playing a role in the GRN divergence of the noisy simulations, it does show it is not necessary for divergence. Mutational robustness exhibits a similar drift-like pattern of alternating increases and decreases to that observed in the noisy simulations.

To observe the divergence in GRNs more closely, we pruned the full GRNs to remove non-functional and redundant genes and interactions, and compared these pruned GRNs between individuals from different clonal populations, as well as with their CA (Fig. 3E). As expected, some regulatory interactions in these networks were highly conserved, e.g., the regulation of gene 3, which is consistently activated by gene 11 and inhibited by gene 1 across all GRNs. In contrast, other interactions were rewired, e.g., the activation of gene 8, which is driven by gene 0 in networks CA, $5_1, 5_3$ but by gene 1 in 5_0 and 5_4, and gene 5 in 5_2. Since these GRNs were pruned to eliminate redundancy and all descend from the same CA, differences in their regulatory interactions reflect functional divergence through neutral evolution, indicating the occurrence of DSD.

To understand the effects of this regulatory turnover on development, we examined the pattern of the proteins that were not subject to selection for a specific pattern. Indeed, we found that the patterns of some free proteins remained conserved, but the expression of other genes underwent significant change (Fig. 3E, S6, S7). For example, the expression of protein 1 was nearly identical among different populations, while proteins 2, 8, and 11 displayed varying degrees of divergence in pattern compared to the CA. Interestingly, protein 2 had varying patterns, but was absent in the pruned networks of CA, $5_1, 5_3$. In line with our findings in figure 3D, this shows that different lineages explored different parts of the neutral evolutionary space, where in some cases TFs were recruited to regulate the expression of the genes under selection, whereas in others the TF remained nonfunctional and accumulated neutral changes. As only 2 out of the 14 genes are under selection for a target pattern, a large number of genes is "free" to evolve, which might contribute to the necessary redundancy for rewiring. Nevertheless, we still observe network divergence in simulations with only 6 instead of 12 "free genes", showing that redundancies can be created even in more constrained GRNs (Fig. S8). Taken together, we found that DSD can drive functional divergence in the underlying GRN resulting in novel spatial expression dynamics of the genes not directly under selection.

Network redundancy creates space for rewiring

To understand how conserved and functional interactions can diverge without disrupting fitness, we traced conserved regulatory interactions over evolutionary time in two of the cloned populations. Consistent with our earlier observations, we found that over time, interactions were lost and new conserved interactions arose (Fig. 4A,B). To investigate the conditions under which an interaction can be rewired, we examined a single interaction ($0 \rightarrow 8$) which is conserved in the lineage of one population but lost in the lineage of a related population, as indicated by the ◀ in Fig. 4A. We measured in each generation how much its removal affected fitness: an importance value of 1 indicates a total loss of fitness after removal, while a value of 0 indicates no change in fitness (see Methods I).

In both lineages, the importance of the interaction ($0 \rightarrow 8$) remained high for the first 60,000 generations. However, in lineage 5_2, a new interaction ($5 \rightarrow 8$) emerged that activated gene 8, which coincides with a drop in the importance of $0 \rightarrow 8$. This reduction in importance provided an opportunity for the deletion of $0 \rightarrow 8$ without a significant loss of fitness. This example shows the general mechanism by which functional redundancy enables the turnover of regulatory interactions, causing neutral GRN divergence over long evolutionary timescales. We even observed this process among interactions regulating the fitness genes (Fig. 4C), which could be rewired multiple times in relatively quick succession. The TF taking over regulation did not necessarily have the same expression pattern as the original: e.g. gene 8, the new regulator of gene 12 at generation 60,000, has a different pattern from gene 4. However, they were both expressed around the center, which is where gene 4's regulation of gene 12 was active. This shows how the evolution of GRNs exploits overlaps in expression patterns to generate redundancies, allowing the rewiring of regulatory interactions.

Finally, we tracked each rewiring event from 4 simulations to investigate more closely how these redundancies emerge in the first place. A general intuition is that gene duplications give rise to redundant but functional copies which can diverge to perform a novel function (Ohno, 1970). We therefore compared the copy number of genes where a conserved interaction was rewired, to the copy number of genes where a non-conserved interaction was rewired; while we would expect that rewiring of conserved interactions is more likely for duplicated genes, we did not find such a bias (Fig. S9).

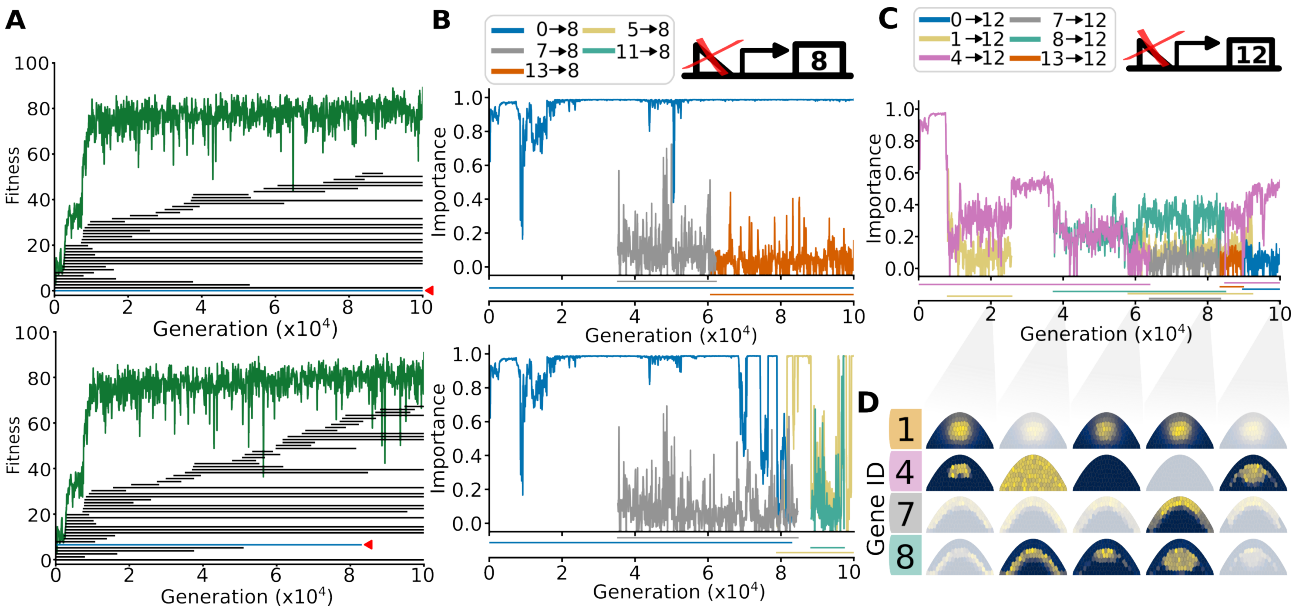

Figure 4. Rewiring of conserved interactions through loss of function. A. Presence of conserved interactions in two related lineages, every bar representing a specific conserved interaction. The green bars indicated with a ◀ show the interaction 0 activates 8. The lineages are 5_1 and 5_2, respectively. **B.** Importance of conserved interactions that upregulate gene 8 expression over evolutionary time of the respective simulations shown in A. The legend shows the TF regulating 8. Indicated under the plot is the persistence time of each interaction. **C.** Importance of conserved interactions upregulating gene 12 (fitness gene). **D.** Expression of subset of genes $(1, 4, 7, 8)$ upregulating expression of gene 12 at generations 20,000, 40,000, 60,000, 80,000 and 100,000. Pattern is transparent if there is no interaction between the respective gene and gene 12 at that time point.

DSD in plant gene regulation: Conservation, loss and emergence of deeply conserved non-coding sequences

Our computational simulations suggest that *cis*-regulatory rewiring plays a significant role in the divergence of developmental programs. To test this hypothesis, we used two available data sets: the Conservatory Project (Hendelman et al., 2021), and Schuster *et al.* Schuster et al. (2026). The Conservatory Project collects conserved non-coding sequences (CNSs) across plant genomes, which we used to investigate the extent of GRN rewiring in flowering plants. Schuster *et al.* measured gene expression in different homologous tissues of several species via bulk RNAseq, which we used to test for gene expression (phenotype) conservation, and how this relates to the GRN rewiring inferred from the CNSs. We constrained our search for regulatory rewiring to the six angiosperm species shared by these two data sets:*Arabidopsis thaliana, Arabidopsis lyrata, Capsella rubella, Eutrema salsugineum, Medicago truncatula,* and *Brachypodium distachyon*. This set of species contains both closely (*A. thaliana, A. lyrata, C. rubella*, *E. salsugineum*) and more distantly related species (*M. truncatula*, *B. distachyon*), which gives insight in both short and long term divergence.

We used CNSs from the Conservatory Project dataset as a proxy for *cis*-regulatory elements (Hendelman et al., 2021). A CNS is defined as a non-coding sequence conserved within the upstream/downstream region of genes within an orthogroup (Emms and Kelly, 2019). The dataset contains 2.3 million CNSs identified across 284 species. Starting with the orthogroup of a gene in the CNS dataset, we mapped all CNSs to their corresponding genes across the six species (Fig. 5A); we focused on 3 orthogroups containing genes relevant to meristem function (see Methods K).

One orthogroup of interest includes the MADS box transcription factors SEPALLATA (SEP) 1–4, which are important for organ specification in the floral meristem (Pelaz et al., 2000; Ditta et al., 2004). The SEP subfamily is monophyletic with a pre-angiosperm origin (Zahn et al., 2005; Ruelens et al., 2013). The different SEP homologs are thought to have emerged from several (whole genome) duplication events (Dreni and Ferrándiz, 2022). Within this SEP orthogroup, we identified 1658 CNSs, of which some are extremely conserved. For instance, 3 CNSs were conserved across all six species, 6 were conserved in eudicots, and 366 were conserved within Brassicaceae (Fig. 5B). The stark difference in shared CNSs between Brassicaceae and more distantly related species highlights a high turnover rate of CNSs, a pattern observed across all genes investigated (Fig. S10). Additionally, the data revealed likely cases of loss of otherwise deeply conserved CNSs in particular species, indicated by all non-monophyletic presence profiles (Fig. 5B). For example, 104 CNSs were lost only in *C. rubella* whilst conserved in the other Brassicaceae. This pattern – CNSs lost in one lineage but conserved in others – resembles the rewiring dynamics we observed in

our simulations.

Next, we examined the CNS composition of individual *SEP* genes to gain a more fine-grained view of regulatory rewiring. We found that CNSs are mostly, but not completely conserved between orthologs (see for instance SEP4 and orthologs, Fig. 5C,D) Paralogs, such as *SEP3* and *SEP4*, tend to have different CNSs, likely due to rewiring after the duplication event. Interestingly, some paralogous gene pairs, such as *SEP1* and *SEP2*, showed partially overlapping CNSs, which may reflect their origin from a relatively recent duplication event. Based on our simulations, we predict that the CNSs still shared between these genes have a higher regulatory importance.

In order to investigate these genetic changes in the light of DSD, the resulting 'phenotype' must be conserved, or more specifically, continuously present since their shared origin (DiFrisco, 2023). To remain close to our computational model (where the gene expression pattern of 2 genes constitutes the phenotype) we here compared gene expression patterns between species in the data set of (Schuster et al., 2026) to measure changes in phenotype, and investigated their correlation with changes in CNS composition (Fig. 5E). We assumed that historical continuity is more likely than repeated loss and gain for similar expression patterns. The SEP genes with similar CNS sets tend to exhibit similar expression patterns (Fig. 5E), which we quantified by calculating the Pearson distance between the expression patterns of all gene pairs, and then comparing this distance to the similarity in CNS set for each pair (Fig. 5F, S12, S13). Notably, genes with any CNS set similarity (> 0) rarely display high Pearson distances, suggesting that shared CNSs strongly correlate with similar expression patterns (Fig. 5F, S12C, S13C). However, dissimilar CNS sets do not necessarily imply different expression patterns (Fig. 5F). Most gene pairs with disjoint CNS sets (similarity $= 0$) exhibit a wide range of expression divergences, including pairs with very similar expression patterns. This is evidence of regulatory rewiring while preserving phenotype — a hallmark of DSD.

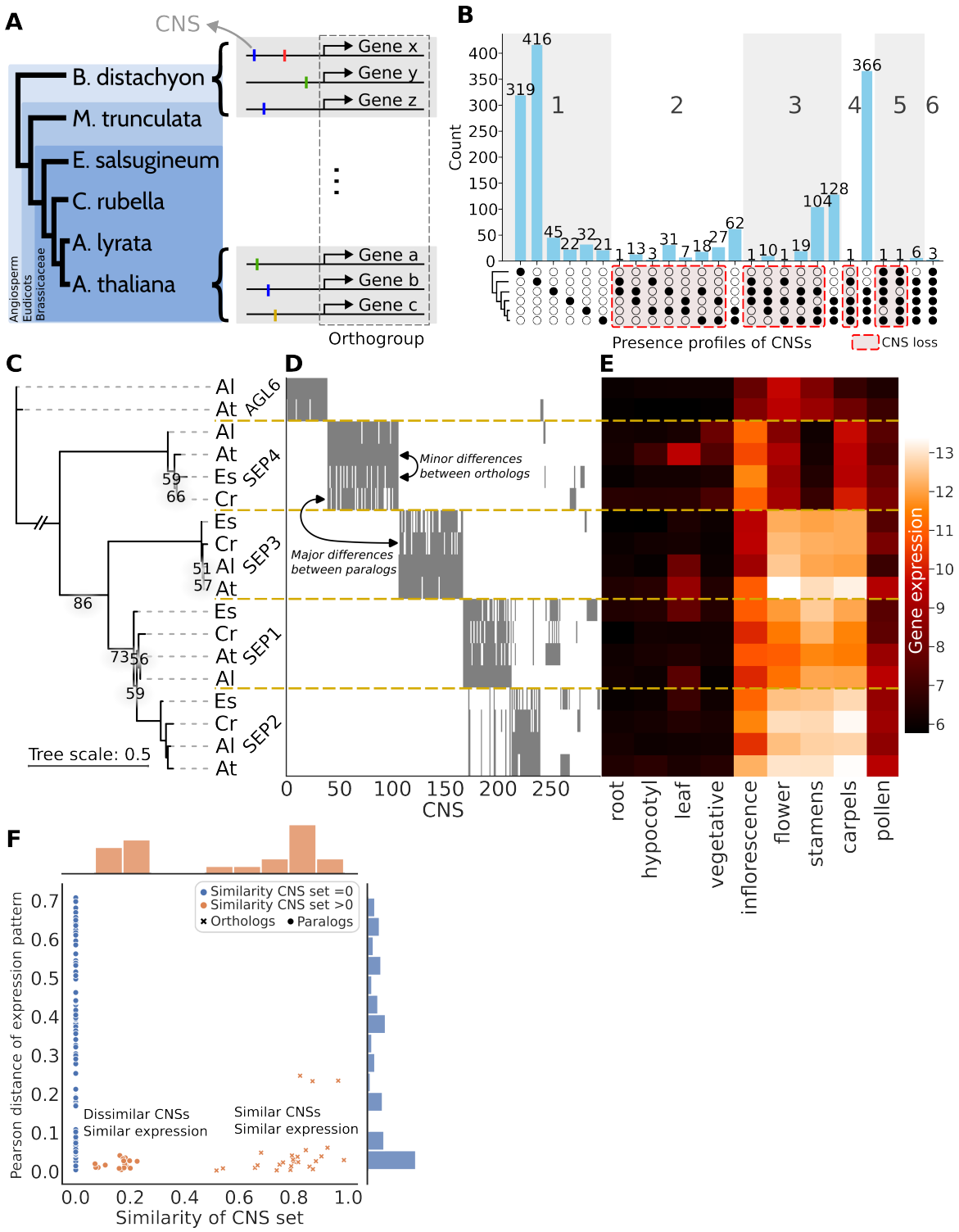

Figure 5. Cis-regulatory rewiring in SEP homologs between different angiosperm species. **A**. Schematic depicting the relation between orthogroup, different species, different genes and CNSs. **B**. UpSet plot of CNS presence in the six angiosperm species. The filled circles indicate in which species the CNSs are present, the barplot shows the count of number of CNSs with that profile. The red dashed boxes indicate cases of CNS loss (all non-monophyletic presence patterns). **C**. Gene tree of primary transcript protein sequences of SEP genes of the four Brassicaceae species, Es: *Eutrema salsugineum*; Cr: *Capsella rubella*; Al: *Arabidopsis lyrata*; At: *Arabidopsis thaliana*. Note that the branch to the outgroups has been truncated for readability, for a full tree see Supplementary Figure S11. Only bootstrap values less than 90% are shown. **D**. Heatmap of CNS presence for each of the SEP genes. Presence of a CNS is shown in grey. **E**. Expression profile of SEP genes in different organs. The expression data shown are variance stabilized mRNA counts taken directly from Schuster et al. (2026). **F**. Pearson distance ($1 - \text{pearson correlation}$) of expression pattern against the similarity of the CNS set for each gene pair in C. A similarity of 0 describes two completely disjoint sets of CNSs, whereas a similarity of 1 describes two identical sets. For each comparison we indicate if this is between orthologs or paralogs.

Discussion

In this study, we investigated developmental system drift (DSD) in a complex genotype-phenotype map (GPM) of plant development. We constructed a model of gene regulation evolution in the plant shoot apical meristem (SAM) to analyze genotypic and phenotypic changes over long evolutionary timescales. In our simulations we observed that a subset of gene interactions was conserved over thousands of generations, while other interactions turned over quickly. These interactions were highly important for generating the target gene expression pattern; nevertheless, mutations could result in redundancies that lowered the importance of such interactions, allowing for their eventual removal from the genome. We also found that such rewiring preserved the expression pattern under selection while facilitating the exploration of diverse expression patterns in genes that were not under direct selection, revealing a highly interconnected genotype space. Using data from the Conservatory project we showed gain and loss of conserved non-coding sequences (CNSs) in six vascular plant species (Hendelman et al., 2021). With the Schuster dataset (Schuster et al., 2026) we showed that these genomic changes did not necessarily result in gene expression (phenotype) changes, a hallmark of DSD.

Step-wise *cis*-regulatory rewiring

By tracking regulatory interactions over evolutionary time, we found that *cis*-regulatory rewiring of the GRN can render previously essential interactions redundant. This redundancy creates opportunities for further rewiring, resulting in a stepwise mechanism of regulatory evolution that can completely change the structure of the GRN over time. Our analysis also showed that the GRN is both mutationally robust (allowing for rewiring)(Wagner, 2008b) and sensitive to mutations at the sites of highly important interactions. This matches bioinformatic analysis of the rice genome, where it was found that while most promoter sites experience low selection pressure, a small but significant subset is under strong selection pressure (Joly-Lopez et al., 2020). The generally low selection pressure on promoter sites suggests that there is sufficient evolutionary space for neutral mutations to accumulate and drive the gradual rewiring of the GRN.

A concrete example of regulatory rewiring in the SAM can be found in two antagonistic peptide signaling pathways involving genes of the CLE family. The CLV3 pathway inhibits stem cell fate, whereas the CLE40 pathway promotes it (Schlegel et al., 2021). Interestingly, the CLV3 and CLE40 proteins are functionally similar, and a *clv3-2* mutant can be rescued by expression of *CLE40* under a *CLV3* promoter (Hobe et al., 2003). Similar results which show regulatory divergence in a conserved pathway have been found for a CNS related to the transcriptional regulator *UNUSUAL FLOWER ORGANS* in tomato and *A. thaliana* (Lanctot et al., 2025). These patterns appear to extend beyond plants, as evidenced by for example the rewiring of *Mcm1* in different yeast species (Tsong et al., 2006). These examples demonstrate that whilst the function of pathways is conserved, their regulatory wiring often is not.

Genetic diversity due to DSD

DSD necessitates the existence of multiple genotypes that map to the same trait. Such many-to-one mappings have been identified in various models of genotype-phenotype maps (GPMs). For example, Cotterell and Sharpe (2010) showed that a three-gene model of stripe formation can generate similar spatial patterns through many distinct regulatory interactions. In their model, alternative solutions clustered in a few disconnected regions in genotype space, corresponding to distinct regulatory mechanisms. This suggests a limited range of genotype space in which DSD could occur (Jiménez et al., 2015). Here, we show that large shifts in gene expression can accompany GRN rewiring during stabilising selection, suggesting that even the underlying regulatory mechanism can change without loss of the expression pattern under selection. This requires rewiring of conserved interactions, which (by definition) happens over long evolutionary timescales as opposed to the rapid rewiring of redundant interactions over few generations.

GRN rewiring through regulatory redundancy requires that genes can regulate and be regulated by multiple TFs. This many-to-many architecture is therefore key to DSD. A previous model by Johnson and Porter (2007) lacked this regulatory flexibility, they displayed DSD under directional selection but not during stabilising selection. Our model's more permissive and realistic network architecture allows for a phenotype to be encoded by a wider range of genotypes, facilitating accumulation of neutral changes. In this way, our system is conceptually similar to RNA folding GPMs, where very different sequences can produce the same secondary structure (Huynen, 1996; Van Nimwegen et al., 1999).

The structure of the GPM and DSD

The structure of the GPM determines the length of neutral paths for any given phenotype. For simpler GPMs, such as that of RNA folding and GRNs governing expression in single cells, neutral paths were shown to be extensive, and connect areas with different levels of mutational robustness as well as different mutant phenotypes (Huynen, 1996; Huynen et al., 1996; Van Nimwegen et al., 1999; Crombach and Hogeweg, 2008; Wagner, 2014). In a more complex GPM, such as morphological development, complex phenotypes were rare and mutations were found to

often reduce complexity, which led to the conclusion that neutral paths and evolvability could be less extensive for complex GPMs and phenotypes (Hagolani et al., 2021). It is possible that morphogenesis adds additional constraints to the types of regulatory changes that can maintain the same phenotype (Cano-Fernández et al., 2024). However, in a model where selection acted on the number of cell types produced by morphogenesis rather than morphology, the resulting fitness landscape was more connected, suggesting that morphogenetic constraints may not be so severe (Hogeweg, 2000a,b). Furthermore, in our model many mutations are also highly deleterious, as seen in Figure 2F, making the mutational landscape similar to that of the complex phenotypes in (Hagolani et al., 2021). It takes multiple generations for the evolutionary process to find and build on those few mutations that are neutral or beneficial. In future work, it will therefore be interesting to explore the length of neutral paths in models of morphogenesis such as (Hagolani et al., 2021) by running evolutionary simulations over extended time scales.

Modelling assumptions and choices

In our model, fitness is based on the spatiotemporal expression of two genes, here called fitness genes. Determining an individual's fitness based on expression patterns is common in evolutionary models (Draghi and Wagner, 2009; Payne and Wagner, 2014; Vroomans et al., 2016; Brun-Usan et al., 2021). However, 'true' fitness is generally defined as being a measure of reproduction and survival (Orr, 2009). Thus a more functional, less explicit, fitness function, described on the level of traits and their contribution to reproduction and survival would be a more accurate description. An example of this is the model in von der Dunk et al. (2022) on the evolution of cell-cycle regulation; there, an individual's fitness resulted from its ability to efficiently organise its cell cycle in concert with the available resources.

In our model, we include 12 genes that are not under direct selection, allowing them to be expressed in any pattern that supports the correct expression of the fitness genes. Even when reduced to only 6 genes not under direct selection, we still observe divergence in the GRNs. This is in contrast with other work, where selection acted on the pattern in one gene of a three-gene GRN (Cotterell and Sharpe, 2013). We argue that the presence of "free" genes aligns with the idea that selection does not directly act on gene expression, but the resulting functioning of the tissue. We therefore expect that the resulting DSD better reflects the drift occurring in naturally developing systems. We do find that in our evolved GRNs, many genes play an indirect role in the correct expression of the fitness genes, resulting in emergent constraints on their evolution. Nevertheless, on long time scales the expression and regulatory interactions of these emergently constrained genes can still undergo DSD.

Limitations of CNSs as CREs

The pattern of conservation and loss of CNSs in the Conservatory Project dataset show a remarkable similarity to our results, and we have thus far assumed CNS to be more or less synonymous with cis-regulatory elements (CREs). CNSs are in fact enriched in CREs (Freeling and Subramaniam, 2009), as also supported by a strong overlap between CNS distribution and ATAC-seq peaks (Hendelman et al., 2021; Amundson et al., 2025), and the fact that mutations within CNSs give rise to phenotypic mutations (Lanctot et al., 2025). However, they do not map one-to-one to CREs, and only provide indirect information about CRE evolution. Since we we cannot directly interpret the presence of a individual CNS as the presence of a conserved regulatory interaction, we only analyse the set of CNSs as a whole. Finally, the CNS data set may miss sequences that are gained and lost at shorter time scales than the sampling of species allows to observe, as well as conserved CRE/CNS that are located farther from the target genes (Ricci et al., 2019; Lu et al., 2019).

Future experimental validation

To test our predictions more quantitatively, we need to know which TFs bind the upstream region of a gene, as direct evidence of a regulatory interaction, instead of inference from CNSs. A recent preprint by Baumgart et al. (2024) exemplifies this approach, presenting experimental validation of TF binding sites across ten different species using a specialized DNA affinity purification sequencing (DAP-seq) technique. While DAP-seq screens do not provide conclusive evidence that the given TF-regulatory site interaction is important in a given developmental context (Bartlett et al., 2017), they do provide more direct insight into the regulatory elements that are conserved or lost across species. Furthermore, experimental work has shown that the effect of mutating a particular CRE strongly depends on the genetic background, also known as epistasis (Schaerli et al., 2018; Aguirre et al., 2023; Ciren et al., 2024). The presence of this cryptic variation, revealed by the differential effect of mutations, suggests that DSD already acts on short time scales. Extending the DAPseq and functional approach to multiple species and multiple CREs would allow us to quantify the extent of such DSD.

The gene expression data used in this study represents bulk expression at the organ level, such as the vegetative meristem (Schuster et al., 2026). This limits our analysis of the phenotypic effects of rewiring to comparisons between organs, which is different to our computational simulations where we look at within-organ gene expression. Additionally, the bulk RNA-seq does not allow us to discern whether the developmental outcome of similar gene expression is the same in all these species. More fine-grained approaches, such as single-cell RNA sequencing or

spatial transcriptomics, will provide a more detailed understanding of how gene expression is modulated spatially and temporally within complex tissues in different organisms, allowing for a closer alignment between computational predictions and experimental observations.

Acknowledgments

We thank Madelaine Bartlett for helpful discussions and support with the conservatory.org dataset. We thank Elliot Meyerowitz and Christoph Schuster for the use of their dataset. We thank the Sainsbury Laboratory Evolution Journal Club and Enrico Sandro Colizzi for their comments and suggestions on the manuscript.

Funding

This work was supported by Gatsby Charitable Foundation (G112566 to RMAV). The funder had no role in design, data collection and analysis, decision to publish, or preparation of the manuscript.

Methods

A. Code availability

All code for the evolutionary model, analysing simulation output, and the bioinformatics analysis is available at https://gitlab.developers.cam.ac.uk/slcu/teamrv/publications/vanderjagt_2025#.

B. Computational model

We developed a computational model of an evolving population of organisms that undergoes cycles of development and selection. Each organism possesses a “beads-on-a-string" genome consisting of genes and upstream regulatory sites (Crombach and Hogeweg, 2008). This genome encodes a gene regulatory network (GRN) that governs gene expression dynamics in a 2D cellularized tissue. A single cycle, or generation, of the population proceeds through the following steps: each individual undergoes development governed by the dynamics of its GRN; this results in gene expression patterns in the tissue, which form the phenotype of the individual; this phenotype is given a fitness value by evaluating how well it meets a target expression pattern; individuals are selected for reproduction based on their fitness value, passing on their genome to the next generation; when a genome is passed on, it is randomly mutated.

C. 2D Tissue generation

Before running the evolutionary simulations, 1000 different tissues (representing the shoot apical meristem, SAM) are generated with different placements of cells. Each tissue contains 130 cells, each cell is initialized as a point in two dimensional space ($x, y \in \mathbb{R}^2$) with domains $y = [0, -0.01x^2]$ and $x = [-61, 61]$. Cells are attached to their neighbours with springs. The potential energy of springs is $U_s = \frac{1}{2}(l - l_r)^2 k$ where l is the current state of the spring, l_r the relaxed state and k the spring constant. To prevent cells from moving out of the predefined tissue shape, a boundary energy is added $U_b = \frac{-10}{q}$ where q is the distance between a cell and the boundary. Cells are distributed further by relaxing the potential energy for 500 steps with a dt of 0.1 . Using these positions, a Voronoi diagram is computed and cells are moved for a further 2000 steps towards their centroid using the Lloyds algorithm (Du et al., 1999). The Voronoi diagrams obtained after this final step are stored and used as template for the developmental simulations. The position and number of cells doesn’t change within the developmental simulations.

D. The genotype

A genome is composed of two components: transcription factor binding sites (TFBSs) and genes (Fig. S14A). Genes encode for transcription factors (TFs) which each bind a specific TFBS type (denoted with an integer identifier). The TFBSs in front of a gene determine which TFs regulate the expression of that gene. Some TFs have specific properties, such as the ability to diffuse between cells or form dimer complexes, see Table 1. There can be multiple copies of genes and TFBSs.

TF type	TF identifier	Description
Diffusing	0,1	can diffuse through the tissue
Cell-Cell	2,3	influence translation of neighboring cells (and not the cell in which the TF is present) in a receptor-ligand type of interaction
Monomer	8,9	can either bind their respective TFBSs as monomer, or dimerise to a dimer
Dimer	7	can not be synthesized through transcription/translation. Formation of the dimer is only possible through synthesis of both monomers within the same cell, which then dimerise
Cell autonomous	4,5,6,10,11	do not diffuse, and only bind to their respective TFBSs within the cell
Fitness	12,13	cell autonomous TFs on which selection for a specific pattern is done.

Table 1. Table with the different TF types and their description used in the model.

The genomes of the individuals in the initial population (generation 0) are initialized with one of each gene type and a random number of TFBSs per gene, see Algorithm 1.

E. Development of genotype into phenotype

Before the development of an individual is computed, their genome is first translated into a reaction network (Fig. S14B). This reaction network is used to compute gene expression in the cells of a multicellular tissue (Fig. S14C).

Algorithm 1: Pseudocode of genome initialization

```
for i = 0 to n                                   // n ← number of gene types
    for j = 0 to r     // r ∈ {1,2,3} number of TFBSs, chosen random with equal probabilities
        create a TFBS of a random type
    create a gene of a random type (from the not yet used types)
```

The reaction network is modelled as a set of stochastic differential equations (SDEs) (Ahmad, 1988; Gillespie, 2007) governing mRNA, protein and TF dimer concentrations. The rate equations for mRNA (Eq. 1) and protein concentrations (Eq. 2) are given by

$$\mathrm{d}G_{g,c}(t) = \left[\tau_{g,c}(t) - \alpha G_{g,c}(t)\right]\mathrm{d}t + \sqrt{\tau_{g,c}(t)}\mathrm{d}W_1 + \sqrt{\alpha G_{g,c}(t)}\mathrm{d}W_2 \tag{1}$$

$$\begin{aligned}\mathrm{d}P_{p,c}(t) = &\left[\upsilon_{g,c}(t) - \beta P_{g,c}(t) + \psi_{g,c}(t) + \frac{D_p}{A_c}\sum_{k\in N_c}\omega_{c,k}[P_{p,k}(t) - P_{p,c}(t)]\right]\mathrm{d}t \\ &+ \sqrt{\upsilon_{g,c}(t)}\mathrm{d}W_3 + \sqrt{\beta P_{p,c}(t)}\mathrm{d}W_4.\end{aligned} \tag{2}$$

Here, $G_{g,c}$ denotes the concentration of mRNA of gene g in cell c, and $P_{p,c}$ denotes the concentration of protein p in cell c. With $\tau(t)$ the transcription rate; $\upsilon(t)$ the translation rate; α and β decay rates; $\psi(t)$ dimerization; the term with D describes diffusion; and the square root terms constitute the stochastic part.
The transcription rate of a gene (Eq. 3) is determined by the transcriptional activity $T_{g,c}$ (Eq. 4) which is the sum of contributions of the respective gene's upstream TFBSs.
The transcriptional contribution of a TFBS (Eq. 5) is described by $\theta_{i,c}(P_p)$ and depends on the Hill constant H_i and sign $w_i \in \{-1, 1\}$ of the respective TFBS, as well as the TF concentration of the TFBS's type (P_p), where the TFBS type is p. If the TFBS type is a cell-cell TF (see Table 1) the input protein concentration is the concentrations of neighbouring cells (N_c) weighted by the length of their contact edge ($\omega_{c,k}$) relative to the sum of edges of the cell (C_c), see Eq. 6.

$$\tau_{g,c}(t) = \tau_{\max,g}\frac{T_{g,c}(t)^2}{T_{g,c}(t)^2 + 1} \tag{3}$$

$$T_{g,c}(t) = \max\left(0, \sum_{i\in \mathrm{TFBS}_g}\theta_{i,c}(P_p)\right) \tag{4}$$

$$\theta_{i,c}(P_p) = \begin{cases} w_i\dfrac{P_{p,c}^2}{P_{p,c}^2 + H_i^2}, & \text{if } p_{\mathrm{type}} \text{ is not cell-cell} \\ w_i\dfrac{\hat{P}_{p,c}^2}{\hat{P}_{p,c}^2 + H_i^2}, & \text{if } p_{\mathrm{type}} \text{ is cell-cell.}\end{cases} \tag{5}$$

$$\hat{P}_{p,c} = \frac{1}{C_c}\sum_{k\in N_c}\omega_{c,k}P_{p,k} \tag{6}$$

The translation rate (Eq. 7) is linear with the amount of available mRNA.

$$\upsilon_{g,c}(t) = G_{g,c}(t) \tag{7}$$

The stochastic terms in Equations 1 and 2 are modelled as independent Wiener processes and are modelled similar to the implementation in (Gillespie, 2007; Loman and Locke, 2023). We only model transcription, translation and decay as stochastic processes, while dimer formation is always assumed to be in equilibrium.

F. Fitness function

During development the fitness of individuals is determined at time points within the intervals $t \in [200, 250]$ and $t \in [450, 500]$. Fitness depends on the protein concentration of the two fitness genes in each cell, P_{cz} and P_{oc} respectively, where $_{\mathrm{cz}}$ and $_{\mathrm{oc}}$ represent protein types 13 and 12, respectively. Expression of the fitness genes in the correct cells results in a positive fitness contribution, whereas miss-expression of fitness genes yields a negative fitness contribution. Each cell is assigned to one of the following four types, a CZ cell (type

1, domain $x = \langle -15, 15 \rangle, y > 28$); a CZ and OC cell (type 2, $x = \langle -15, 15 \rangle, y = \langle 24, 28 \rangle$); an OC cell (type 3, $x = \langle -15, 15 \rangle, y = \langle 16, 24 \rangle$); other cell, see Figure 1C. In order to prevent overexpression of fitness genes, the positive term of the fitness equations is capped at a concentration of $p = 400$, after which the fitness of a cell is set to 100. For each cell c their fitness contribution (f_c) is computed (see Eq. 8), and the sum of all cells is the total fitness for a time point. We sum the fitness contributions of each time point (t) and divide the resulting number by the number of cells (130) and number of time points in the fitness interval (100) to gain a fitness score (F) between 0 and 100, see Eq. 9.

$$f_c(t) = \begin{cases} \frac{100 \cdot P_{\text{cz},c}(t)}{P_{\text{cz},c}(t)+100} - \frac{50 \cdot P_{\text{oc},c}(t)}{P_{\text{oc},c}(t)+200}, & \text{if } c_{\text{type}} == 1 \\ \frac{50 \cdot P_{\text{cz},c}(t)}{P_{\text{cz},c}(t)+100} + \frac{50 \cdot P_{\text{oc},c}(t)}{P_{\text{oc},c}(t)+100}, & \text{if } c_{\text{type}} == 2 \\ \frac{100 \cdot P_{\text{oc},c}(t)}{P_{\text{oc},c}(t)+100} - \frac{50 \cdot P_{\text{cz},c}(t)}{P_{\text{cz},c}(t)+200}, & \text{if } c_{\text{type}} == 3 \\ -\frac{50 \cdot P_{\text{oc},c}(t)}{P_{\text{oc},c}(t)+200} - \frac{50 \cdot P_{\text{cz},c}(t)}{P_{\text{cz},c}(t)+200}, & \text{else.} \end{cases} \tag{8}$$

$$F = \max\left(0, \frac{1}{130 \cdot 100} \sum_t \sum_{c \in N} f_c(t)\right) \tag{9}$$

G. Selection and Mutation

Individuals are selected for reproduction proportional to their fitness F, with the probability (ρ) of selecting an individual i from population P:

$$\rho(i) = \frac{F_i}{\sum_{j \in P} F_j}. \tag{10}$$

Offspring inherits their parent's genome with mutations, such as duplications, deletions, de novo TFBS emergence (Stone and Wray, 2001; Duque and Sinha, 2015), and TFBS type switches (Tsong et al., 2006). All implemented mutations are given in Table 2.

Gene	TFBS
Deletion	Deletion
Duplication	Duplication
Transcription rate change ($\tau_{\max}$)	Hill constant change (H)
mRNA decay rate change (α)	Sign change (w)
	Type change
	De novo emergence
Whole genome duplication	

Table 2. The different mutation types implemented in the model. Indicated between brackets is the changed parameter if applicable, see E.

Besides mutations in the genome, we also allow for changes in the diffusion, association and dissociation constants of the diffusing and dimerising proteins and whole genome duplications. Whole genome duplications do not result in polyploidization, but simply double the genetic material by appending the genome with a full copy of itself. Mutations to parameters are done through addition of the old value and a random number from a normal distribution, where $\mu = 0$ and σ^2 is parameter specific, see Table S1. Additionally, if the new value is outside of a predefined domain (see Table S1), the value is set to it's closest bound, yielding the general equation for a mutation of a given parameter λ as

$$\lambda_{\text{new}} = \begin{cases} \lambda_{\max}, & \text{if } \lambda_{\text{old}} + \mathcal{N}(0, \sigma_\lambda^2) > \lambda_{\max} \\ \lambda_{\min}, & \text{if } \lambda_{\text{old}} + \mathcal{N}(0, \sigma_\lambda^2) < \lambda_{\min} \\ \lambda_{\text{old}} + \mathcal{N}(0, \sigma_\lambda^2), & \text{else.} \end{cases} \tag{11}$$

Only one mutation per element is possible per offspring, see Algorithm 2.

H. Functional network analysis

The evolved GRNs contain redundant and/or non-functional interactions. We obtain a 'functional network' from a full evolved network through a pruning algorithm. In the pruning algorithm all possible single deletions are performed on

Algorithm 2: Pseudocode of genome mutations upon offspring creation

```
n ← length of genome G
for i = 0 to n
    r ← U[0,1]      // random floating point number from a uniform distribution over [0,1)
    if G_i == Gene then
        k ← 0
        for ∀m ∈ M_Gene                                  // for each possible gene mutation
            k = k + ρ(m)                        // ρ(m) is the probability of the given mutation
            if r < k then
                perform mutation m on genomic element G_i
    else
        for ∀m ∈ M_TFBS                                  // for each possible TFBS mutation
            k = k + ρ(m)
            if r < k then
                perform mutation m on genomic element G_i
Check for TFBS innovations                      // only one innovation possible per offspring
Check for whole genome duplications
```

a genome, development run for this set of genomes (5 times per genome to account for stochastic development), and fitness is computed for all resulting expression patterns. The genome with the highest average fitness is accepted and the procedure is iterated until no mutations maintain a fitness within 5 points from the original fitness.

I. Importance analysis

The importance I of a regulatory interaction is calculated as the fraction of fitness loss after the interaction is removed. Since the model is stochastic, development is repeated R times after removal of the interaction and the average fitness is taken.

$$I = \frac{1}{R} \sum_{r \in R} \frac{F_0 - F_r}{F_0}. \tag{12}$$

Here, F_0 is the original fitness and F_r is the fitness when the regulatory interaction is removed. For all analysis R was set to 500.

J. Divergence analysis

The divergence analysis aims to describe how diverged two regulatory networks are. We use this metric to determine if DSD is present by comparing network divergence along ancestral lineages where the fitness phenotype remained constant. There are multiple approaches we considered as divergence metric:

- by graph edit distance, which describes the number of 'edits' (structural mutations) needed to transform one graph into another.
- through an adjacency matrix, which describes the network as an interaction matrix in which rows describe a gene type, columns the TFBS type and entries in the matrix describe the number of interactions (number of TFBSs). By comparing adjacency matrices we can find both quantitative and qualitative differences, but can not distinguish between them.
- through a Boolean adjacency matrix, which is similar to the adjacency matrix but matrix entries only describe if an interaction is present, not the number of interactions. Therefore, comparing Boolean matrices only describes qualitative differences between to networks.

The Boolean matrix approach is most conservative and only describes qualitative differences, whereas the other methods also regard quantitative differences. In Appendix S3 we compare these different methods in more detail. Although quantitative differences are also regarded to be DSD, we chose to report the most conservative, qualitative approach.
The divergence between two networks is computed by first rewriting each network as a Boolean adjacency matrix. To do this we find for each unique gene type the set of TFBS types which activate/inhibit the respective gene type (we

separate activating and inhibiting interactions). A row in the boolean matrix denotes a gene type and the columns describe interactions of which TFBS type are present, see Fig S17 for examples. We compute the difference between two of these matrices, A and B, as follows

$$d(A,B) = \frac{\sum_i^N \sum_j^N |A_{i,j} - B_{i,j}|}{\sum_i^N \sum_j^N (A_{i,j} + B_{i,j})}. \tag{13}$$

Using this difference we can define the divergence between the GRNs of two individuals, C and D as:

$$\mathrm{div}(C,D) = \frac{1}{2}\Big[d(C_{\mathrm{act}}, D_{\mathrm{act}}) + d(C_{\mathrm{inh}}, D_{\mathrm{inh}})\Big], \tag{14}$$

where the GRNs of C and D are transformed into 2 Boolean matrices containing all activating ($C_{\mathrm{act}}, D_{\mathrm{act}}$) and inhibiting ($C_{\mathrm{inh}}, D_{\mathrm{inh}}$) interactions.

K. Conserved Non-coding Sequence analysis

For the counts of CNS presence profiles analysis we created a set with all the CNS ids which belong to a specific orthogroup. For each CNS in this set we assessed its presence in the given orthogroup of each species.
The phylogeny was made with the approach described in Man et al. (2020). In short, annotated proteins of interest were aligned using MAFFT and conserved regions identified (Katoh and Standley, 2013). For each conserved region, a BLAST search was performed (Altschul et al., 1990), and a hidden Markov model (HMMs) was generated with HMMER version 3.4 (Potter et al., 2018). Hits from both the BLAST and HMMER searches were taken together and a cutoff was determined using a known outgroup. Finally, using this set of hits, a maximum likelihood (ML) tree was generated using IQ-TREE 2 (Minh et al., 2020) with 100 bootstraps.

Bibliography

Aguirre, L., Hendelman, A., Hutton, S. F., McCandlish, D. M., and Lippman, Z. B. (2023). Idiosyncratic and dose-dependent epistasis drives variation in tomato fruit size. *Science*, 382(6668):315–320.

Ahmad, R. Introduction to stochastic differential equations, (1988).

Alberch, P. (1991). From genes to phenotype: dynamical systems and evolvability. *Genetica*, 84(1):5–11. doi: 10.1007/BF00123979.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, 215(3):403–410.

Amundson, K. R., Hendelman, A., Ciren, D., Yang, H., de Neve, A. E., Tal, S., Sulema, A., Jackson, D., Barlett, M. E., Lippman, Z. B., et al. (2025). A deep-time landscape of plant cis-regulatory sequence evolution. *bioRxiv*, pages 2025–09.

Bartlett, A., O'Malley, R. C., Huang, S.-s. C., Galli, M., Nery, J. R., Gallavotti, A., and Ecker, J. R. (2017). Mapping genome-wide transcription-factor binding sites using dap-seq. *Nature protocols*, 12(8):1659–1672.

Baumgart, L. A., Morales-Cruz, A., Greenblum, S. I., Wang, P., Zhang, Y., Yang, L., Chen, C., Dilworth, D. J., Garretson, A. C., Grosjean, N., et al. (2024). An atlas of conserved transcription factor binding sites reveals the cell type-resolved gene regulatory landscape of flowering plants. *bioRxiv*, pages 2024–10.

Brun-Usan, M., Rago, A., Thies, C., Uller, T., and Watson, R. A. (2021). Development and selective grain make plasticity'take the lead'in adaptive evolution. *BMC ecology and evolution*, 21(1):205.

Brun-Usan, M., Zimm, R., and Uller, T. (2022). Beyond genotype-phenotype maps: Toward a phenotype-centered perspective on evolution. *BioEssays*, 44(9):2100225.

Cano-Fernández, H., Brun-Usan, M., Tissot, T., and Salazar-Ciudad, I. (2024). A morphospace exploration using a general model of development reveals a basic set of morphologies for early animal development and evolution. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*.

Ciliberti, S., Martin, O. C., and Wagner, A. (2007). Innovation and robustness in complex regulatory gene networks. *Proceedings of the National Academy of Sciences*, 104(34):13591–13596.

Ciren, D., Zebell, S., and Lippman, Z. B. (2024). Extreme restructuring of cis-regulatory regions controlling a deeply conserved plant stem cell regulator. *PLoS Genetics*, 20(3):e1011174.

Cotterell, J. and Sharpe, J. (2010). An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients. *Molecular systems biology*, 6(1):425.

Cotterell, J. and Sharpe, J. (2013). Mechanistic explanations for restricted evolutionary paths that emerge from gene regulatory networks. *PloS one*, 8(4):e61178.

Crombach, A. and Hogeweg, P. Evolution of evolvability in gene regulatory networks, (2008). ISSN 1553734X.

Crombach, A., Wotton, K. R., Jiménez-Guri, E., and Jaeger, J. (2016). Gap gene regulatory dynamics evolve along a genotype network. *Molecular biology and evolution*, 33(5):1293–1307.

DiFrisco, J. (2023). Toward a theory of homology: development and the de-coupling of morphological and molecular evolution. *The British Journal for the Philosophy of Science*, 74(3):771–810.

Ditta, G., Pinyopich, A., Robles, P., Pelaz, S., and Yanofsky, M. F. (2004). The sep4 gene of arabidopsis thaliana functions in floral organ and meristem identity. *Current biology*, 14(21):1935–1940.

Draghi, J. and Wagner, G. P. (2009). The evolutionary dynamics of evolvability in a gene network model. *Journal of evolutionary biology*, 22(3):599–611.

Dreni, L. and Ferrándiz, C. (2022). Tracing the evolution of the sepallata subfamily across angiosperms associated with neo-and sub-functionalization for reproductive and agronomically relevant traits. *Plants*, 11(21):2934.

Du, Q., Faber, V., and Gunzburger, M. (1999). Centroidal voronoi tessellations: Applications and algorithms. *SIAM review*, 41(4): 637–676.

Duque, T. and Sinha, S. (2015). What does it take to evolve an enhancer? a simulation-based study of factors influencing the emergence of combinatorial regulation. *Genome biology and evolution*, 7(6):1415–1431.

Emms, D. M. and Kelly, S. (2019). Orthofinder: phylogenetic orthology inference for comparative genomics. *Genome biology*, 20: 1–14.

Fouracre, J. P. and Harrison, C. J. (2022). How was apical growth regulated in the ancestral land plant? insights from the development of non-seed plants. *Plant Physiology*, 190(1):100–112.

Frank, M. H., Edwards, M. B., Schultz, E. R., McKain, M. R., Fei, Z., Sørensen, I., Rose, J. K., and Scanlon, M. J. (2015). Dissecting the molecular signatures of apical cell-type shoot meristems from two ancient land plant lineages. *New Phytologist*, 207(3):893–904.

Freeling, M. and Subramaniam, S. (2009). Conserved noncoding sequences (cnss) in higher plants. *Current opinion in plant biology*, 12(2):126–132.

Gillespie, D. T. (2007). Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.*, 58(1):35–55.

Hagolani, P. F., Zimm, R., Vroomans, R., and Salazar-Ciudad, I. (2021). On the evolution and development of morphological complexity: A view from gene regulatory networks. *PLoS computational biology*, 17(2):e1008570.

Harrison, C. J. (2015). Shooting through time: new insights from transcriptomic data. *Trends in Plant Science*, 20(8):468–470.

Hendelman, A., Zebell, S., Rodriguez-Leal, D., Dukler, N., Robitaille, G., Wu, X., Kostyun, J., Tal, L., Wang, P., Bartlett, M. E., et al. (2021). Conserved pleiotropy of an ancient plant homeobox gene uncovered by cis-regulatory dissection. *Cell*, 184(7): 1724–1739.

Hobe, M., Müller, R., Grünewald, M., Brand, U., and Simon, R. (2003). Loss of cle40, a protein functionally equivalent to the stem cell restricting signal clv3, enhances root waving in arabidopsis. *Development genes and evolution*, 213:371–381.

Hogeweg, P. (2000). Evolving mechanisms of morphogenesis: on the interplay between differential adhesion and cell differentiation. *Journal of theoretical biology*, 203(4):317–333.

Hogeweg, P. (2000). Shapes in the shadow: evolutionary dynamics of morphogenesis. *Artificial Life*, 6(1):85–101.

Hogeweg, P. (2012). Toward a theory of multilevel evolution: long-term information integration shapes the mutational landscape and enhances evolvability. *Evolutionary systems biology*, pages 195–224.

Huynen, M. A. (1996). Exploring phenotype space through neutral evolution. *Journal of molecular evolution*, 43:165–169.

Huynen, M. A., Stadler, P. F., and Fontana, W. (1996). Smoothness within ruggedness: the role of neutrality in adaptation. *Proceedings of the National Academy of Sciences*, 93(1):397–401.

Jaeger, J. (2018). Shift happens: the developmental and evolutionary dynamics of the gap gene system. *Current Opinion in Systems Biology*, 11:65–73.

Jill Harrison, C. (2017). Development and genetics in the evolution of land plant body plans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1713):20150490.

Jiménez, A., Cotterell, J., Munteanu, A., and Sharpe, J. (2015). Dynamics of gene circuits shapes evolvability. *Proceedings of the National Academy of Sciences*, 112(7):2103–2108.

Johnson, N. A. and Porter, A. H. (2007). Evolution of branched regulatory genetic pathways: directional selection on pleiotropic loci accelerates developmental system drift. *Genetica*, 129:57–70.

Joly-Lopez, Z., Platts, A. E., Gulko, B., Choi, J. Y., Groen, S. C., Zhong, X., Siepel, A., and Purugganan, M. D. (2020). An inferred fitness consequence map of the rice genome. *Nature plants*, 6(2):119–130.

Katoh, K. and Standley, D. M. (2013). Mafft multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, 30(4):772–780.

Kimura, M. (1979). The neutral theory of molecular evolution. *Scientific American*, 241(5):98–129.

Krol, A. J., Roellig, D., Dequéant, M.-L., Tassy, O., Glynn, E., Hattem, G., Mushegian, A., Oates, A. C., and Pourquié, O. (2011). Evolutionary plasticity of segmentation clock networks. *Development*, 138(13):2783–2792.

Lanctot, A., Hendelman, A., Udilovich, P., Robitaille, G. M., and Lippman, Z. B. (2025). Antagonizing cis-regulatory elements of a conserved flowering gene mediate developmental robustness. *Proceedings of the National Academy of Sciences*, 122(8): e2421990122.

Lipman, D. J. and Wilbur, W. J. (1991). Modelling neutral and selective evolution of protein folding. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 245(1312):7–11.

Loman, T. E. and Locke, J. C. (2023). The σb alternative sigma factor circuit modulates noise to generate different types of pulsing dynamics. *PLoS Computational Biology*, 19(8):e1011265.

Lu, Z., Marand, A. P., Ricci, W. A., Ethridge, C. L., Zhang, X., and Schmitz, R. J. (2019). The prevalence, evolution and chromatin signatures of plant regulatory elements. *Nature Plants*, 5(12):1250–1259.

Man, J., Gallagher, J. P., and Bartlett, M. (2020). Structural evolution drives diversification of the large lrr-rlk gene family. *New Phytologist*, 226(5):1492–1505.

McColgan, Á. and DiFrisco, J. (2024). Understanding developmental system drift. *Development*, 151(20).

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., and Lanfear, R. (2020). Iq-tree 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular biology and evolution*, 37(5): 1530–1534.

Ohno, S. *Evolution by gene duplication*. Springer Science & Business Media, (1970).

Orr, H. A. (2000). Adaptation and the cost of complexity. *Evolution*, 54(1):13–20.

Orr, H. A. (2009). Fitness and its role in evolutionary genetics. *Nature Reviews Genetics*, 10(8):531–539.

Pavlicev, M. and Wagner, G. P. (2012). A model of developmental evolution: selection, pleiotropy and compensation. *Trends in Ecology & Evolution*, 27(6):316–322.

Payne, J. L. and Wagner, A. (2014). Latent phenotypes pervade gene regulatory circuits. *BMC systems biology*, 8(1):64.

Payne, J. L., Moore, J. H., and Wagner, A. (2014). Robustness, evolvability, and the logic of genetic regulation. *Artificial life*, 20 (1):111–126.

Pelaz, S., Ditta, G. S., Baumann, E., Wisman, E., and Yanofsky, M. F. (2000). B and c floral organ identity functions require sepallata mads-box genes. *Nature*, 405(6783):200–203.

Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., and Finn, R. D. (2018). Hmmer web server: 2018 update. *Nucleic acids research*, 46(W1):W200–W204.

Ricci, W. A., Lu, Z., Ji, L., Marand, A. P., Ethridge, C. L., Murphy, N. G., Noshay, J. M., Galli, M., Mejía-Guerra, M. K., Colomé-Tatché, M., et al. (2019). Widespread long-range cis-regulatory elements in the maize genome. *Nature plants*, 5 (12):1237–1249.

Ruelens, P., De Maagd, R. A., Proost, S., Theißen, G., Geuten, K., and Kaufmann, K. (2013). Flowering locus c in monocots and the tandem origin of angiosperm-specific mads-box genes. *Nature communications*, 4(1):2280.

Schaerli, Y., Jiménez, A., Duarte, J. M., Mihajlovic, L., Renggli, J., Isalan, M., Sharpe, J., and Wagner, A. (2018). Synthetic circuits reveal how mechanisms of gene regulatory networks constrain evolution. *Molecular Systems Biology*, 14(9):e8102.

Schiffman, J. S. and Ralph, P. L. (2022). System drift and speciation. *Evolution*, 76(2):236–251.

Schlegel, J., Denay, G., Wink, R., Pinto, K. G., Stahl, Y., Schmid, J., Blümke, P., and Simon, R. G. (2021). Control of arabidopsis shoot stem cell homeostasis by two antagonistic cle peptide signalling pathways. *Elife*, 10:e70934.

Schuster, C., Gabel, A., Drost, H.-G., Grosse, I., Leyser, O., and Meyerowitz, E. M. (2026). Evolutionary transcriptomics unveils rapid changes of gene expression patterns in flowering plants. *Cell*.

Schuster, P., Fontana, W., Stadler, P. F., and Hofacker, I. L. (1994). From sequences to shapes and back: a case study in rna secondary structures. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 255(1344):279–284.

Somssich, M., Je, B. I., Simon, R., and Jackson, D. (2016). Clavata-wuschel signaling in the shoot meristem. *Development (Cambridge)*, 143:3238–3248. doi: $10.1242/\mathrm{dev}.133645$.

Stone, J. R. and Wray, G. A. (2001). Rapid evolution of cis-regulatory sequences via local point mutations. *Molecular biology and evolution*, 18(9):1764–1770.

Szövényi, P., Waller, M., and Kirbis, A. Chapter one - evolution of the plant body plan. In Grossniklaus, U., editor, *Plant Development and Evolution*, volume 131 of *Current Topics in Developmental Biology*, pages 1 – 34. Academic Press, (2019). doi: $\mathrm{https://doi.org/10.1016/bs.ctdb.2018.11.005}$.

ten Tusscher, K. H. and Hogeweg, P. (2011). Evolution of networks for body plan patterning; interplay of modularity, robustness and evolvability. *PLoS Computational Biology*, 7. doi: $10.1371/\mathrm{journal.pcbi.1002208}$.

True, J. R. and Haag, E. S. (2001). Developmental system drift and flexibility in evolutionary trajectories. *Evolution & development*, 3(2):109–119.

Tsong, A. E., Tuch, B. B., Li, H., and Johnson, A. D. (2006). Evolution of alternative transcriptional circuits with identical logic.

Nature, 443(7110):415–420.

Van Nimwegen, E., Crutchfield, J. P., and Huynen, M. (1999). Neutral evolution of mutational robustness. *Proceedings of the National Academy of Sciences*, 96(17):9716–9720.

Verster, A. J., Ramani, A. K., McKay, S. J., and Fraser, A. G. (2014). Comparative rnai screens in c. elegans and c. briggsae reveal the impact of developmental system drift on gene function. *PLoS genetics*, 10(2):e1004077.

von der Dunk, S. H., Snel, B., and Hogeweg, P. (2022). Evolution of complex regulation for cell-cycle control. *Genome biology and evolution*, 14(5):evac056.

Vroomans, R. M., Hogeweg, P., and Ten Tusscher, K. H. (2016). In silico evo-devo: reconstructing stages in the evolution of animal segmentation. *EvoDevo*, 7:1–20.

Vroomans, R. M., Hogeweg, P., and Ten Tusscher, K. H. (2018). Around the clock: gradient shape and noise impact the evolution of oscillatory segmentation dynamics. *EvoDevo*, 9:1–20.

Wagner, A. (2008). Neutralism and selectionism: a network-based reconciliation. *Nature Reviews Genetics*, 9(12):965–974.

Wagner, A. (2008). Robustness and evolvability: a paradox resolved. *Proceedings of the Royal Society B: Biological Sciences*, 275(1630):91–100.

Wagner, A. (2014). Mutational robustness accelerates the origin of novel rna phenotypes through phenotypic plasticity. *Biophysical journal*, 106(4):955–965.

Weiss, K. M. and Fullerton, S. M. (2000). Phenogenetic drift and the evolution of genotype–phenotype relationships. *Theoretical population biology*, 57(3):187–195.

Whitewoods, C. D., Cammarata, J., Nemec Venza, Z., Sang, S., Crook, A. D., Aoyama, T., Wang, X. Y., Waller, M., Kamisugi, Y., Cuming, A. C., Szövényi, P., Nimchuk, Z. L., Roeder, A. H., Scanlon, M. J., and Harrison, C. J. (2018). Clavata was a genetic novelty for the morphological innovation of 3d growth in land plants. *Current Biology*, 28(15):2365–2376.e5. doi: https://doi.org/10.1016/j.cub.2018.05.068.

Wotton, K. R., Jiménez-Guri, E., Crombach, A., Janssens, H., Alcaine-Colet, A., Lemke, S., Schmidt-Ott, U., and Jaeger, J. (2015). Quantitative system drift compensates for altered maternal inputs to the gap gene network of the scuttle fly megaselia abdita. *Elife*, 4:e04785.

Zahn, L. M., Kong, H., Leebens-Mack, J. H., Kim, S., Soltis, P. S., Landherr, L. L., Soltis, D. E., Depamphilis, C. W., and Ma, H. (2005). The evolution of the sepallata subfamily of mads-box genes: a preangiosperm origin with multiple duplications throughout angiosperm history. *Genetics*, 169(4):2209–2223.

Zheng, J., Payne, J. L., and Wagner, A. (2019). Cryptic genetic variation accelerates evolution by opening access to diverse adaptive peaks. *Science*, 365(6451):347–353.

Zheng, J., Guo, N., and Wagner, A. (2020). Selection enhances protein evolvability by increasing mutational robustness and foldability. *Science*, 370(6521):eabb5962.

Supplementary Information

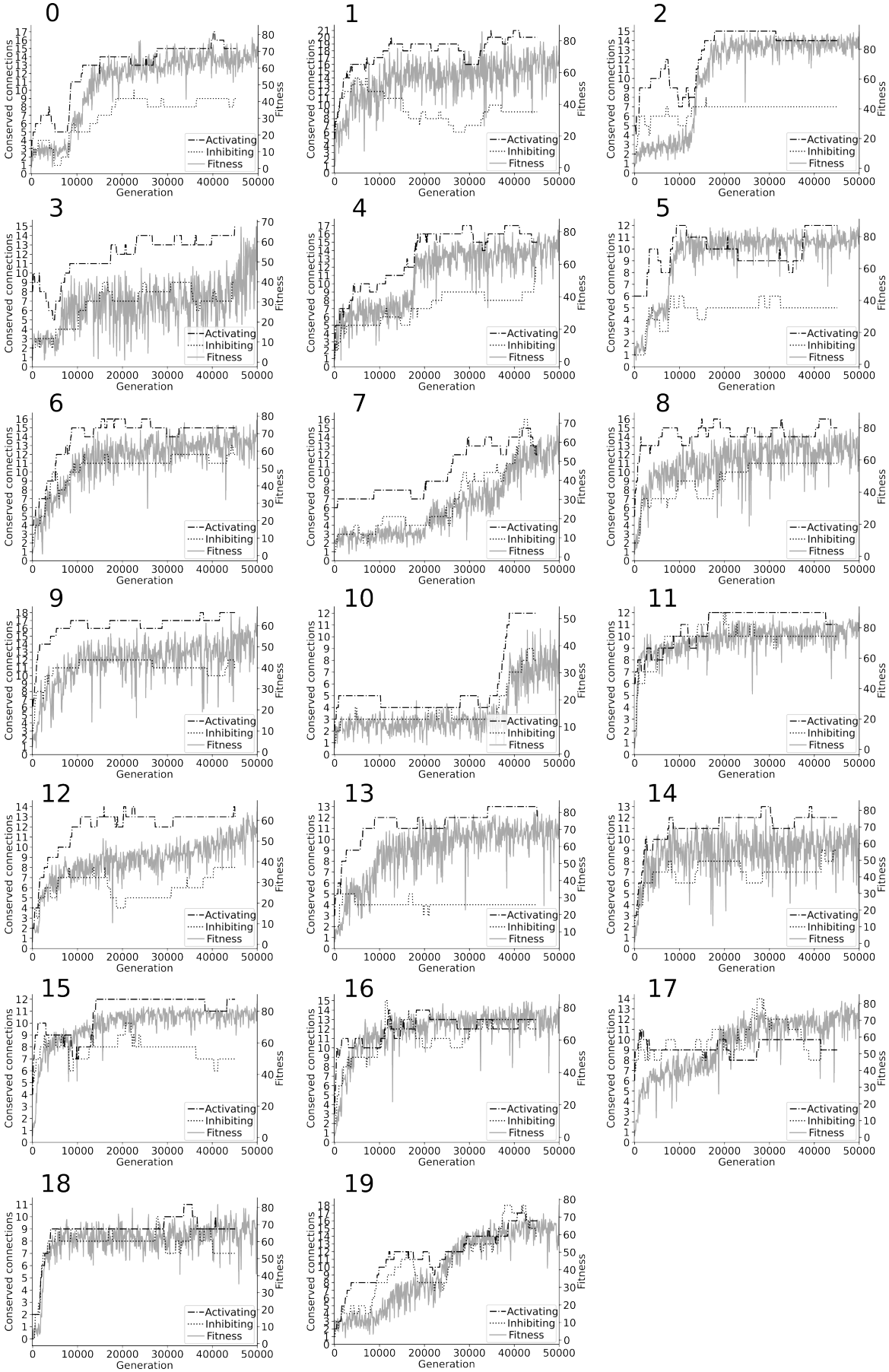

Figure S1. Fitness and conservation of interactions over evolutionary time. An interaction (either activating or inhibiting) is considered conserved if it persists for more than 5000 generations in the GRNs of the ancestral lineage.

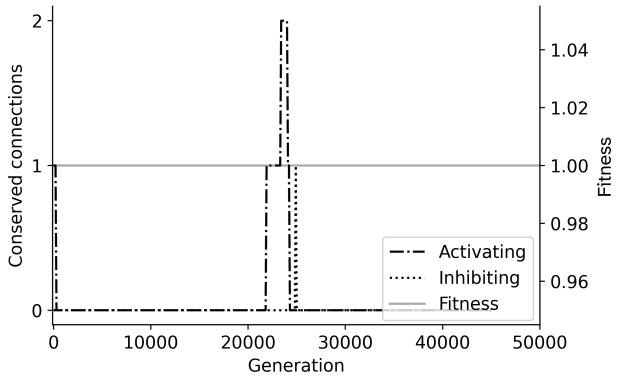

Figure S2. A simulation with random selection does not yield conserved interactions. In a simulation in which individuals are chosen for reproduction at random without selecting for a pattern, no clear conservation was observed.

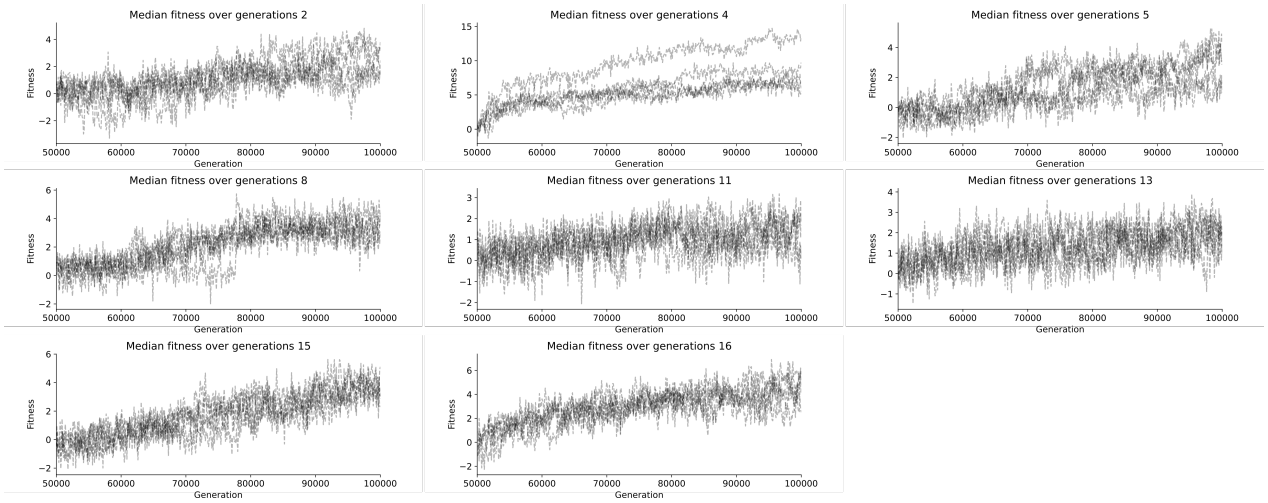

Figure S3. Median fitness increase in simulations after cloning the population. The median fitnesses of all 5 cloned populations in simulations $2, 4, 5, 8, 11, 13, 15, 16$. Populations 4 and 16 were excluded for analysis of stabilising evolution due to their fitness increase.

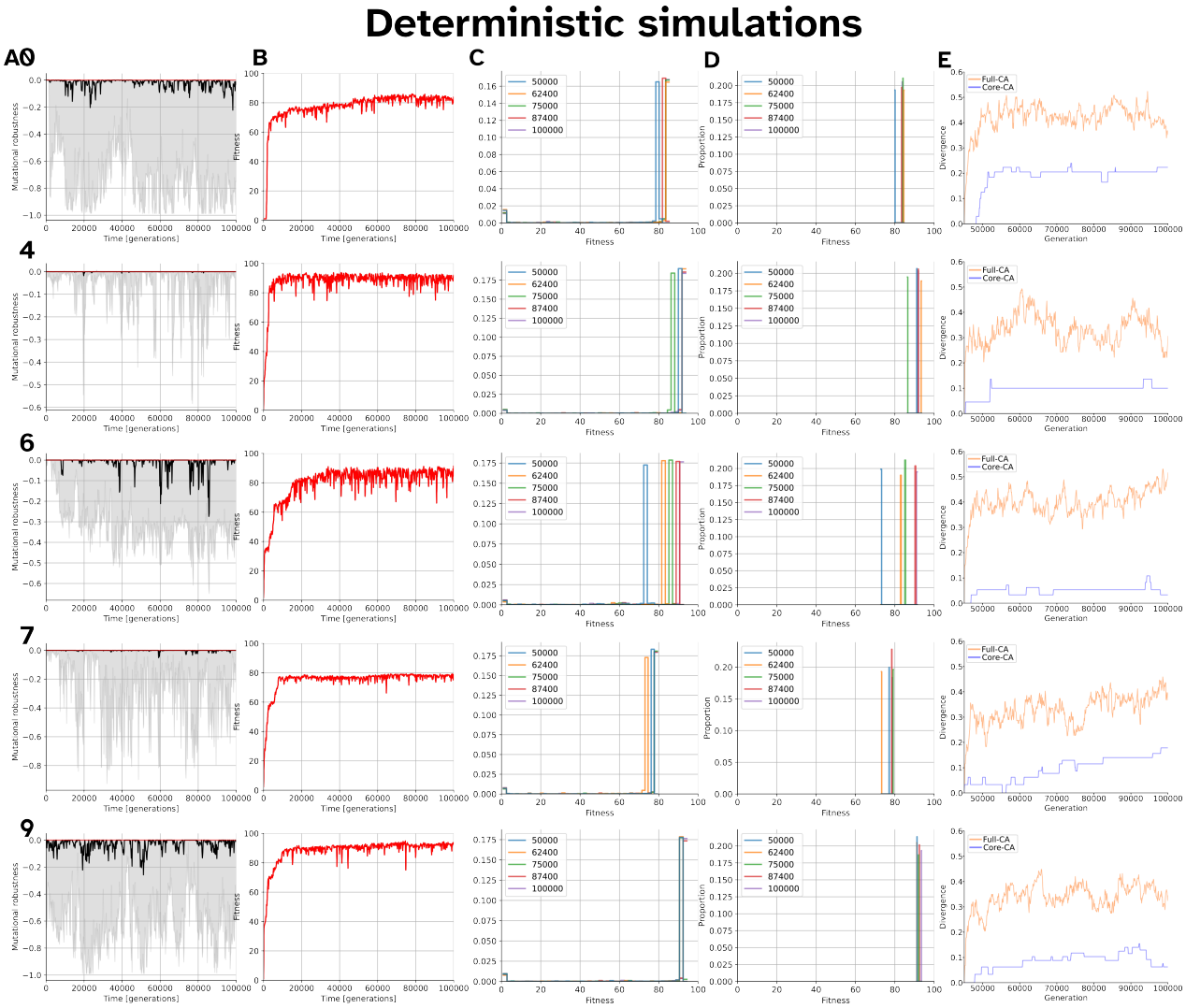

Figure S4. Stabilising selection of deterministic simulations. All plots show individuals along an ancestral lineage of deterministic simulations with their simulation ID indicated at the left of each row. **A.** Mutational robustness (see eq. S7 over generational time. Red considers all offspring; black considers only the mutated offspring, IQR of the mutated offspring is shown in grey. **B.** Fitness along ancestral lineage. **C.** Fitness distribution of 10,000 offspring. **D.** Fitness distribution of non-mutated offspring in E. **E.** GRN divergence along ancestral lineage compared to the individual at generation 45,000.

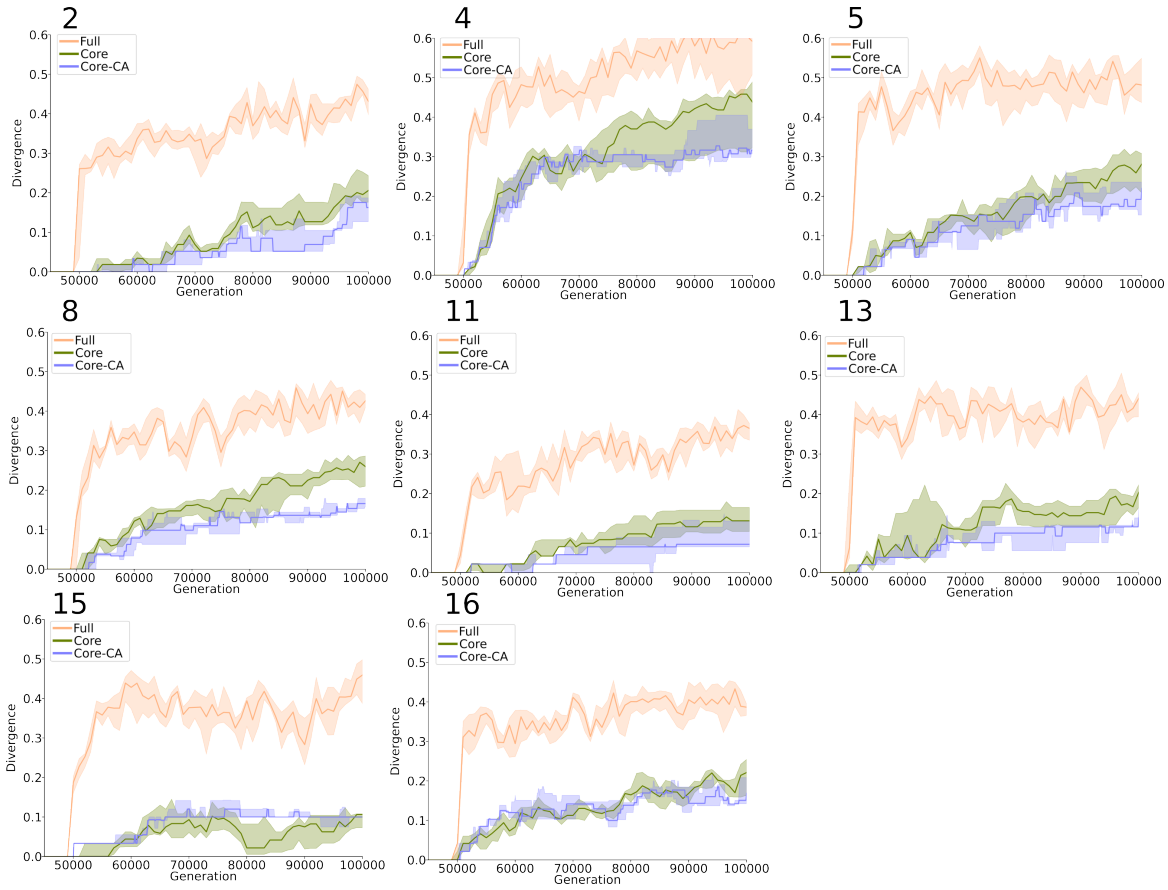

Figure S5. Network divergence in the cloned lineages of 8 populations.. Network divergence of simulations $2, 4, 5, 8, 11, 13, 15, 16$, which were cloned 5 times and evolved independently until generation 100000. Full: pairwise GRN divergence between all pairs of cloned lineages of the population. GRNs are taken from the ancestral lineage of the fittest individual in the final generation. Core: comparison of only the conserved interactions of the GRN (interactions > 5000). Core-CA: comparison of an ancestral GRN in a particular generation with the common ancestor (CA) of all lineages.

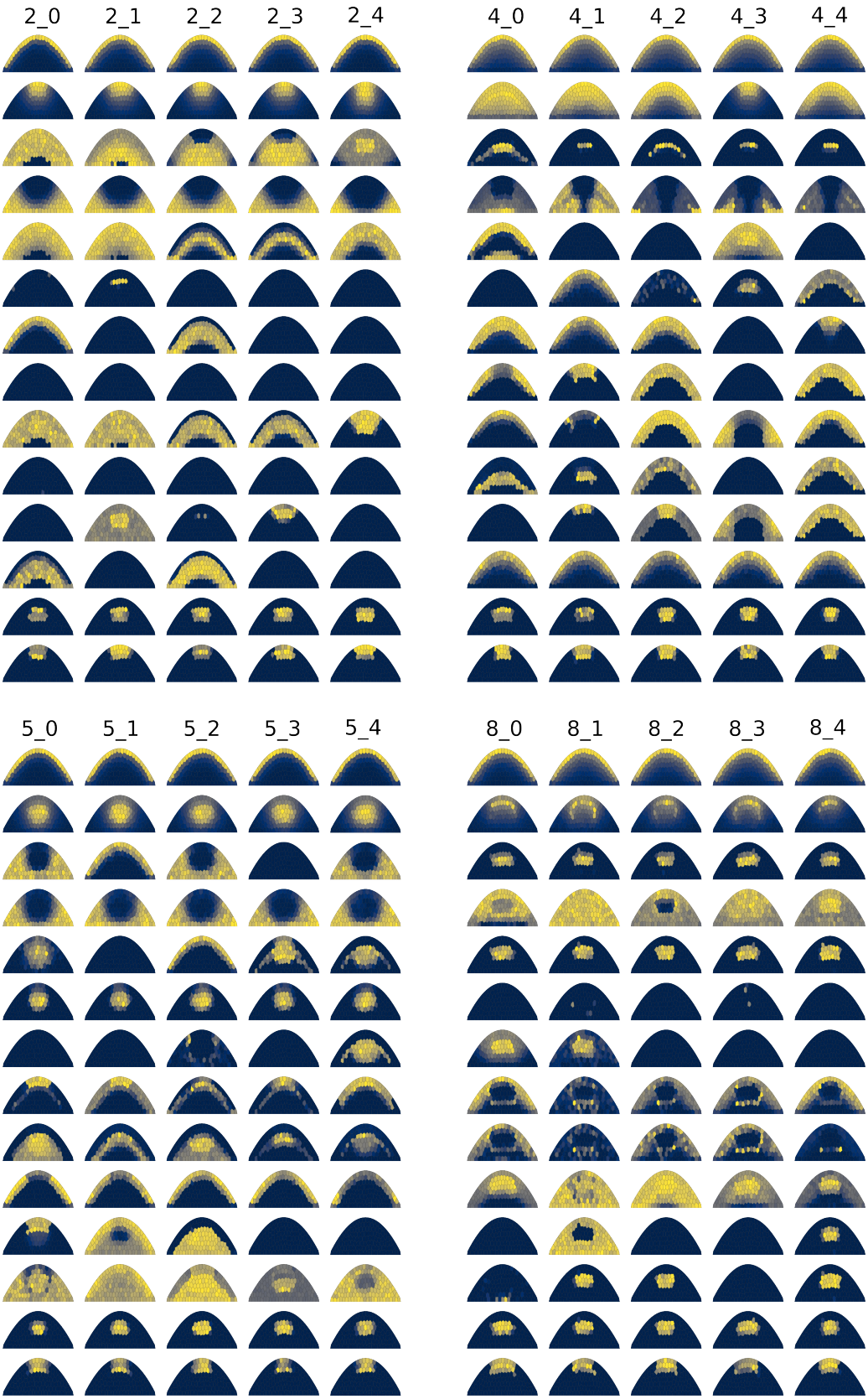

Figure S6. Expression patterns in cloned lineages. Expression patterns of all 14 genes (0-13, top to bottom) after development for the highest fitness individuals in the final cloned populations (generation 100000). In the label, the first number indicates the original simulation from which the clones were generated (generation $0-50000$), the second number indicates the clone number (generation $50000-100000$).

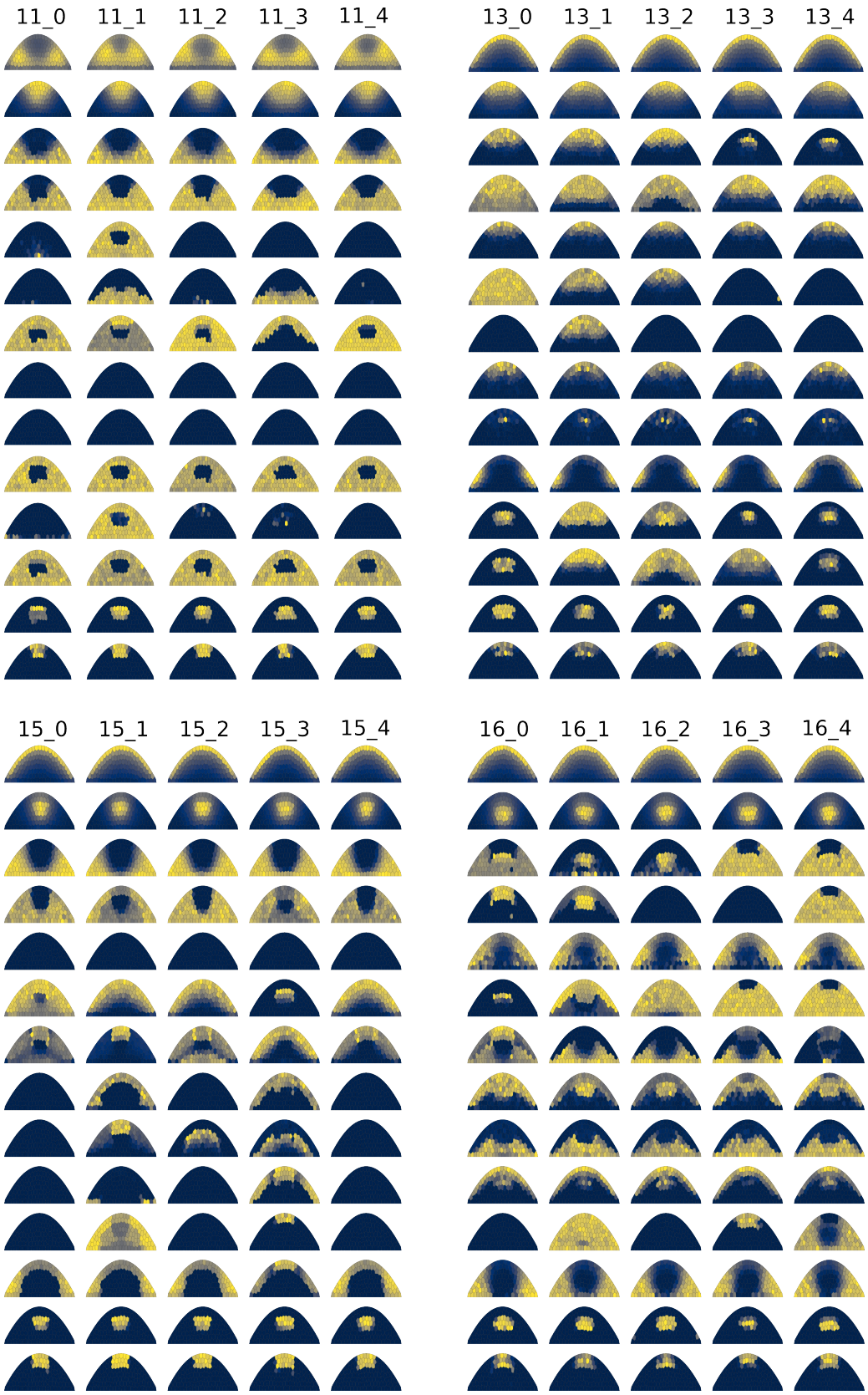

Figure S7. Expression patterns, cont'd.. Expression patterns of all 14 genes (0-13, top to bottom) after development for the highest fitness individuals in the final cloned populations (generation 100000). In the label, the first number indicates the original simulation from which the clones were generated (generation $0-50000$), the second number indicates the clone number (generation $50000-100000$).

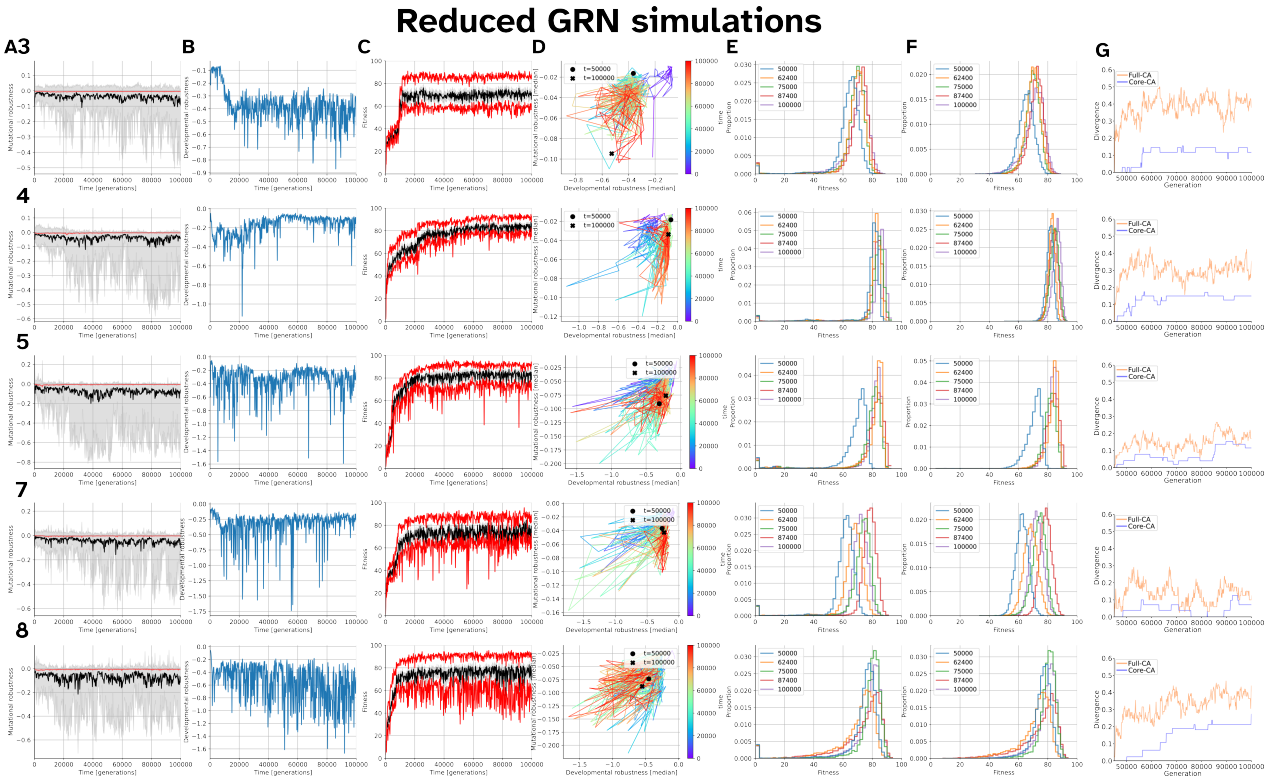

Figure S8. Stabilising selection of stochastic simulations with a reduced GRN. All plots show individuals along an ancestral lineage of deterministic simulations with their simulation ID indicated at the left of each row. These simulations contained 2 diffusing genes, 2 cell-cell genes, 2 cell autonomous genes, and 2 fitness genes. **A.** Mutational robustness (see eq. S7 over generational time. Red considers all offspring; black considers only the mutated offspring, IQR of the mutated offspring is shown in grey. **B.** Developmental robustness (see eq. S8) over generational time. **C.** Median fitness of clonal repeats (black) with IQR (grey) and upper 95% (red).
textbfD. Mutational robustness plotted against developmental robustness for several timepoints along the line of descent (indicated by color). For mutational robustness we only considered mutated offspring. **E.** Fitness distribution of 10,000 offspring. **F.** Fitness distribution of only the non-mutated offspring in E. **G.** GRN divergence along ancestral lineage compared to the individual at generation 45,000.

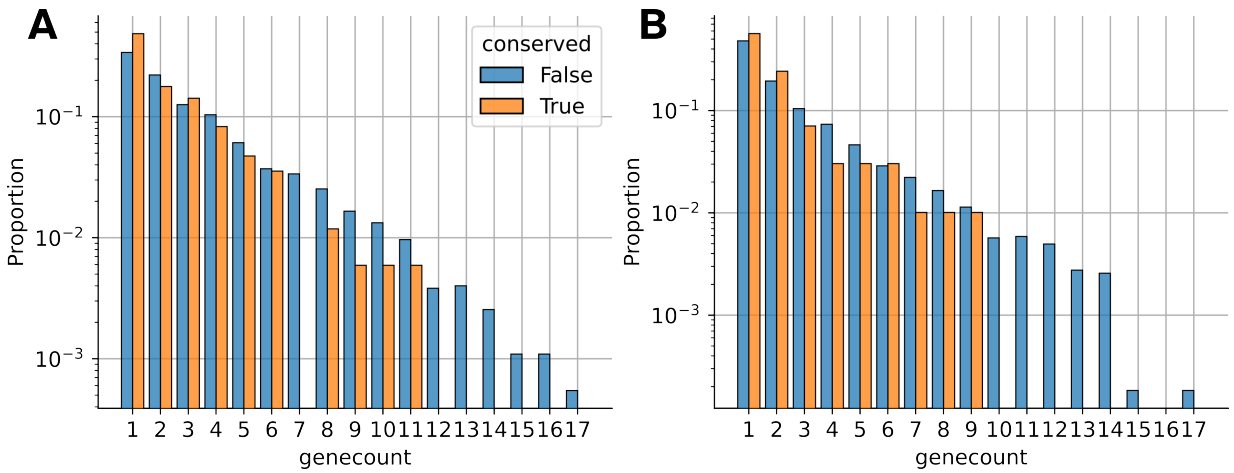

Figure S9. Rewiring of conserved interactions. Proportion of gained (**A**) or lost (**B**) interactions by gene copy (so, how many copies of a gene were present when the connection was gained/lost). No difference in proportion per gene count can be observed in rewiring events between not conserved and conserved interactions.

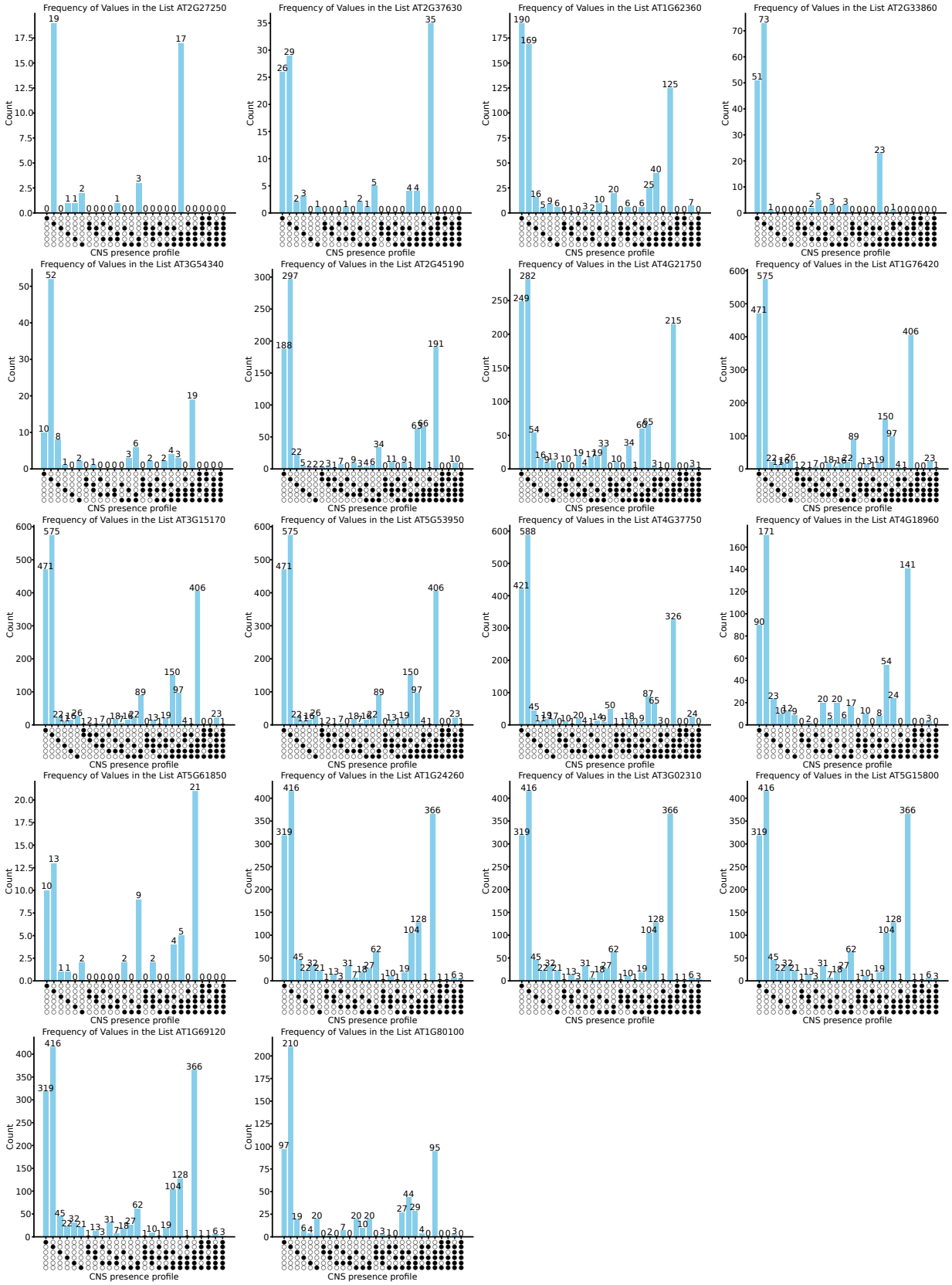

Figure S10. CNS presence profiles for different orthogroups in six land plant species. Presence of a CNS is indicated as a filled circle, species are from top to bottom: *Brachypodium distachyon*, *Medicago truncatula*, *Eutrema salsugineum*, *Capsella rubella*, *Arabidopsis lyrata*, *Arabidopsis thaliana*. The gene ID of the *Arabidopsis thaliana* gene in the respective orthogroup is shown on top of the plots. Some genes are within the same orthogroup resulting in identical patterns.

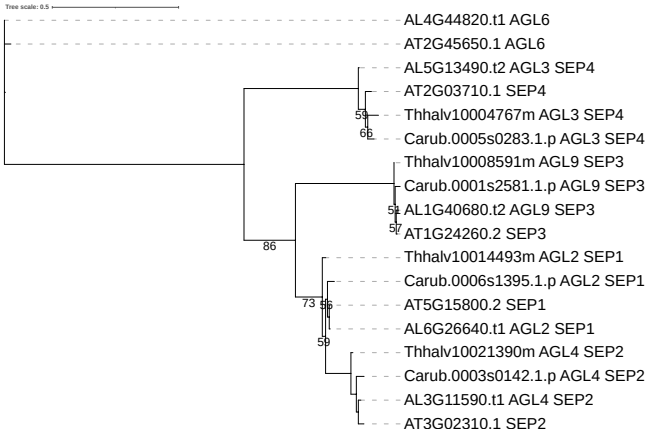

Figure S11. Full gene tree of protein sequences derived from SEP primary transcripts. Figure matches the gene tree in Figure 5C.

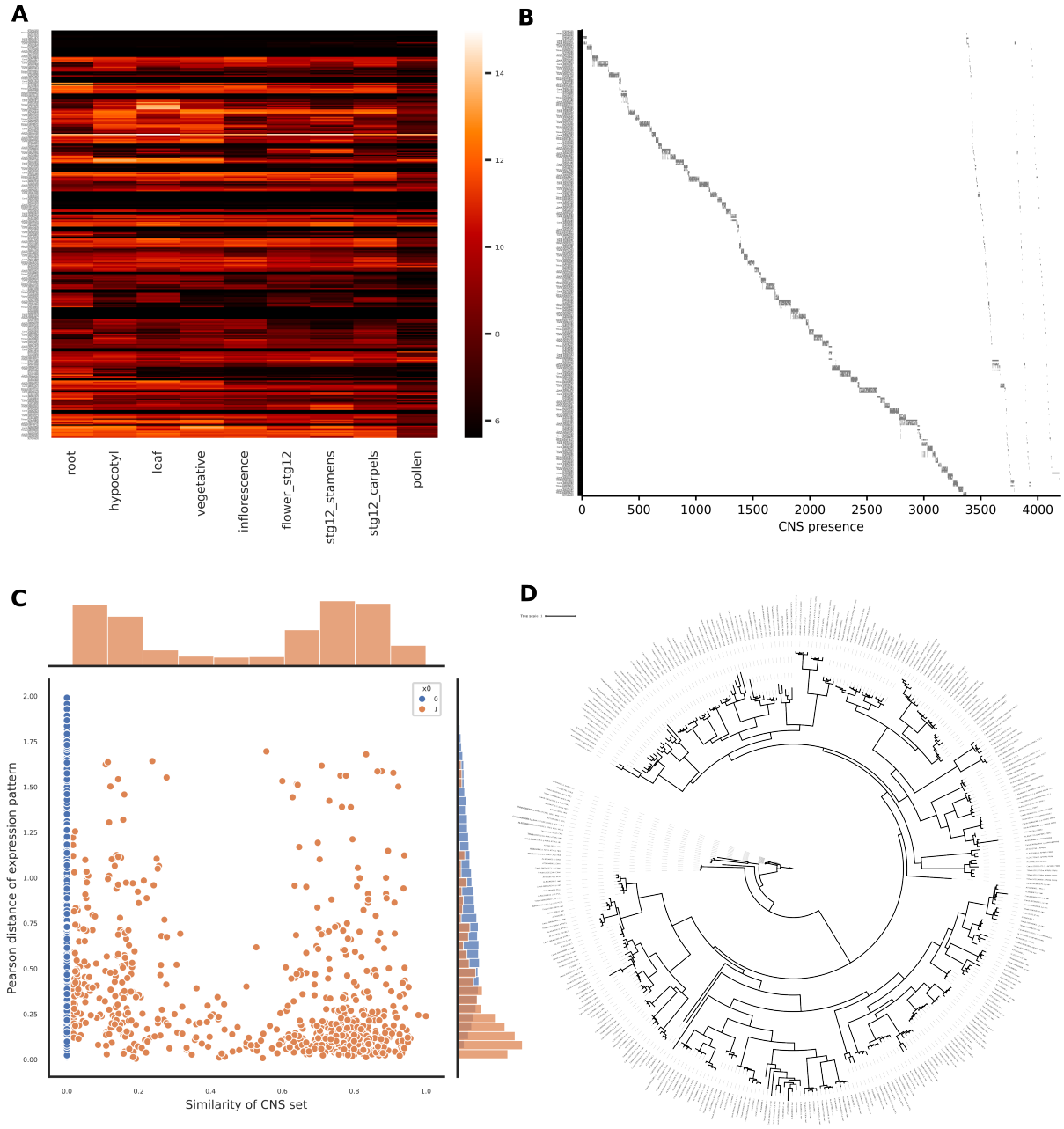

Figure S12. CNS profile and expression of B-ARR genes in four Brassicaceae species. A. Expression profile of B-ARR genes in different organs. The expression data shown are variance stabilized mRNA counts. **B.** Heatmap of CNS presence for each of the B-ARR genes. Presence of a CNS is shown in grey. **C.** Pearson distance of expression against the similarity of the CNS set for each gene pair in D. Histograms show the distribution of the pearson distances (right) and similarity of CNS set (top). The data of disjoint CNS sets is shown in blue. **D.** Gene tree of primary transcript protein sequences of B-ARR genes of the four Brassicaceae species: *Eutrema salsugineum; Capsella rubella; Arabidopsis lyrata; Arabidopsis thaliana*.

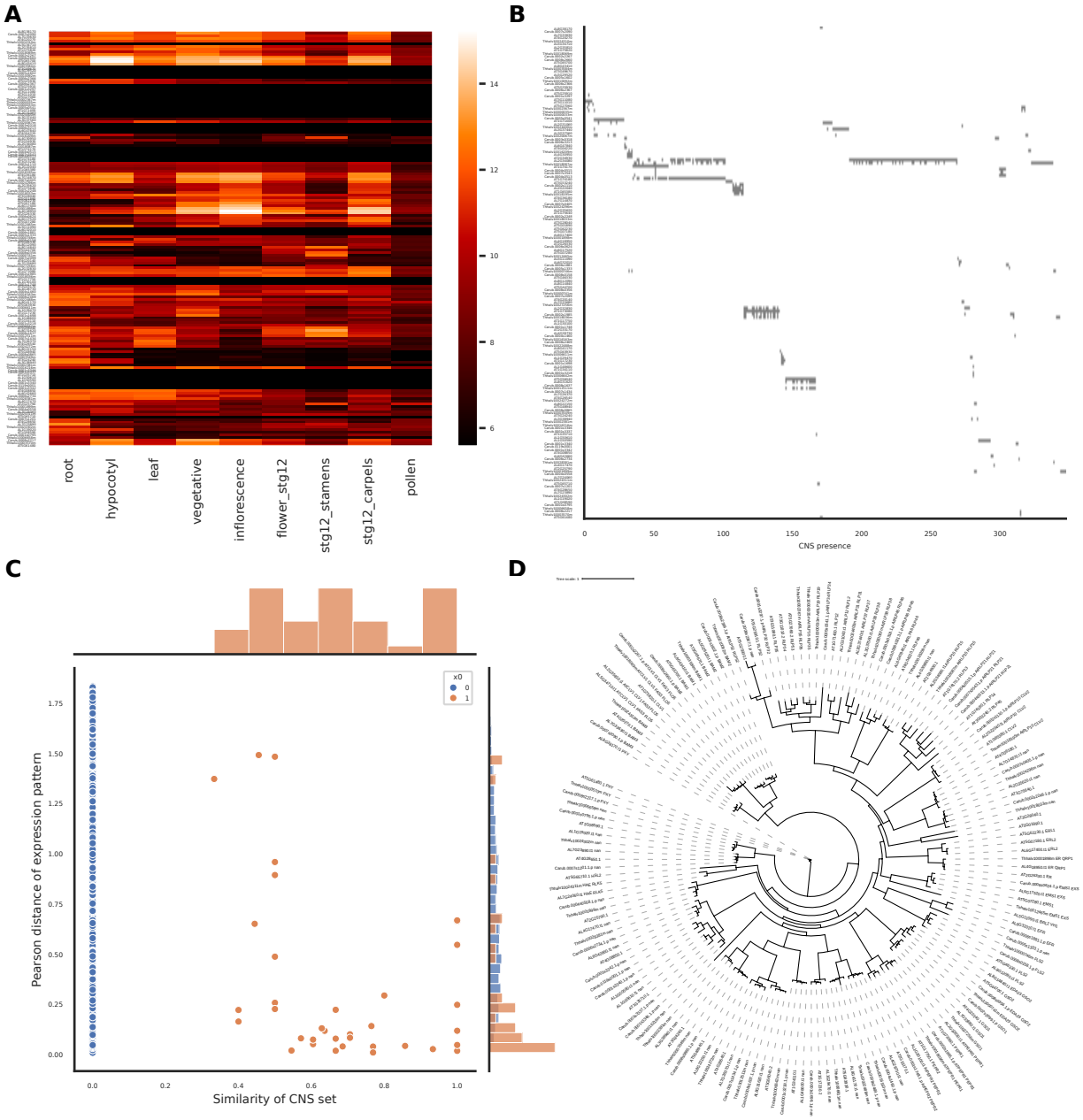

Figure S13. CNS profile and expression of CLE genes in four Brassicaceae species. A. Expression profile of CLE genes in different organs. The expression data shown are variance stabilized mRNA counts. **B.** Heatmap of CNS presence for each of the CLE genes. Presence of a CNS is shown in grey. **C.** Pearson distance of expression against the similarity of the CNS set for each gene pair in D. Histograms show the distribution of the pearson distances (right) and similarity of CNS set (top). The data of disjoint CNS sets is shown in blue. **D.** Gene tree of primary transcript protein sequences of CLE genes of the four Brassicaceae species: *Eutrema salsugineum; Capsella rubella; Arabidopsis lyrata; Arabidopsis thaliana*.

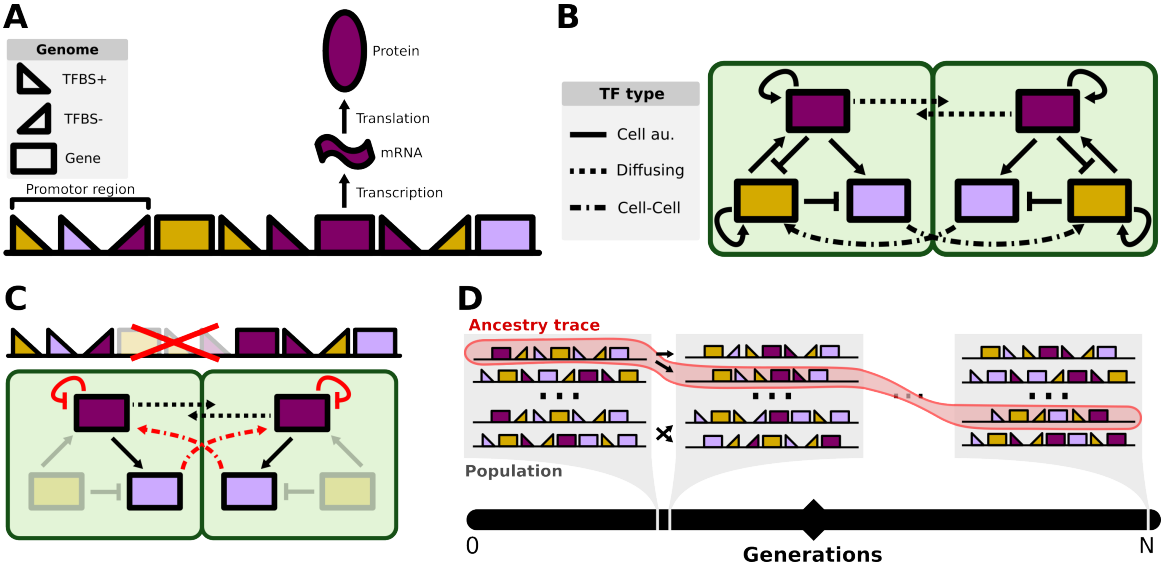

Figure S14. Schematic representation of the genome and the reaction network used for the computational evo-devo simulations. A. The genome, which consists of genes and TFBSs. The TFBSs in front of a gene make up the regulatory region of that gene and determine which TFs regulate its expression. A TFBS can be either enhancing (+) or inhibiting (−) the expression of its downstream gene. Transcription of a gene results in the production of a specific type of mRNA, which in turn is translated to produce a specific TF. **B.** The genome can be translated into a reaction network. TFs can have different properties, they either only act within their own cell, can diffuse to other cells or can affect transcription in directly neighbouring cells as a form of cell-cell communication. **C.** Example of how mutations to a genome influence the reaction networks within and between cells. **D.** Example of an ancestry trace of a lineage through generations shown in red. Arrows indicate a parent-offspring relationship.

Appendix S1 Drift vs adaptation

To observe DSD, we cloned our populations at the apparent fitness plateau at 50,000 generations. In the subsequent 50,000 generations, we observed a very slight fitness increase in most populations (Fig S3). To determine if this increase was due to specific adaptive mutations, we analyzed the probability that a mutation with a given fitness increase a would become fixed within the population versus the probability that a mutation providing no fitness gain would become fixed. For a mutation to become fixed the given lineage has to be selected for offspring in successive generations.

The probability of an individual to be picked for generating offspring is $\frac{f_q}{\sum_i^N f_i}$ for a single trial. An individual with a mutation resulting in fitness gain a will have probability $p_q = \frac{f_q+a}{(\sum_i^N f_i)+a}$ of being picked, again for a single trial. For $a << f_i$ we can assume

$$\sum_i^N (f_i) + a \approx \sum_i^N (f_i + a) \approx \sum_i^N f_i, \qquad \textbf{(eq. S1)}$$

which means that we can treat p_q as a constant instead of having to account for fitness gain of the total population. Our selection (roulette wheel selection) can be described as a Bernoulli process which has the following probability mass function

$$f(k,n,p) = Pr(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad \textbf{(eq. S2)}$$

describing the probability of exactly k successes in n trials with probability p success. For ease of notation we describe the probability of m individuals with the mutation to be present in generation g as $P_{m,g}$. Assuming only a single individual gains the mutation we can denote the probability vector of generation 0 as $P_{1,0} = 1$ and $P_{i,0} = 0 \forall i \neq 1$.

$$P_{m,g}(p_q) = \sum_i [P_{i,g-1} \cdot f(m,n,i \cdot p_q)]. \qquad \textbf{(eq. S3)}$$

This function can also be adapted to allow for fitness gain over time through

$$p_q(g) = \frac{f_q + ag}{(\sum_i^N f_i)} \qquad \textbf{(eq. S4)}$$

$$P_{m,g}(p_q(g)) = \sum_i [P_{i,g-1} \cdot f(m,n,i \cdot p_q(g))] \qquad \textbf{(eq. S5)}$$

Now we can describe the probability that a given mutation is still present in the population at generation g as $1 - P_{0,g}$. This system converges to $P_g \approx P_{g-1}$. We investigated P_g after 1000 generations, or when $\mathrm{sum}(P_g - P_{g-1})^2 \leq 0.0001$, after which we call $1 - P_{0,1000}$ the fixation probability. If for a mutation with a given fitness increase a the fixation probability is significantly higher than for a mutation without that fitness increase, we assume a selective regime. If a fitness gain a results in a negligible increase in fixation probability we assume a neutral regime.

Using the data from simulation 2 we estimated the fitness gain at each time point by fitting a line ($ax + b = y$) through the next 5 fitness values. As we only sampled the population every 100 generations this slope might be an underestimate of the actual fitness gain achieved. To get a best-case upper bound we therefore multiplied this slope by 100, which assumes that all the fitness gain within that window was gained in just 5 generations. At each generation we compute the log fold change of fixation probability

$$\log_2 \left[(1 - P_{0,g} \left(\frac{f_q + a}{(\sum_i^N f_i)} \right)) / (1 - P_{0,g} \left(\frac{f_q}{(\sum_i^N f_i)} \right)) \right], \qquad \textbf{(eq. S6)}$$

for the given slope and for $a = 0$ as neutral baseline. Figure S15A shows the fitness of the population at intervals of 10,000 generations, and the median fitness of the population over time. Using the median population fitness we gain a at each timepoint and compute the fold change in fixation probability (Fig S15B). For example, in generation 12,500 the fixation probability of the genome of a median fit individual is 3.97$\cdot 10^{-7}$, compared to the fixation probability of the genome of a median fit individual plus fitness gain 0.36. Whereas the fixation probability of the genome of a median fit individual at generation 50,000 is 0.144 versus the probability with fitness gain 0.140 (a was slightly negative at this point). Only in the beginning of the simulation up until around generation 15,000 there is a notable difference in fixation probability. The fitness gain after generation 50,000 is too shallow for selection to outweigh drift. This method, however, only detects direct adaptive fitness gains and will miss out on higher order effects such as increases in mutational or developmental robustness. During the shallow fitness slope such higher order effects are much more likely to play a dominant role.

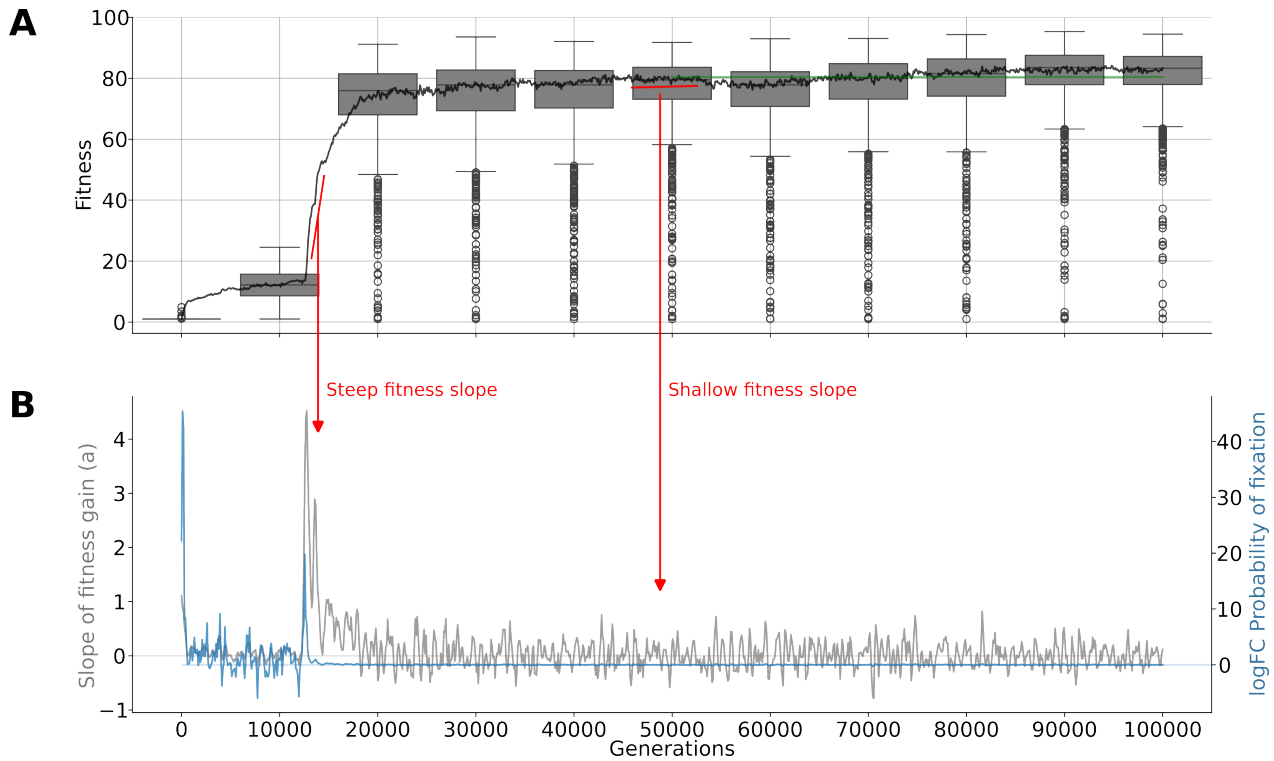

Figure S15. Neutral vs adaptive regimes of fitness gain. (A) Fitness of the whole population shown in box plots at intervals of 10,000 generations. Median fitness over generations shown per 100 generations as line. In green the median fitness of population at generation 50,000. **(B)** The slope of fitness gain per generation (grey) and the log2 fold change in fixation probability of a mutation providing fitness gain a versus without, see eq. S6.

Appendix S2 Evolution of Robustness during stabilising selection

We investigated the changes in either mutational robustness or developmental robustness over evolutionary time, especially during stabilising selection.

Mutational robustness (R_m) is defined as

$$R_m = \frac{\mathrm{med}(F_{\mathrm{offspring}}) - \mathrm{med}(F_{\mathrm{parent}})}{\mathrm{med}(F_{\mathrm{parent}})}, \qquad \textbf{(eq. S7)}$$

where $\mathrm{med}(F_{\mathrm{offspring}})$ denotes the median offspring fitness and $\mathrm{med}(F_{\mathrm{parent}})$ the median parent fitness. In other work, such as in (**?**), mutational robustness is described in terms of only the mutated offspring. However, the fluctuating genome sizes in our simulations result in different probabilities for offspring to be mutated. Only considering mutated offspring in Equation eq. S7 does not take into account this variability. On the other hand, considering all offspring might hide concomitant changes in developmental robustness.

Developmental robustness (R_d) is described as the variance proportional to the mean of clonal individuals. We transform this value such that a positive slope describes an increase in robustness.

$$R_d = \log\left[\frac{1}{1 + \frac{\mathrm{var}(F_{\mathrm{clones}})}{\mathrm{mean}(F_{\mathrm{clones}})}}\right] \qquad \textbf{(eq. S8)}$$

During stabilising selection (after generation 50,000), both the mutational and developmental robustness fluctuate, but neither does clearly increase or decrease overall (Fig S16). An exception to this is simulation 15_0 where a clearer gain in fitness can be observed at generation 90,000, which is also reflected in the distributions describing mutational and developmental robustness. This highlights the importance of neutral drift and how it interplays with adaptive steps, described before in **?**. In deterministic simulations divergence of the GRNs is still observed (Fig S4). Although this does not rule out a role for selection for developmental robustness in the divergence of the stochastic simulations, it does show that it is not a necessary driver of divergence. Reducing the size of the GRN made evolving developmental programs more challenging, showing the importance of available degrees of freedom in the genome (Fig S8). Nevertheless, similar neutral drift patterns can be seen in the mutational and developmental robustness.

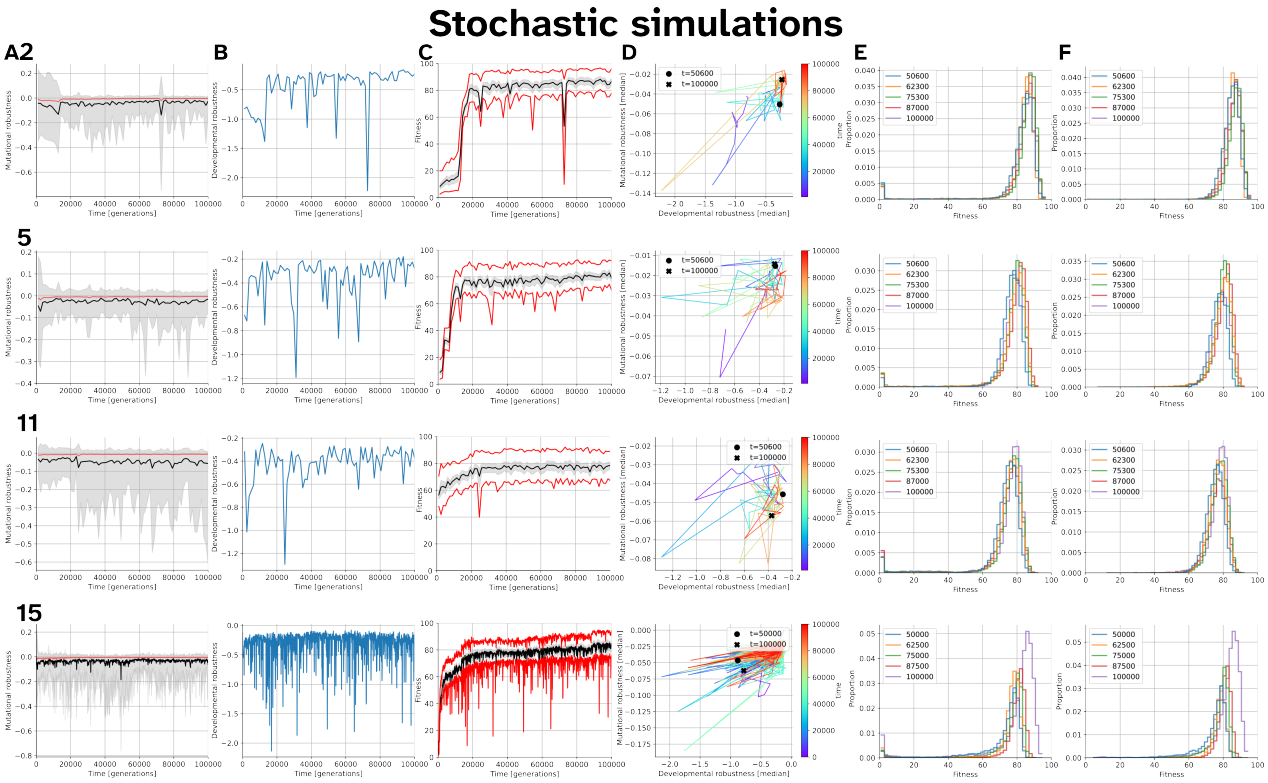

Figure S16. Stabilising selection of stochastic simulations. All plots show the stats of individuals along an ancestral lineage of stochastic simulations with their simulation ID indicated at the left of each row. **(A)** Mutational robustness (see eq. S7 over generational time. Red considers all offspring; black considers only the mutated offspring, IQR of the mutated offspring is shown in grey. **(B)** Developmental robustness (see eq. S8) over generational time. **(C)** Median fitness of clonal repeats (black) with IQR (grey) and upper 95% (red). **(D)** Mutational robustness plotted against developmental robustness for several timepoints along the line of descent (indicated by color). For mutational robustness we only considered mutated offspring. **(E)** Fitness distribution of 10,000 offspring. **(F)** Fitness distribution of only the non-mutated offspring in E.

References

Appendix S3 Divergence metrics

To measure the divergence of different GRNs we considered the following metrics: graph edit distance (GED), adjacency matrix difference, and Boolean adjacency matrix difference (see Fig S17). The GED is a measure of the minimal number of edits a graph needs to undergo in order to be similar to another graph. For the two matrix based comparisons we first transform the graph into two NxN matrices (one for activating and one for inhibiting regulatory interactions) for N unique gene types. If multiple copies of a gene exist, their incoming interactions of a specific type are summed. The difference between the adjacency matrix and the Boolean matrix is that any input > 0 in the adjacency matrix is set to 1 for the Boolean matrix. $\forall(i,j), M^B_{i,j} = \min(1, M^{adj}_{i,j})$ where M^{adj} denotes the adjacency matrix and M^B denotes the Boolean matrix.

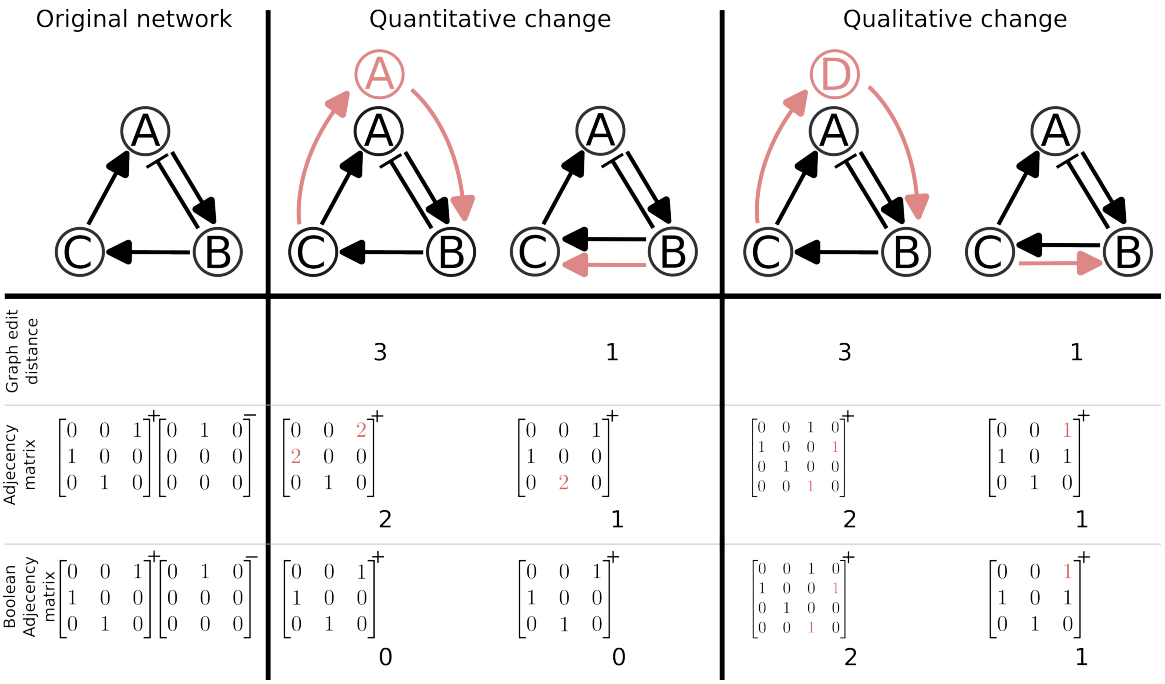

Figure S17. Different distance metrics for comparison of networks. The matrix notation of the networks also distinguishes activating ($+$) and inhibiting ($-$) interactions by storing them in separate matrices. The rows in the matrix represent genes and the columns by which gene they are activated/inhibited, so for example if gene 0 activates gene 2 then $M^+_{0,2} = 1$.

Supplementary Tables

	Value	Unit	Use
Mechanical parameters			
x_bounds	[-61, 61]	[um, um]	Horizontal bounds of the tissue
y_bounds	[0, 40]	[um, um]	Vertical bounds of the system
F_bound	10	a.f.u.	Boundary force
F_spring	1	a.f.u.	Spring constant
eq_dist	0.9	um	Equillibrium distance of the spring
Molecular parameters			
decay mRNA (α)	0.2 [0.05, 0.9]	t^{-1}	Initial decay rate [min, max]
decay protein (β)	0.2	t^{-1}	Decay rate of proteins
connections	3		Initial avg nr of connections per gene
TFBS hill constant (H)	40 [5, 100]	a.m.u.	Initial Hill constant [min, max]
Max gene expression rate (τ)	25 [5, 80]	a.m.u. t^{-1}	Initial max expression rate [min, max]
Dimer k_{on}	0.01 [0.00001, 0.1]		Initial k_{on} of the dimer gene [min, max]
Dimer k_{off}	0.001 [0.00001, 0.1]		Initial k_{off} of the dimer gene [min, max]
Diffusion coefficient (D)	1 [0.5, 5]	$um^2\ t^{-1}$	Diffusion coefficient [min, max]
Simulation parameters			
N_pop	1000		Population size
dev_steps	500		Developmental timesteps
generations	50000		Number of generations (initially)
N_genes	14		Number of unique genes
N_diff	2		Number of diffusing gene products
N_cellcell	2		Number of cellcell gene products
Mutational parameters			
gene duplication probability	0.002		per gene
gene deletion probability	0.003		per gene
gene parameter change	0.0005		Probability of changing decay mRNA or E
TFBS sign change	0.0002		per TFBS
TFBS type switch	0.0002		Probability of changing TFBS type per TFBS
TFBS duplication probability	0.0005		per TFBS
TFBS deletion probability	0.00075		per TFBS
TFBS innovation probability	0.00004 * genome size		max once per offspring
TFBS H change probability	0.002		per TFBS
Whole genome duplication probability	0.0005		per offspring
σ^2 max gene expression (τ) change	5		
σ^2 TFBS hill constant (H) change	10		
σ^2 mRNA decay rate (α) change	0.01		

Table S1. Table of the standard parameters used. a.f.u., arbitrary force unit. a.m.u., arbitrary molar unit. t, developmental time step.

Term	Abbreviation	Meaning
Developmental System Drift	DSD	The phenomenon where a trait maintains a constant phenotype but shows divergence in its underlying developmental pathway
Shoot Apical Meristem	SAM	Stem cell niche located in the apical tip of a (vascular) plant
Transcription Factor	TF	DNA binding protein which enhances/inhibits transcription
Transcription Factor Binding Site	TFBS	Site in the promoter where a TFBS can bind
Common Ancestor	CA	A shared ancestor between 2 or more individuals
Expression phenotype		The expression patterns of genes **not** included in the fitness function (genes $0-11$)
Fitness phenotype		The expression patterns of genes included in the fitness function (genes $12, 13$)
Gene Regulatory Network	GRN	
Conserved interaction		Regulatory interaction which was present continuously in an ancestry trace for at least 5000 generations
Pruned network		A minimal GRN which is obtained by iteratively removing genomic elements from a genome and accepting the removal if fitness is not impacted
Ancestry trace		A trace back through evolutionary time from an individual to it's first ancestor (Fig. 1)
WUSCHEL	WUS	MADS box TF associated with the stem cell pool in the OC
CLAVATA3	CLV3	CLE TF associated with stem cell differentiation and proliferation in the CZ
Conserved non-coding sequence	CNS	A non-coding sequence which is found to be conserved within a group of orthologs
CNS landscape		The set of CNSs associated with a specific gene
genotype-phenotype map	GPM	The function mapping from genotype space to phenotype space

Table S2. Terms and abbreviations used in this work.