nature genetics

Review article

https://doi.org/10.1038/s41588-024-01806-7

Rewiring gene circuitry for plant improvement

Received: 29 March 2024

Accepted: 17 May 2024

Alexander T. Borowsky & Julia Bailey-Serres

Aspirations for high crop growth and yield, nutritional quality and bioproduction of materials are challenged by climate change and limited adoption of new technologies. Here, we review recent advances in approaches to profile and model gene regulatory activity over developmental and response time in specific cells, which have revealed the basis of variation in plant phenotypes: both redeployment of key regulators to new contexts and their repurposing to control different slates of genes. New synthetic biology tools allow tunable, spatiotemporal regulation of transgenes, while recent gene-editing technologies enable manipulation of the regulation of native genes. Ultimately, understanding how gene circuitry is wired to control form and function across varied plant species, combined with advanced technology to rewire that circuitry, will unlock solutions to our greatest challenges in agriculture, energy and the environment.

Plants are essential to human life on earth. They have important roles from the planetary scale to societal and cultural practices. They modulate global atmospheric CO_2 and oxygen levels, cool our cities and nourish our bodies. Like animals, plants are wired to respond to diverse environmental cues that enable them to thrive in constantly changing environments. However, plants differ from animals in their decision making in important ways. First, plants lack a central nervous system to interpret and respond to stimuli. Second, plants undergo continuous development, maintaining populations of pluripotent stem cells called meristems, which differentiate and continue to produce new aboveground and belowground organs throughout the plant's life, unlike most animals, which may continue to grow in size but have a determinant body plan. This means that plants can respond to environmental cues or challenges by modifying their development. While an animal can respond to its environment through neuronal networks that influence behavior, plant responses are hardcoded in their genomes and comprise networks of genes, which are turned on and off in specific contexts through signaling pathways to tune development and physiology.

A deep understanding of the gene regulatory circuitry underlying this plasticity presents a unique opportunity to address grand challenges of our time. Climate change is causing more extreme and frequent weather events, including droughts, floods and temperature oscillations, which already reduce crop productivity[1–3]. Fertilizer use is inefficient and damaging to aquatic ecosystems, and its production is unsustainable and its use unaffordable for many smallholder farmers in the Global South[4,5]. Remarkably, plants have adapted to thrive in varied and even extreme ecosystems. Knowledge of the genetic, developmental and physiological mechanisms of this trait variation and plasticity has provided important insights into strategies for crop improvement[3,6,7].

Understanding the gene regulatory circuitry by which environmental cues are integrated into developmental and physiological decisions is only the first step in manipulating desired traits. It is well known that specific genes are conditionally programmed in specific cells over time and space. Single-cell technologies that provide DNA, RNA, protein and metabolite readouts are rapidly magnifying this understanding, yet wide deployment of genetically engineered crops has mostly been limited to traits resulting from the constitutive expression of genes that convey herbicide tolerance or pest resistance. Technological strategies enabled by CRISPR–Cas genome editing and synthetic transcriptional circuits provide an avenue to manipulate gene regulatory circuitry in new ways and in specific cells, which can move mechanisms of resilience and productivity across species or augment or otherwise engineer existing circuitry. Here, we review recent findings in defining and manipulating gene regulatory circuitry, which could be used to engineer improved fitness in adverse environments or develop plant-based solutions in the bioeconomy.

Defining plant gene regulatory networks

Transcription factors (TFs) are proteins that bind to a specific DNA sequence (TF binding site; TFBS) in promoters or genic regions[8,9] and

Center for Plant Cell Biology, Department of Botany and Plant Sciences, University of California, Riverside, Riverside, CA, USA. e-mail: serres@ucr.edu

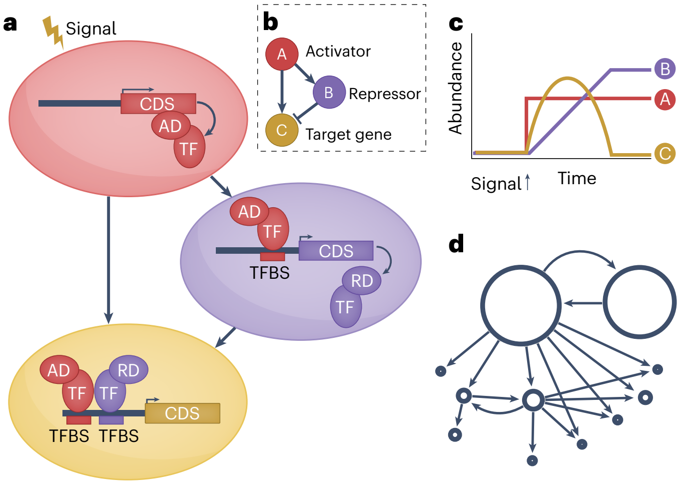

Fig. 1 | Gene regulatory network architecture. **a**, Model of a simple gene regulatory network (GRN). TFs bind DNA of targets at specific short sequences (TFBSs) and either promote, through activation domains (ADs), or prevent, through repression domains (RDs), the transcription of genes, including those encoding other TFs. The coding sequence (CDS) of the transcribed and processed mRNA can be translated into proteins with functional roles in the cell. **b**, A simplified version of the network represented in **a**, forming the motif of an incoherent feedforward loop. **c**, Theoretical activity of the GRN over time, where TF A is turned on by a signal. This functions as a pulse generator, as gene C is active only before it can be repressed by B. **d**, A theoretical hierarchical network in which the size of each node (circle) is weighted by its connections using the PageRank algorithm. For a GRN, this would reveal driver or high-level TFs that likely are upstream components as larger nodes. Data are from ref. 99.

promote, prevent or finetune the transcription of one or more genes. TFBSs of many plant TFs are known because of high-throughput assays, including protein-binding microarrays and DNA affinity purification and sequencing (DAP-seq)[10,11]. The DNA-binding domains of TFs are generally well conserved; therefore, TFs conserved across species can be grouped by homology and associated TFBSs can be inferred[12]. At the simplest level, the promotion or prevention of transcription is largely achieved through activation domains or repression domains of TFs or associated proteins, which either recruit or prevent the recruitment of transcriptional machinery to the transcription start site (Fig. 1a). Activation and repression domains are more difficult to identify and are less conserved at the level of protein sequence, but recent work has begun to systematically characterize these domains in plant TFs[13]. TFs themselves are encoded by genes and thus are regulated by other TFs. This gene regulatory logic can be represented at the system level as a gene regulatory network (GRN), in which nodes represent genes, including TFs, which are connected by edges that represent regulation (Fig. 1a,b). This abstraction necessitates a simplification of additional levels of gene regulatory control (post-transcriptional and post-translational), which can exist within the nodes of a GRN. These post-transcriptional modes of control can disproportionally affect TFs, notably by microRNA-mediated transcript degradation[14] and translational control by upstream open reading frames[15,16].

Functionally, two major modalities have been used to infer GRNs. The first approach is to consider transcriptome data. TFs and their targets are often coexpressed, and the cohort of TFs for which mRNA abundance, across time, tissues, developmental stages or conditions, best predicts the abundance of a target gene can be predicted as its regulators[17]. Single-cell and single-nucleus transcriptome data provide an opportunity to map these relationships at a fine scale, particularly as cell identities change throughout development[18]. The second approach is to consider protein–DNA interaction data, captured through direct evidence from DAP-seq, chromatin immunopurification followed by high-throughput sequencing (ChIP–seq) or yeast one-hybrid approaches or indirectly by searching for sequence matches to a known TFBS[11,19]. Of these, only TF ChIP–seq evaluates TF binding within the native chromatin context, albeit at low throughput[19]. DAP-seq has the advantage of capturing potential binding locations independent of chromatin states and also can characterize simultaneous binding of multiple TFs through the newer double DAP-seq method[20]. Ultimately, combining these modalities, by considering which TFs both bind DNA proximal to a target and predict its expression, can better resolve GRNs[21–23].

Other genome-wide epigenomic data can narrow down regulatory regions of DNA, including those obtained from assay for transposase-accessible chromatin with sequencing (ATAC–seq)[24]. This approach identifies open chromatin regions that often coincide with functional TFBSs and has been widely implemented at the single-cell level[25]. ChIP–seq recognizes histone modifications on nucleosomes in genic regions and can also indicate active or repressed transcriptional states[26,27]. A frontier is to identify pioneer TFs, which initiate changes in chromatin state, to refine cause-or-consequence relationships with chromatin accessibility and TF binding. An example is LEAFY, which controls the vegetative-to-reproductive growth transition in the *Arabidopsis* shoot meristem by binding to nucleosomal DNA and recruiting chromatin modifiers[28,29]. Systematic identification of this special class of TFs with pioneer activity, which has begun in mammalian systems[30], will ultimately enhance understanding of GRNs.

Architecture of GRNs

The inherent structure of networks can reveal their function (Fig. 1). Network motifs, specific patterns of connections between a subset of nodes, can play specific roles in information processing. For example, the feedforward loop network motif is one way through which three nodes can be connected in a network. As characterized in bacteria and yeast, an incoherent feedforward loop consisting of an activator, repressor and target (Fig. 1b) can serve as a pulse generator; if an external cue activates A, C will be activated until B is transcribed and translated to then repress C[31] (Fig. 1c). Coherent feedforward loops (protein A activates B; A and B activate C) can conversely serve as 'persistence detectors' in which the signal activating A must remain at or above a certain threshold to ensure that the synthesis of B is sufficient to activate C. This persistence phenomenon can be important in cell fate commitment: feedforward loops are abundant in the regulation of xylem development, the vascular tissue that transports water and nutrients from the root system to the body of the plant[32]. The TF VASCULAR-RELATED NAC DOMAIN 7 participates in more than one of these feedforward loops and is sufficient to activate bistable switch behavior in committing cells to xylem identity in *Arabidopsis*[33], consistent with the persistence detector function of feedforward loops. Globally, enrichment or depletion of certain network motifs is a hallmark of biological networks. In fact, human cell type-specific GRNs share similar patterns of motif enrichment and depletion with networks that map neuronal connectivity in *Caenorhabditis elegans*[34].

Time is often critical to the function of GRNs. The core circadian clock GRN is a negative feedback loop composed of the partially redundant MYB TFs CIRCADIAN CLOCK-ASSOCIATED 1 (CCA1) and LATE ELONGATED HYPOCOTYL (LHY) that activate the regulator TIMING OF CAB EXPRESSION 1 (TOC1) to subsequently repress CCA1 and LHY[35]. This module acts as a time-sensitive gatekeeper to permit or restrict activity of GRNs across day cycles. For many genes, the magnitude of response to heat varies by time of day[36]. On a finer time scale, cascades (successive activation of TFs) in GRNs temporally tune responses to nitrogen availability in *Arabidopsis*. Here, by fine-scale transcriptional profiling, TFBS enrichment analysis and GRN construction, a transcriptional regulatory hierarchy that cascades within 2 h could be established[37]. Another example of transience is the TF bZIP1, which regulates many more genes in response to nitrogen than it stably binds to. Working by so-called 'hit-and-run', it transiently interacts with a *cis*-element to promote transcription, while licensing the promoter for activity by other TFs that bind to co-occurring TFBSs[38]. Transient response networks are likely to be essential for tuning responses to external cues.

Often, predicted networks contain a large number of TFs, and further analysis is necessary to identify which are the most crucial components. Network motifs enriched in biological systems have been used to identify important regulators in plants by prioritizing TFs that participate more frequently in motifs with functional importance[39]. Other network science approaches can be implemented to identify important nodes of global GRNs. For example, driver TFs were predicted using the PageRank algorithm[40] in a mouse embryo network[41]. This method, developed to prioritize search engine results, involves computationally reassigning weight along edges in a network (in the case of websites, hyperlinks; for GRNs, TF–target gene interactions). By initializing the network with expression values and then running the PageRank algorithm, TFs with more connections to expressed genes receive higher PageRank scores and TFs upstream of those TFs receive even higher PageRank scores[41] (Fig. 1d). This method was applied to predict key regulators of xylem development in rice, recovering orthologs to many key regulators from *Arabidopsis*[42]. If we assume that the architecture of GRNs enables their function in developmental and response regulation, this can guide the functional identification of regulators of phenotypes. Also, the types of network structures (motifs) that control different developmental or response programs can guide efforts to rewire and augment these networks in a desired manner.

GRN conservation and variation in the plant kingdom

While plants display an immense diversity of form and function, the gene regulatory grammar underlying the networks that encode this variation is generally conserved. A core concept of evolutionary developmental biology is that the presence or absence of *cis*-elements, largely TFBSs, is more mutable than genes themselves and that TFs are a toolkit, the spatiotemporal deployment of which contributes to morphological variation across organisms[43,44]. Studies defining gene regulatory programs across land plant species reinforce the tenet discovered in animals that cognate TFBSs of homologous TFs across species tend to be conserved[12], and the repertoire of TFBSs deployed across plant genomes is largely similar[45]. These conserved parts are likely deployed in similar contexts, with transcriptomes of specific land plant organs showing conserved identity, although to varying degrees, depending on the organ and species[46]. This conservation may highlight constrained steps in development where conservation is deeper, such as in the development of grass inflorescences[47] and root meristems[48]. The conserved toolkit of GRN parts enables the application of common approaches across species to study and compare GRNs, the extensive diversification of which mirrors the diversity of plant forms.

Differences in the deployment of key regulatory TFs can explain differences in developmental forms (Fig. 2a). A TF and its targets can be referred to as a ‘regulon’. A regulon can be partially or entirely conserved, but differences in the expression and location of the TF controlling it can drive major changes. While *Arabidopsis* develops only one layer of cortex cells in its roots, most grasses develop several (Fig. 2b). A key regulator of this process is SHORTROOT (SHR), a member of the GRAS TF family[49]. In *Arabidopsis*, SHR moves one cell layer outward from its site of synthesis in the stele to the endodermis to regulate the division and identity of the endodermis and the cortex. By mapping the activity of orthologous genes in maize roots by single-cell RNA sequencing and use of transcriptional and translational reporters, it was shown that *SHR1* is transcribed in the root endodermis and the protein then moves outward through the cortical cell layers toward the epidermis[50]. Mutants in these genes in both maize and the grass *Setaria viridis* have fewer cortical cell layers than wild type, suggesting that the differences seen in the transcriptional domain and subsequent protein localization for SHR in these plants contribute to the morphological differences observed in these species[50].

Another example of varied regulon deployment is the transcriptional network driving suberin biosynthesis (Fig. 2c). Suberin is a waxy polymer deposited into the cell wall that provides a diffusion barrier and plays an important role in plant–environment interactions[51,52]. In *Arabidopsis* roots, suberin is deposited in the endodermis and is regulated by a redundant set of MYB family TFs[53]. Tomato lacks a suberized endodermis but instead forms a suberized exodermis, a subspecialized outer cortex layer[54]. Its suberin biosynthesis genes are expressed in the exodermis and not in the endodermis and possess promoters enriched for MYB TFBSs, indicative that homologous MYBs, also preferentially expressed in the exodermis, regulate this suberization[4854]. This provides an example of transcriptional regulators maintaining conserved target genes. Other suberization patterns exist, including in rice, which has a suberized endodermis and an environmentally dependent suberized exodermis[42,55], likely also due to varied deployment of MYB regulons. Indeed, regulator redeployment to new cell types may be a common mode of cellular divergence across species. Using single-cell and single-nucleus RNA sequencing followed by GRN inference in the grasses maize, *S. viridis* and sorghum, 15 regulons showed the same cell type specificity in all three species, whereas 50 regulons had swapped cell types[56]. An example is a mucilage-related gene regulon repurposed for expression in the maize columella from ancestral expression in the cortex.

Another example of regulator redeployment occurs in the development of nodules, root organs in legumes that enable symbiosis with nitrogen-fixing bacteria. In both *Medicago truncatula* and *Lotus japonicus*, the symbiosis signaling pathway converges on the lateral root development regulatory LATERAL ORGAN BOUNDARY (LOB) domain TF LBD16[57,58] (Fig. 2d). First characterized in *Arabidopsis*, AUXIN RESPONSE FACTOR TFs promote LBD16 expression to regulate lateral root formation[59]. In an innovation specific to nodule formation, the symbiosis-driven TF NODULE INCEPTION (NIN) also activates LBD16, driving a GRN that promotes cell proliferation[57,58]. In this case, redeployment of the regulator might be a preferable mode of adaptation, because repurposing of the program for lateral root initiation produces an output that only required modest tuning to give rise to nodulation, which suggests that this rewiring could be reproduced to engineer this symbiosis in non-legumes.

Variation is not limited to the redeployment of regulators. GRN rewiring also involves *cis*-regulatory changes for target genes. For example, C4 photosynthesis has evolved repeatedly as a mode of higher-efficiency carbon fixation under lower atmospheric CO_2 concentrations from ancestral C3 photosynthesis species[60]. In C3 plants, carbon fixation occurs predominantly in mesophyll cells; however, the enzyme RuBisCO, which incorporates CO_2, also reacts with O_2 in photorespiration, limiting the capacity to fix carbon[60]. In C4 plants, this carboxylation reaction is located in adjacent bundle sheath cells, and a metabolic ‘pump’ concentrates CO_2 in these cells, reducing photorespiration[60]. This metabolic compartmentation necessitated specific spatial patterning of photosynthesis gene expression not found in the ancestral state[60] (Fig. 2e). While many modes of regulatory evolution might produce this novel partitioning, recent evidence suggests that target genes within C4 grass species acquired or lost TFBSs that placed them under regulation of TFs for which cell type-specific expression is conserved. One key TF family implicated in regulation of photosynthetic genes across species is the GOLDEN2-LIKE (GLK) TFs[61]. Initially, it appeared that duplication and subfunctionalization of these MYB TFs contributed to C4 evolution; however, recent analysis of *glk* mutants and complementation lines in *S. viridis* suggests that both GLK1 and GLK2 retain ancestral function[62]. TF binding across two eudicot and three monocot species indicates that very few GLK-bound targets are highly conserved, demonstrating widespread *cis*-regulatory rewiring[63]. More recent studies resolved gene activity through single-nucleus ATAC–seq

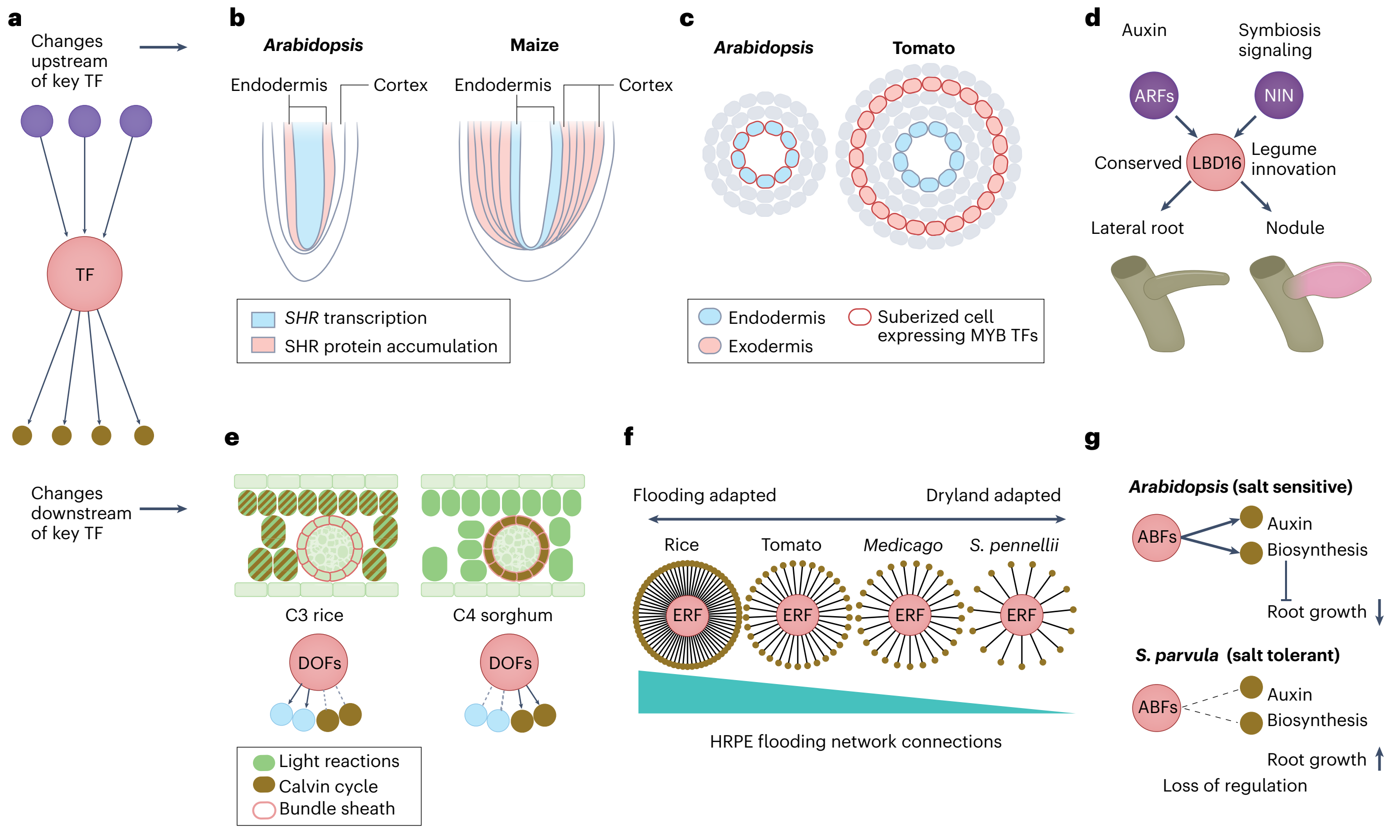

Fig. 2 | Modes of GRN evolution that drive variation in development and environmental plasticity. **a**, Changes to GRNs can either have consequences upstream, such as redeployment of regulators (**b**–**d**) of key TFs, or downstream, for example, *cis*-regulatory changes of targets (**e**–**g**). **b**, Redeployment of the TF SHR in maize contributes to the development of additional root cortical cell layers. **c**, Redeployment of MYB TFs from the root endodermis to the exodermis of the tomato contributes to the function of barrier layers in controlling the flow of water and nutrients into the plant. **d**, Lateral root transcriptional programming by TFs that mediate the regulation of auxin signaling via AUXIN RESPONSE FACTORS (ARFs) and symbiosis via NIN converges on the TF LATERAL ORGAN BOUNDARY DOMAIN 16 (LBD16), which is repurposed for nodulation in legumes. **e**, C4 species such as sorghum express distinct photosynthetic gene programs for light reactions in the mesophyll and the Calvin cycle in the bundle sheath, while, in C3 plants, all photosynthetic programs occur in the mesophyll. Conserved bundle sheath TFs, such as DNA BINDING WITH ONE FINGER (DOF), activate one set of genes in C3 rice (light blue circles) and a different set in C4 sorghum, which facilitates distinct metabolic programs (brown circles). **f**, Networks of genes upregulated by short-term submergence under the control of group VII ERF TFs across species, with more widespread regulation in flooding-adapted rice[67]. Each node represents two targets. **g**, ABA RESPONSE FACTOR (ABF) TFs are conserved in *Arabidopsis thaliana* and *Schrenkiella parvula*, a salt-tolerant relative. In *Arabidopsis*, these TFs activate auxin biosynthesis genes (brown circles), which increase auxin levels and reduce growth, while, in *S. parvula*, the ABFs do not regulate auxin biosynthesis, allowing growth to be maintained.

and RNA sequencing in C3 and C4 grasses to provide the clearest evidence to date of *cis*-regulatory rewiring in C4 evolution[64,65]. Both studies identified DNA BINDING WITH ONE FINGER (DOF) TFs as key drivers of bundle sheath expression; however, bundle sheath transcriptomes of C3 rice and C4 sorghum were the most divergent of all cell types[65]. These data demonstrate that remodeling of *cis*-regulatory regions aided the evolution of C4 photosynthesis, a crucial understanding that may enable engineering more efficient C4 metabolism into C3 crops including rice.

Cis-regulatory wiring not only contributes to inherent physiological differences between species, it is also important in the context of environmental responses. For example, a short-term early submergence response network orchestrated by conserved sets of TFs across species is distinct in flooding-adapted rice, which was domesticated from semi-aquatic progenitors. In the conserved response network of rice, *Medicago*, domesticated tomato and dryland-adapted *Solanum pennellii*, rice relies more on the group VII ETHYLENE RESPONSE FACTOR (ERF) TFs that bind hypoxia-responsive promoter elements (HRPEs)[66,67] (Fig. 2f). Responsive genes in rice show greater increases in the accessibility of TFBSs and concomitantly stronger transcriptional induction than in the less flooding-adapted *Medicago*[67]. One consequence of long-term submergence in rice roots is a cessation of cell division in root meristem cells until desubmergence[42], but any connection to the short-term hypoxia-response network involved in controlling central carbon and anaerobic metabolism is not yet defined at the GRN level. Ultimately, expanding the HRPE network in flooding-intolerant crops could help reduce yield losses due to floods.

Osmotic stress adaptation is another example for *cis*-regulatory rewiring. Looking across *Brassicaceae*, the salt-resilient extremophile *Schrenkiella parvula* responds to abscisic acid (ABA), the hormone crucial for the response to desiccation or osmotic stress, not by reducing root growth, as expected from the model *Arabidopsis*, but rather with increased root growth[68]. A DAP-seq approach determined that ABA-response element (ABRE)-BINDING FACTOR (ABF) TFs that directly respond to ABA signaling and their target TFBSs are conserved across the species examined. Notably, sequence changes in *cis*-regulatory regions were found to contribute to interspecies differences in response (Fig. 2g). In *Arabidopsis*, ABFs activate auxin biosynthesis and signaling, which suppresses root growth. By contrast, *S. parvula* has lost ABF binding in the promoters of many auxin genes, resulting in reduced auxin signaling and maintained root growth[67]. This illustrates that *cis*-regulatory variations can confer opposed outcomes that impact development across species and suggests that targeted modification of TF activity in *cis* or *trans* might help promote growth of crops in extreme environments.

Synthetic biology for plant GRN manipulation

Plant biotechnology exploded with the advent of *Agrobacterium*-mediated and tissue bombardment methods to insert engineered genes randomly in the plant nuclear genome. These technologies are sufficient to manipulate plant GRNs. Chimeric promoter–coding sequence fusions can add new nodes under specific cellular or conditional regulation. If the coding sequence used encodes a TF, a new node can target downstream native components of the network. Constitutive overexpression of TFs has contributed to crop improvement[69–71], but sometimes it may be more desirable to exert spatiotemporal control. For example, driving MYB and bHLH TFs from snapdragon with a fruit-specific promoter activated anthocyanin biosynthesis in tomato fruits, turning them purple and enriching their nutritional quality[72]. Recent developments in synthetic biology greatly expand the capacity to generate large, multi-gene cassettes made of modular parts, making more complex additions to the genome possible[73].

Introduction of orthogonal (unreactive with native components) and synthetic (new-to-nature) parts into plants to form GRNs is well established. The earliest *in planta* demonstration used the GAL4 TF from yeast and its corresponding TFBS (upstream activation sequence)[74]; by coexpressing *GAL4* under an endogenous cell type- or condition-specific promoter, a second gene with upstream activation sequence *cis*-elements could be spatiotemporally controlled. The current repertoires of orthogonal and synthetic TF and promoter combinations include tunable expression[75] and synthetic repression[76]. These tools enable engineering of logic gates, systems that respond specifically and predictably to the presence or absence of multiple inputs. By using logic gates based on a recombinase or integrase system, inputs can activate stable genetic changes within somatic cells[77,78]. Additionally, one- or two-input logic gates using specific protein–DNA combinations can act transiently. Single-input BUFFER gate promoters, which contain operators (bacterial TFBSs) with varied affinity for the synthetic TF AmtR-VP16 were used to drive varied levels of transcription of *solitary root (slr)* in lateral root meristem cells in *Arabidopsis*[79] (Fig. 3). Gain-of-function mutations of *SLR* like *slr-1* result in primary roots that lack lateral roots, along with pleiotropic effects in primary root growth, hair development and gravitropism[80]. These phenotypes were avoided by limiting *slr-1* transcription to the pericycle and lateral root meristems. In this manner, quantitative tuning of *SLR* yielded phenotypic variation in lateral root density[79], elegantly engineering gene regulatory circuitry to control plant form.

CRISPR–Cas systems also present a fertile opportunity to create synthetic TFs in plants. Using a deactivated Cas9 (dCas9) protein fused to activation or repression domains, transcriptional repression or activation can be recruited to specific endogenous genes by targeting with guide RNAs (gRNAs) to tune transcription[81]. Because gRNAs can be transcribed in arrays, a single construct can be used for dCas9 activation of a suite of target genes[81].

These orthogonal systems rely on endogenous promoters; yet few of these promoters are well characterized, and they all are restricted to the existing sequence space and spatiotemporal patterning. Also, promoter sequences can be quite large (>2 kb) and cumbersome for engineering. Synthetic promoters that recruit endogenous TFs thus can expand the repertoire of promoters available to engineers. One approach is to leverage known *cis*-elements or TFBSs upstream of a minimal promoter that are known to recruit a single TF family under a specific condition. This approach has enabled the design of auxin-responsive *DR5*[82], ABA-responsive *6xABRE*[83], hypoxia-responsive *3xHRPE*[66] and general stress-responsive *4xRSRE*[84] promoters. A powerful approach may be to include more varied TFBSs in a synthetic promoter owing to their cooperativity by combining TFBSs from different TF families to achieve higher levels of transcription than an equal number of TFBSs from a single family. This strategy of combining TFBSs has already produced constitutive or near-constitutive synthetic promoters with predictable strengths[85].

The throughput in measuring promoter activity is a limiting factor for assessing promoter function. STARR-seq (self-transcribing active regulatory region sequencing) and similar 'massively parallel reporter assays'[86] have helped generate informative data for synthetic promoter design. In the application of STARR-seq to plants, the enhancer activity of DNA sequences is measured for a library of fragments cloned into the promoter region of a reporter construct, transfected into protoplasts or *Nicotiana benthamiana* leaves, before the capacity of the fragments to promote transcription is quantified by RNA sequencing of transcribed barcodes[87]. In this manner, the enhancer activity of all core promoters of *Arabidopsis* and maize genes could be quantified, and subsequent machine learning models were used to design synthetic minimal promoters with predictable strengths[88]. However, to date, these approaches have measured bulk rather than cell-specific transcription, and more work is needed to design promoters with both predictable spatiotemporal activity and strength. As an example, a cell type-specific enhancer design was achieved with machine learning approaches that built on single-cell ATAC–seq data in animals[89,90]. It will be valuable to evaluate whether these approaches are also effective in plants to finetune beneficial phenotypes.

Manipulating GRNs through gene editing

CRISPR–Cas can be used to manipulate GRNs through loss-of-function mutations in the gene body or mutations in the promoter. Double-strand breaks and resultant indels induced in the coding region of a TF gene can remove a node from the network entirely. Currently, loss-of-function mutagenesis is an imprecise but useful tool for genetic study, complicated by gene redundancy and undesired pleiotropic consequences. Therefore, it can be more desirable to tune the expression of a TF or other target genes in *cis* (Fig. 3b). In tomato, the peptide CLAVATA3 (*Sl*CLV3) provides a negative feedback loop to restrict the domain of the cell proliferation-promoting TF WUSCHEL (WUS) in shoot meristems[91]. Scanning mutagenesis of the *SlCLV3* promoter with tiled CRISPR–Cas9 gRNAs generated alleles with variable-sized deletions and inversions. These conferred a spectrum of *SlCLV3* expression in the floral meristem, with reductions allowing higher *WUS* expression and more meristem cell proliferation, resulting in a larger fruit with more locules[92,93] (Fig. 3c). While the full set of TFs regulating *SlCLV3* is not known, some of the deletions almost certainly resulted in the loss of regulation by specific TFs.

CRISPR–Cas-mediated promoter editing has been used in a number of subsequent studies, some of which leveraged information on chromatin accessibility in the selection of promoter regions for targeting. Targeted deletion of the promoter region of *IDEAL PLANT ARCHITECTURE 1* (*IPA1*) in rice ultimately demonstrated that the deletion of a single edge of a GRN can improve crop productivity[94]. Here, the authors started from the knowledge that gain-of-function alleles of *IPA1* that eliminate microRNA targeting and broadly increase mRNA levels confer large panicles and higher yield but reduce tillering[95,96]. By tiling the *IPA1* promoter with CRISPR–Cas9 gRNAs, deletions were identified that promote elevated mRNA in the young panicle but not in the shoot base where tillers originate[94]. These conferred advantageous large panicles without the tradeoff of reduced tillers. Identification of the binding site for the TF AWNLESS1 (AN1) within the deleted region prompted demonstration by transactivation assays that AN1 binds this region of the *IPA1* promoter to limit its transcription[94] (Fig. 3d). Another rice study targeting genes involved in grain and plant architecture traits used a predictive framework based on sequence content, chromatin state and evolutionary conservation to hone in on regulatory regions. The authors then used Cas12a, which is more likely to generate large deletions than Cas9, to generate agronomically relevant quantitative trait variation[97].

These pioneering studies screened deletions that were targeted to broad regulatory regions, rather than precise TFBSs. Improvements in editing technologies, greater knowledge of temporal and

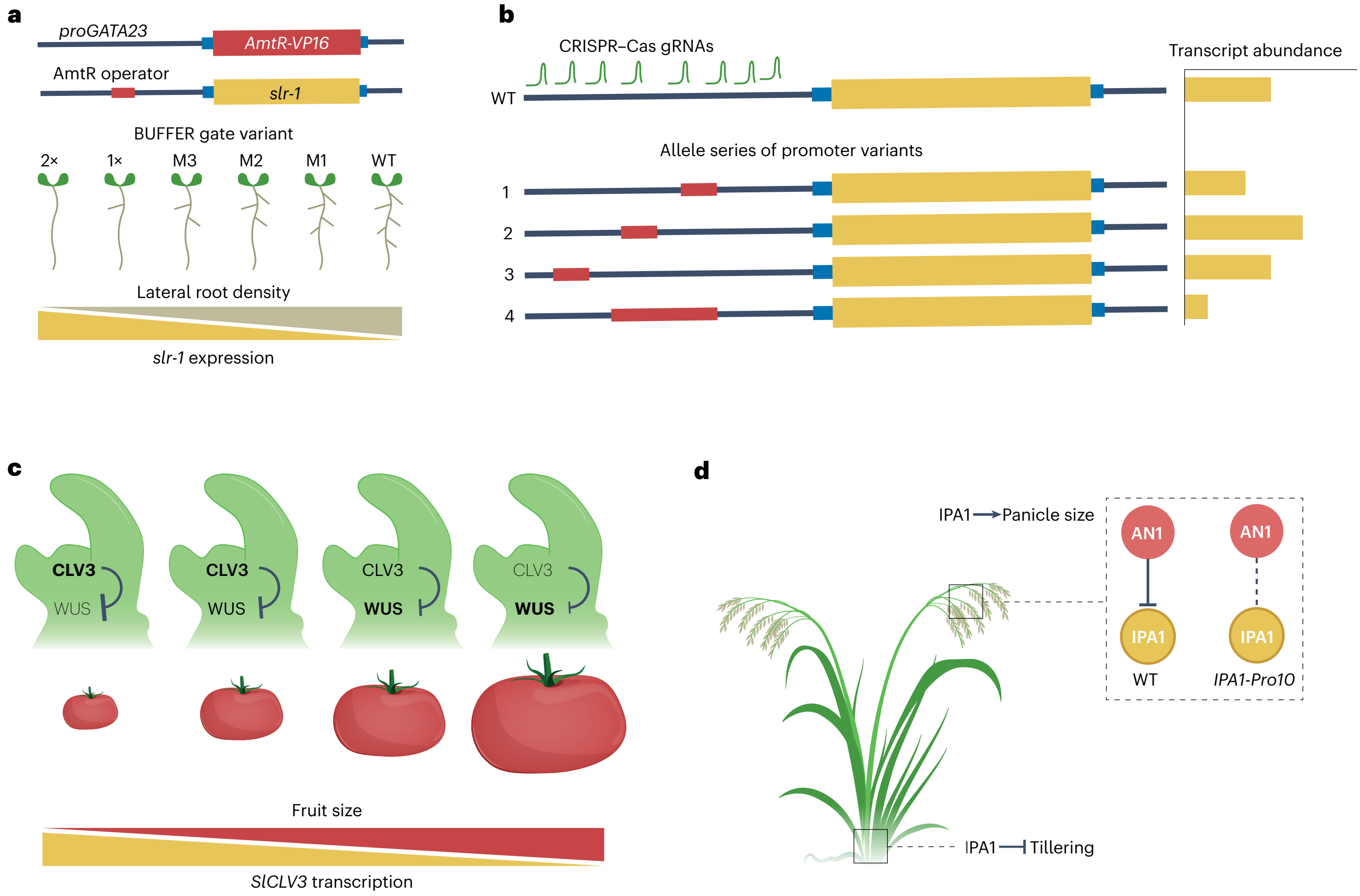

Fig. 3 | Synthetic biology and gene-editing approaches to network manipulation. **a**, Synthetic gene circuits to generate quantitative trait variation. The *Arabidopsis GATA23* promoter (proGATA23) directs transcription of the gene encoding the synthetic TF AmtR-VP16, a bacterial DNA-binding TetR family protein fused to the VP16 transactivation domain from herpes simplex virus, in lateral root meristems. By using variants of BUFFER gates (tools to reduce or amplify the signal of an input), promoters with one (1×) or two (2×) copies of the AmtR operator sequence or a copy with various mutations (M1 to M3), different levels of transcription of *slr-1*, encoding a dominant repressor of lateral root emergence, are induced. These levels give rise to quantitative variations in lateral root density. WT, wild type. **b**, Use of genome editing for network manipulation, exemplified here with a general strategy to generate *cis*-regulatory variation by using CRISPR–Cas9. gRNA species distributed across the promoter are delivered into plants together with Cas9 for gene editing. The resulting genotypes with insertions, deletions or inversions in regulatory regions can give rise to either gain or loss of regulation by TFs, altering transcription of the target gene. **c**, CRISPR–Cas9-mediated gene editing of *SlCLV3* results in a variation in tomato fruit size, proportional to the reduction in the level of *Sl*CLV3. **d**, Gene editing of *IPA1* in rice increases panicle size without pleiotropic reductions in tillering. IPA1 is expressed in the shoot base, where it represses the formation of new branches called tillers. It is also expressed in the young panicle, where it promotes proliferation and therefore yield. *IPA-Pro10*, CRISPR mutant with a deletion in the *IPA1* promoter that includes a binding site for the TF AN1. Removal of this GRN edge promotes elevated IPA1 expression in the panicle but not in the shoot base.

cell type-specific chromatin accessibility and GRN modeling combining omic data types should enhance the prediction of the *cis*-elements that contribute to the spatial distribution, timing and amplitude of transcription and thus their editing. These could also help to predict which TFBSs could be inserted, and at which locations, to have a consequential effect on transcription. This level of control may be required to expand quantitative trait variation beyond that of extant germplasm or species, which could be essential for fitness under climate change, which creates environments at the extremes of species' adaptive ranges.

The path forward

Advancements in plant functional genomics, single-cell biology and modeling will continue to define and refine GRNs. As these technologies become more commonplace, biologists will be able to map out the genes and networks underlying the variation seen across the plant kingdom. Studying how diverse species make developmental and physiological determinations in the natural environment or diverse agricultural ecosystems can help inform engineering of crops for the breadth of desirable traits, from biomass and yield to sustainability and nutrition.

The pioneering work described here provides a roadmap to rewiring genetic circuitry with specific goals in mind. First, phenotypic variation across or within species for a trait of interest is identified. Next, using advanced omic techniques, GRNs underlying those traits are predicted, before the structure, variation and conservation of network components is exploited to identify changes necessary for the desired phenotype. Finally, synthetic biology and gene editing are deployed to achieve effective rewiring of gene circuitry.

While such an approach is entirely possible now, advancements are imperative. First, our understanding of plant phenotypic diversity across species at the GRN level is a frontier. Here, plant genetic engineers can collaborate with botanists, physiologists and ecologists to identify opportunities to further expand our knowledge of GRN wiring. This should include an equal and fair exchange with indigenous scientists and stewards of native species, crops and landraces, which provide opportunities to explore adaptive genetic strategies as well as to apply engineering tools.

Second, despite advances in GRN modeling, the existing studies have characterized either only small subsets or specific contexts of the potential GRN space. Further developments in modeling methods are needed to generate comprehensive models of GRNs that include their activity throughout daily circadian cycles and the overall life cycle of the plant in natural, often stressful environments. In the future, as omic resources expand across species, the rules of genome organization and function might be defined to such an extent that GRNs can be confidently predicted from genome sequence alone. In the meantime, target gene (rather than TF)-focused molecular approaches to query the cohort of TFs and TFBSs that regulate a gene of interest[98] could provide invaluable information for engineering traits.

Finally, tools to rewire genetic circuitry must expand to allow precise and targeted changes without the need for cumbersome transformation and selection methods. For *cis*-regulatory regions, improvements in GRN modeling could guide targeted insertion or deletion of TFBSs with predictable impacts on gene activity, and improvements in editing technologies could deploy widespread GRN rewiring across greater numbers of genetic loci in parallel. For synthetic *cis*-regulatory sequences that could be designed with additional advances, a transgene could be activated with predictable cell type, environment and strength characteristics, purely by recruiting endogenous factors; for now, the synthetic tools that are available allow engineers to modulate one or more inputs through a synthetic circuit. Lastly, it is important to acknowledge that transcriptional regulation is not the only mode of control of gene activity, let alone phenotype. Factors and features that control DNA methylation, chromatin structure, and all cotranscriptional and post-transcriptional steps of RNA biology, including mRNA modification, maturation, translation, sequestration, decay and regulation by long noncoding and microRNAs, are important and provide additional engineering opportunities.

In sum, the application of plant functional genomics for gene regulatory circuitry discovery, along with new developments in genetic engineering and synthetic biology, provide valuable opportunities for plant scientists and engineers to meaningfully and expeditiously contribute to solving the grand challenges of our time.

References

1. Challinor, A. J. et al. A meta-analysis of crop yield under climate change and adaptation. *Nat. Clim. Change* **4**, 287–291 (2014).
2. Lobell, D. B., Schlenker, W. & Costa-Roberts, J. Climate trends and global crop production since 1980. *Science* **333**, 616–620 (2011).
3. Bailey-Serres, J., Parker, J. E., Ainsworth, E. A., Oldroyd, G. E. D. & Schroeder, J. I. Genetic strategies for improving crop yields. *Nature* **575**, 109–118 (2019).
4. Vitousek, P. M., Aber, J. D. & Howarth, R. W. Human alteration of the global nitrogen cycle: sources and consequences. *Ecol. Appl.* **7**, 737–750 (1997).
5. Mueller, N. D. et al. Closing yield gaps through nutrient and water management. *Nature* **490**, 254–257 (2012).
6. Mickelbart, M. V., Hasegawa, P. M. & Bailey-Serres, J. Genetic mechanisms of abiotic stress tolerance that translate to crop yield stability. *Nat. Rev. Genet.* **16**, 237–251 (2015).
7. Jez, J. M., Lee, S. G. & Sherp, A. M. The next green movement: plant biology for the environment and sustainability. *Science* **353**, 1241–1244 (2016).
8. Meng, F. et al. Genomic editing of intronic enhancers unveils their role in fine-tuning tissue-specific gene expression in *Arabidopsis thaliana*. *Plant Cell* **33**, 1997–2014 (2021).
9. Reyna-Llorens, I. et al. Ancient duons may underpin spatial patterning of gene expression in C4 leaves. *Proc. Natl Acad. Sci. USA* **115**, 1931–1936 (2018).
10. Weirauch, M. T. et al. Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **158**, 1431–1443 (2014).
11. O'Malley, R. C. et al. Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell* **165**, 1280–1292 (2016).
12. Lambert, S. A. et al. Similarity regression predicts evolution of transcription factor sequence specificity. *Nat. Genet.* **51**, 981–989 (2019).
13. Hummel, N. F. C. et al. The *trans*-regulatory landscape of gene networks in plants. *Cell Syst.* **14**, 501–511.e4 (2023).
14. Rhoades, M. W. et al. Prediction of plant microRNA targets. *Cell* **110**, 513–520 (2002).
15. Hayden, C. A. & Jorgensen, R. A. Identification of novel conserved peptide uORF homology groups in *Arabidopsis* and rice reveals ancient eukaryotic origin of select groups and preferential association with transcription factor-encoding genes. *BMC Biol.* **5**, 32 (2007).
16. van der Horst, S., Filipovska, T., Hanson, J. & Smeekens, S. Metabolite control of translation by conserved peptide uORFs: the ribosome as a metabolite multisensor. *Plant Physiol.* **182**, 110–122 (2020).
17. Huynh-Thu, V. A., Irrthum, A., Wehenkel, L. & Geurts, P. Inferring regulatory networks from expression data using tree-based methods. *PLoS ONE* **5**, e12776 (2010).
18. Nolan, T. M. & Shahan, R. Resolving plant development in space and time with single-cell genomics. *Curr. Opin. Plant Biol.* **76**, 102444 (2023).
19. Gaudinier, A. & Brady, S. M. Mapping transcriptional networks in plants: data-driven discovery of novel biological mechanisms. *Annu. Rev. Plant Biol.* **67**, 575–594 (2016).
20. Li, M. et al. Double DAP-seq uncovered synergistic DNA binding of interacting bZIP transcription factors. *Nat. Commun.* **14**, 2600 (2023).
21. Pratapa, A., Jalihal, A. P., Law, J. N., Bharadwaj, A. & Murali, T. M. Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. *Nat. Methods* **17**, 147–154 (2020).
22. Staut, J., Pérez, N. M., Depuydt, T., Vandepoele, K. & Lukicheva, S. MINI-EX version 2: cell-type-specific gene regulatory network inference using an integrative single-cell transcriptomics approach. Preprint at *bioRxiv* https://doi.org/10.1101/2023.12.24.573246 (2023).
23. Cirrone, J., Brooks, M. D., Bonneau, R., Coruzzi, G. M. & Shasha, D. E. OutPredict: multiple datasets can improve prediction of expression and inference of causality. *Sci. Rep.* **10**, 6804 (2020).
24. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
25. Marand, A. P. & Schmitz, R. J. Single-cell analysis of *cis*-regulatory elements. *Curr. Opin. Plant Biol.* **65**, 102094 (2022).
26. Schmitz, R. J., Grotewold, E. & Stam, M. *Cis*-regulatory sequences in plants: their importance, discovery, and future challenges. *Plant Cell* **34**, 718–741 (2022).
27. Lu, Z. et al. The prevalence, evolution and chromatin signatures of plant regulatory elements. *Nat. Plants* **5**, 1250–1259 (2019).
28. Jin, R. et al. LEAFY is a pioneer transcription factor and licenses cell reprogramming to floral fate. *Nat. Commun.* **12**, 626 (2021).
29. Lai, X. et al. The LEAFY floral regulator displays pioneer transcription factor properties. *Mol. Plant* **14**, 829–837 (2021).
30. Zhao, Y. et al. 'Stripe' transcription factors provide accessibility to co-binding partners in mammalian genomes. *Mol. Cell* **82**, 3398–3411 (2022).
31. Mangan, S. & Alon, U. Structure and function of the feed-forward loop network motif. *Proc. Natl Acad. Sci. USA* **100**, 11980–11985 (2003).
32. Taylor-Teeples, M. et al. *An Arabidopsis* gene regulatory network for secondary cell wall synthesis. *Nature* **517**, 571–575 (2015).

33. Turco, G. M. et al. Molecular mechanisms driving switch behavior in xylem cell differentiation. *Cell Rep.* **28**, 342–351 (2019).
34. Neph, S. et al. Circuitry and dynamics of human transcription factor regulatory networks. *Cell* **150**, 1274–1286 (2012).
35. Alabadí, D. et al. Reciprocal regulation between *TOC1* and *LHY*/*CCA1* within the *Arabidopsis* circadian clock. *Science* **293**, 880–883 (2001).
36. Bonnot, T. & Nagel, D. H. Time of the day prioritizes the pool of translating mRNAs in response to heat stress. *Plant Cell* **33**, 2164–2182 (2021).
37. Varala, K. et al. Temporal transcriptional logic of dynamic regulatory networks underlying nitrogen signaling and use in plants. *Proc. Natl Acad. Sci. USA* **115**, 6494–6499 (2018).
38. Para, A. et al. Hit-and-run transcriptional control by bZIP1 mediates rapid nutrient signaling in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **111**, 10371–10376 (2014).
39. Clark, N. M. et al. Stem-cell-ubiquitous genes spatiotemporally coordinate division through regulation of stem-cell-specific gene networks. *Nat. Commun.* **10**, 5574 (2019).
40. Page, L., Brin, S., Motwani, R. & Winograd, T. *The PageRank Citation Ranking: Bringing Order to the Web* Technical Report (Stanford InfoLab, 1999).
41. Zhang, K., Wang, M., Zhao, Y. & Wang, W. Taiji: system-level identification of key transcription factors reveals transcriptional waves in mouse embryonic development. *Sci. Adv.* **5**, eaav3262 (2019).
42. Reynoso, M. A. et al. Gene regulatory networks shape developmental plasticity of root cell types under water extremes in rice. *Dev. Cell* **57**, 1177–1192 (2022).
43. Carroll, S. B. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**, 25–36 (2008).
44. Britten, R. J. & Davidson, E. H. Gene regulation for higher cells: a theory. *Science* **165**, 349–357 (1969).
45. Tu, X. et al. Reconstructing the maize leaf regulatory network using ChIP–seq data of 104 transcription factors. *Nat. Commun.* **11**, 5089 (2020).
46. Julca, I. et al. Comparative transcriptomic analysis reveals conserved programmes underpinning organogenesis and reproduction in land plants. *Nat. Plants* **7**, 1143–1159 (2021).
47. Leiboff, S. & Hake, S. Reconstructing the transcriptional ontogeny of maize and sorghum supports an inverse hourglass model of inflorescence development. *Curr. Biol.* **29**, 3410–3419 (2019).
48. Kajala, K. et al. Innovation, conservation, and repurposing of gene function in root cell type development. *Cell* **184**, 3333–3348 (2021).
49. Wu, S. et al. A plausible mechanism, based upon SHORT-ROOT movement, for regulating the number of cortex cell layers in roots. *Proc. Natl Acad. Sci. USA* **111**, 16184–16189 (2014).
50. Ortiz-Ramírez, C. et al. Ground tissue circuitry regulates organ complexity in maize and *Setaria*. *Science* **374**, 1247–1252 (2021).
51. Barberon, M. et al. Adaptation of root function by nutrient-induced plasticity of endodermal differentiation. *Cell* **164**, 447–459 (2016).
52. Salas-González, I. et al. Coordination between microbiota and root endodermis supports plant mineral nutrient homeostasis. *Science* **371**, eabd0695 (2021).
53. Shukla, V. et al. Suberin plasticity to developmental and exogenous cues is regulated by a set of MYB transcription factors. *Proc. Natl Acad. Sci. USA* **118**, e2101730118 (2021).
54. Cantó-Pastor, A. et al. A suberized exodermis is required for tomato drought tolerance. *Nat. Plants* **10**, 118–130 (2024).
55. Shiono, K. et al. RCN1/OsABCG5, an ATP-binding cassette (ABC) transporter, is required for hypodermal suberization of roots in rice (*Oryza sativa*). *Plant J.* **80**, 40–51 (2014).
56. Guillotin, B. et al. A pan-grass transcriptome reveals patterns of cellular divergence in crops. *Nature* **617**, 785–791 (2023).
57. Schiessl, K. et al. NODULE INCEPTION recruits the lateral root developmental program for symbiotic nodule organogenesis in *Medicago truncatula*. *Curr. Biol.* **29**, 3657–3668 (2019).
58. Soyano, T., Shimoda, Y., Kawaguchi, M. & Hayashi, M. A shared gene drives lateral root development and root nodule symbiosis pathways in *Lotus*. *Science* **366**, 1021–1023 (2019).
59. Okushima, Y., Fukaki, H., Onoda, M., Theologis, A. & Tasaka, M. ARF7 and ARF19 regulate lateral root formation via direct activation of *LBD*/*ASL* genes in *Arabidopsis*. *Plant Cell* **19**, 118–130 (2007).
60. Niklaus, M. & Kelly, S. The molecular evolution of C4 photosynthesis: opportunities for understanding and improving the world's most productive plants. *J. Exp. Bot.* **70**, 795–804 (2019).
61. Rossini, L., Cribb, L., Martin, D. J. & Langdale, J. A. The maize *golden2* gene defines a novel class of transcriptional regulators in plants. *Plant Cell* **13**, 1231–1244 (2001).
62. Lambret-Frotte, J., Smith, G. & Langdale, J. A. *GOLDEN2-LIKE1* is sufficient but not necessary for chloroplast biogenesis in mesophyll cells of C4 grasses. *Plant J.* **117**, 416–431 (2024).
63. Tu, X. et al. Limited conservation in cross-species comparison of GLK transcription factor binding suggested wide-spread cistrome divergence. *Nat. Commun.* **13**, 7632 (2022).
64. Mendieta, J. P. et al. Investigating the *cis*-regulatory basis of C3 and C4 photosynthesis in grasses at single-cell resolution. Preprint at *bioRxiv* https://doi.org/10.1101/2024.01.05.574340 (2024).
65. Swift, J. et al. Single nuclei sequencing reveals C4 photosynthesis is based on rewiring of ancestral cell identity networks. Preprint at *bioRxiv* https://doi.org/10.1101/2023.10.26.562893 (2023).
66. Gasch, P. et al. Redundant ERF-VII transcription factors bind to an evolutionarily conserved *cis*-motif to regulate hypoxia-responsive gene expression in *Arabidopsis*. *Plant Cell* **28**, 160–180 (2016).
67. Reynoso, M. A. et al. Evolutionary flexibility in flooding response circuitry in angiosperms. *Science* **365**, 1291–1295 (2019).
68. Sun, Y. et al. Divergence in the ABA gene regulatory network underlies differential growth control. *Nat. Plants* **8**, 549–560 (2022).
69. González, F. G. et al. Field-grown transgenic wheat expressing the sunflower gene *HaHB4* significantly outyields the wild type. *J. Exp. Bot.* **70**, 1669–1681 (2019).
70. Wu, J. et al. Overexpression of *zmm28* increases maize grain yield in the field. *Proc. Natl Acad. Sci. USA* **116**, 23850–23858 (2019).
71. Wei, S. et al. A transcriptional regulator that boosts grain yields and shortens the growth duration of rice. *Science* **377**, eabi8455 (2022).
72. Butelli, E. et al. Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nat. Biotechnol.* **26**, 1301–1308 (2008).
73. Patron, N. J. et al. Standards for plant synthetic biology: a common syntax for exchange of DNA parts. *New Phytol.* **208**, 13–19 (2015).
74. Haseloff, J. GFP variants for multispectral imaging of living cells. In *Methods in Cell Biology* Vol. 58 (eds Sullivan, K. F. & Kay, S. A.) Ch. 9, 139–151 (Academic Press, 1998).
75. Belcher, M. S. et al. Design of orthogonal regulatory systems for modulating gene expression in plants. *Nat. Chem. Biol.* **16**, 857–865 (2020).
76. Schaumberg, K. A. et al. Quantitative characterization of genetic parts and circuits for plant synthetic biology. *Nat. Methods* **13**, 94–100 (2016).
77. Lloyd, J. P. B. et al. Synthetic memory circuits for stable cell reprogramming in plants. *Nat. Biotechnol.* **40**, 1862–1872 (2022).

78. Guiziou, S., Maranas, C. J., Chu, J. C. & Nemhauser, J. L. An integrase toolbox to record gene-expression during plant development. *Nat. Commun.* **14**, 1844 (2023).
79. Brophy, J. A. N. et al. Synthetic genetic circuits as a means of reprogramming plant roots. *Science* **377**, 747–751 (2022).
80. Fukaki, H., Tameda, S., Masuda, H. & Tasaka, M. Lateral root formation is blocked by a gain-of-function mutation in the *SOLITARY-ROOT/IAA14* gene of *Arabidopsis*. *Plant J.* **29**, 153–168 (2002).
81. Pan, C. et al. CRISPR–Act3.0 for highly efficient multiplexed gene activation in plants. *Nat. Plants* **7**, 942–953 (2021).
82. Ulmasov, T., Murfett, J., Hagen, G. & Guilfoyle, T. J. Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant Cell* **9**, 1963–1971 (1997).
83. Wu, R. et al. The *6xABRE* synthetic promoter enables the spatiotemporal analysis of ABA-mediated transcriptional regulation. *Plant Physiol.* **177**, 1650–1665 (2018).
84. Benn, G. et al. A key general stress response motif is regulated non-uniformly by CAMTA transcription factors. *Plant J.* **80**, 82–92 (2014).
85. Cai, Y.-M. et al. Rational design of minimal synthetic promoters for plants. *Nucleic Acids Res.* **48**, 11845–11856 (2020).
86. Klein, J. C. et al. A systematic evaluation of the design and context dependencies of massively parallel reporter assays. *Nat. Methods* **17**, 1083–1091 (2020).
87. Jores, T. et al. Identification of plant enhancers and their constituent elements by STARR-seq in tobacco leaves. *Plant Cell* **32**, 2120–2131 (2020).
88. Jores, T. et al. Synthetic promoter designs enabled by a comprehensive analysis of plant core promoters. *Nat. Plants* **7**, 842–855 (2021).
89. de Almeida, B. P. et al. Targeted design of synthetic enhancers for selected tissues in the *Drosophila* embryo. *Nature* **626**, 207–211 (2023).
90. Taskiran, I.I., Spanier, K.I., Dickmänken, H. et al. Cell-type-directed design of synthetic enhancers. *Nature* **626**, 212–220 (2024).
91. Somssich, M., Je, B. I., Simon, R. & Jackson, D. CLAVATA–WUSCHEL signaling in the shoot meristem. *Development* **143**, 3238–3248 (2016).
92. Wang, X. et al. Dissecting *cis*-regulatory control of quantitative trait variation in a plant stem cell circuit. *Nat. Plants* **7**, 419–427 (2021).
93. Rodríguez-Leal, D., Lemmon, Z. H., Man, J., Bartlett, M. E. & Lippman, Z. B. Engineering quantitative trait variation for crop improvement by genome editing. *Cell* **171**, 470–480 (2017).
94. Song, X. et al. Targeting a gene regulatory element enhances rice grain yield by decoupling panicle number and size. *Nat. Biotechnol.* **40**, 1403–1411 (2022).
95. Jiao, Y. et al. Regulation of *OsSPL14* by OsmiR156 defines ideal plant architecture in rice. *Nat. Genet.* **42**, 541–544 (2010).
96. Miura, K. et al. *OsSPL14* promotes panicle branching and higher grain productivity in rice. *Nat. Genet.* **42**, 545–549 (2010).
97. Zhou, J. et al. An efficient CRISPR–Cas12a promoter editing system for crop improvement. *Nat. Plants* **9**, 588–604 (2023).
98. Wang, Z. et al. A reverse chromatin immunoprecipitation technique based on the CRISPR–dCas9 system. *Plant Physiol.* **191**, 1505–1519 (2023).
99. Stonedahl, F. & Wilensky, U. *NetLogo PageRank Model* ccl.northwestern.edu/netlogo/models/PageRank (2009).

Acknowledgements

We thank G. Akmakjian, C. Sepulveda, A. Morris and S. Borowsky for helpful suggestions. Our research is supported by the US National Science Foundation (IOS-211980 to J.B.-S. and DGE-1922642 to J.B.-S. and A.T.B.) and the US Department of Agriculture, US National Institute of Food and Agriculture—Agriculture and Food Research Initiative grants 2019-67013-29313 (to J.B.-S.) and 1026477 (to A.T.B.).

Author contributions

A.T.B. and J.B.-S. conceptualized the paper. A.T.B. wrote the original draft. A.T.B. and J.B.-S. revised and edited the paper.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Julia Bailey-Serres.

Peer review information *Nature Genetics* thanks Yiping Qi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.