



## OPEN Regulation of meristem and hormone function revealed through analysis of directly-regulated SHOOT MERISTEMLESS target genes

Tamara Lechon, Nicholas A. Kent, James A. H. Murray & Simon Scofield✉

The Arabidopsis *Knotted1*-like homeobox (*KNOX*) gene *SHOOT MERISTEMLESS* (*STM*) encodes a homeodomain transcription factor that operates as a central component of the gene regulatory network (GRN) controlling shoot apical meristem formation and maintenance. It regulates the expression of target genes that include transcriptional regulators associated with meristem function, particularly those involved in pluripotency and cellular differentiation, as well as genes involved in hormone metabolism and signaling. Previous studies have identified *KNOX*-regulated genes and their associated *cis*-regulatory elements in several plant species. However, little is known about *STM*-DNA interactions in the regulatory regions of target genes in Arabidopsis. Here, we identify and map *STM* binding sites in the Arabidopsis genome using global ChIP-seq analysis to reveal potential directly-regulated *STM* target genes. We show that in the majority of target loci, *STM* binds within 1 kb upstream of the TSS, with other loci showing *STM* binding at more distal enhancer sites, and we reveal enrichment of DNA motifs containing a TGAC and/or TGAT core in *STM*-bound target gene *cis*-regulatory elements. We further demonstrate that many *STM*-bound genes are transcriptionally responsive to altered levels of *STM* activity, and show that among these, transcriptional regulators with key roles in meristem and hormone function are highly represented. Finally, we use a subset of these target genes to perform Bayesian network analysis to infer gene regulatory associations and to construct a refined GRN for *STM*-mediated control of meristem function.

Shoot growth in higher plants is an iterative process that relies on populations of undifferentiated, pluripotent cells located in specialised stem cell-containing structures termed meristems to enable continued production of lateral organs such as leaves and flowers throughout the plant life cycle<sup>1,2</sup>. Cell fate decisions in the SAM are controlled by transcription factors and other regulatory proteins that operate in gene regulatory networks (GRNs) to maintain pluripotency and stem cell homeostasis via repression of cell differentiation, or to initiate lateral organ development through the repression of pluripotency and activation of organ-specific differentiation programmes in incipient lateral organ primordia<sup>3–5</sup>.

Class-1 *Knotted1*-like homeobox (*KNOX*) genes such as *KNOTTED1* (*KN1*) from maize and its ortholog *SHOOT MERISTEMLESS* (*STM*) from Arabidopsis encode TALE homeodomain transcription factors that are essential for conferring pluripotency to cells in the SAM<sup>6,7</sup>. *KNOX* genes are expressed throughout the SAM, including the central zone of stem cells and the surrounding peripheral zone of transit amplifying cells, but are transcriptionally repressed in incipient organ primordia and developing organs by multiple factors including the transcriptional repressor AS1, the LOB-domain protein AS2, members of the YABBY transcription factor family, the Polycomb Repressive Complex (PRC) and the class-2 TCP transcriptional regulators<sup>8–17</sup>. *STM* has previously been shown to be essential for embryonic SAM formation and continued SAM maintenance throughout the plant life-cycle, with loss-of-function mutants either failing to develop a SAM during embryogenesis, or developing defective SAMs that exhibit stem cell depletion caused by inappropriate organogenesis from the central zone of stem cells<sup>18,19</sup>. Conversely, ectopic expression of *STM* inhibits proper cellular differentiation in leaves, leading to leaf lobing and to *de novo* shoot meristem formation on the adaxial leaf surface<sup>20–22</sup>. This highlights the need for repression of *KNOX* gene expression in simple leaf species such as Arabidopsis, though reactivation of *KNOX*

School of Biosciences, Cardiff University, Cardiff CF10 3AX, UK. ✉email: ScofieldS@cardiff.ac.uk

gene expression is required for leaflet formation in plant species with compound leaves<sup>23,24</sup>. Taken together, these observations suggest a critical role for STM in the control of pluripotency associated with SAM function.

Several studies have sought to understand the *KNOX* gene regulatory network by identifying *KNOX*-responsive target genes and examining the *KNOX*-DNA interactions at target gene promoters. For example, in maize and rice the principal target genes of the *KNOX* proteins KN1 and OSH1 were identified as being involved in phytohormone biosynthesis, signalling and response, especially relating to auxin and brassinosteroids<sup>25–27</sup>. In *Arabidopsis* and barley, *STM* or its ortholog *BKN3* have also been implicated in the regulation of auxin biosynthesis, transport and response, with several auxin-associated factors showing differential expression following ectopic expression<sup>28,29</sup>. Auxin accumulation in incipient organ primordia is coincident with down-regulation of *STM* expression, and auxin response factors (ARFs) promote histone deacetylation leading to transcriptional repression of *STM* in incipient floral primordia<sup>30</sup>, suggesting antagonism between auxin and *KNOX* gene function.

*KNOX* genes have also been shown to repress the biosynthesis of gibberellic acid (GA), accumulation of which is detrimental to sustained SAM function, through direct transcriptional repression of genes encoding GA biosynthetic enzymes<sup>31–35</sup>. Furthermore, in *Arabidopsis*, *STM* has been shown to promote cytokinin biosynthesis by activating expression of several members of the *ISOPENTYL TRANSFERASE (IPT)* gene family<sup>32,36</sup>. Cytokinin is required for proper SAM function and promotes cell division through the Cyclin D pathway<sup>32,36–38</sup>. Hence, *KNOX* proteins impinge on several phytohormone pathways that affect cell division, cell expansion and cell differentiation processes associated with SAM development/maintenance and lateral organ formation.

In *Arabidopsis*, *STM* was shown to regulate the expression of many transcription factors associated with meristem development and the control of pluripotency<sup>29</sup>. These included the class-1 *KNOX* genes *KNAT1/BP* and *KNAT2*, the AP2 family gene *PLT7/AIL7*, which promotes pluripotency and the regulation of phyllotactic patterning<sup>39–41</sup>, *CUC1* and *BOP2*, which are associated with meristem-organ boundary establishment, regulation of *KNOX* gene expression and organ polarity<sup>42–49</sup> and *HB25* which promotes shoot identity<sup>50</sup>. *STM* was also shown to repress expression of the class-2 TCP transcription factor-encoding genes *TCP3*, *TCP4* and *TCP10*, which promote cellular differentiation and themselves repress *KNOX* gene expression during leaf development and antagonise meristem function when ectopically expressed<sup>15,51–54</sup>, revealing a functional mechanism for the inhibition of cellular differentiation associated with leaf formation by *KNOX* transcriptional regulators.

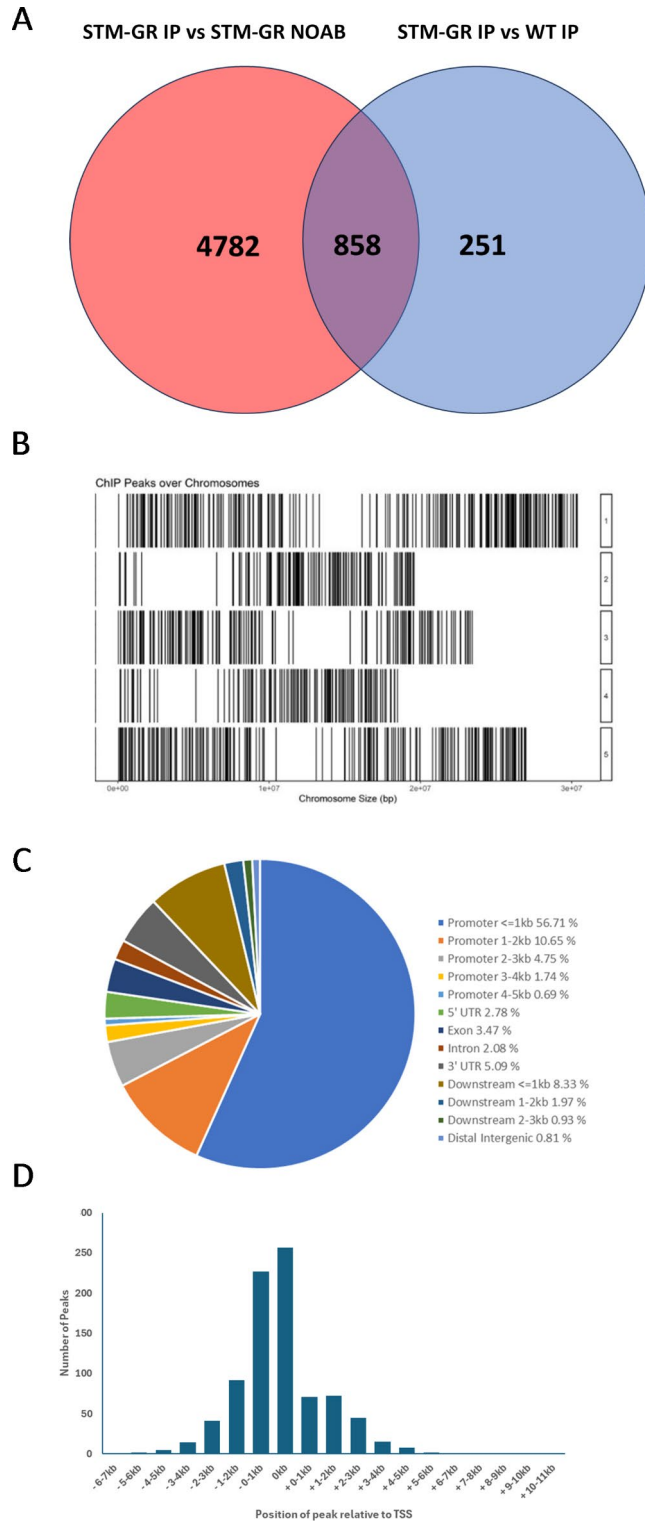
In this study, we comprehensively map the *STM* binding sites in the *Arabidopsis* genome using chromatin immunoprecipitation of *STM*-bound DNA followed by whole genome next-generation DNA sequencing (ChIP-seq). We reveal distinct *STM* binding sites in target gene regulatory regions, with positions ranging from proximal promoter to distal regions several kilobases upstream or downstream of the transcriptional start site (TSS). We identify several putative *STM*-bound *cis*-regulatory motifs, all of which contain the TGAC or TGAT cores previously shown to be required for the binding of homeodomain proteins to DNA. We then correlate *STM* binding sites with genes previously identified as being transcriptionally responsive to *STM*, revealing direct regulation of many genes that are involved in SAM development. To investigate *STM* GRN topology, we perform Bayesian network analysis to infer regulatory relationships among the *STM* target genes that encode transcriptional regulators. Together our results provide key information on the *STM*-DNA interactions that are associated with the promotion of pluripotency and the repression of differentiation in the SAM of *Arabidopsis*.

## Results

### Genome-wide identification of putative *STM*-binding sites by ChIP-seq

*KNOX* homeodomain transcriptional regulators typically bind to *cis*-acting DNA elements containing a TGAC core motif, often located in the promoter region upstream of the target locus coding sequence<sup>25</sup>. Previous studies have shown that *STM* binds to a *cis*-regulatory element ~200 bp upstream from the translation initiation codon of the *CUC1* gene, which encodes a NAC domain transcription factor that is essential for activation of *STM* expression during embryogenesis and for the proper establishment of meristem-organ boundary zones<sup>29,42</sup>. Other *STM* targets such as in *BOP2* or *ATHB25* show *STM* binding at positions further upstream of the promoter region, while other *KNOX* proteins bind to *cis*-regulatory elements in the gene body, such as in the *GA20ox1* gene in maize<sup>29,55</sup>. Hence, *STM* binding elements can occur at a range of distances from the transcriptional start site, necessitating a global NGS driven ChIP-seq approach to discover *STM* binding sites rather than targeted ChIP-PCR analysis based on *in silico* predictions.

To perform a global identification of the DNA binding sites for *STM* in the *Arabidopsis* genome, we performed a ChIP-seq experiment using a fusion of the *STM* coding sequence to the rat glucocorticoid receptor (GR), placed under control of the constitutive CaMV 35S promoter. CaMV 35S:*STM-GR* (*STM-GR*) plants were treated with the glucocorticoid analogue dexamethasone (DEX; allowing nuclear import of the *STM-GR* fusion protein) and immunoprecipitated with the anti-GR epitope antibody PA-516. We compared DEX-treated *STM-GR* anti-GR immunoprecipitated samples (*STM-GR* IP) with mock-immunoprecipitated no-antibody (NOAB) samples from the same DEX-treated *STM-GR* line (*STM-GR* NOAB), and with DEX-treated WT immunoprecipitated and mock-immunoprecipitated no-antibody samples (WT IP and WT NOAB) which served as no-GR epitope controls. By performing these two types of comparison we were better able to identify only the most robust putative *STM*-binding sites in the genome which were used exclusively for further analysis. This yielded a total of 858 ChIP-seq peaks, containing putative *STM* binding sites, distributed across all five chromosomes (Fig. 1A, B), which were then assigned to the closest genomic locus. The 858 peaks were associated with 859 loci, as some peaks were potentially associated with 2 loci, for example when the peak was located in the downstream intergenic region approximately equidistant between two loci (6 in total). Of the 859 loci, 854 had a single associated peak containing at least one putative *STM*-binding site, while five had two distinct peaks at different positions (Supplementary Table S1).



**Fig. 1.** ChIP-seq analysis of putative STM-binding sites in the *Arabidopsis thaliana* genome. **(A)** Venn diagram showing the overlap between the number of ChIP-seq peaks identified by comparison of the 35S::STM-GR sample immunoprecipitated with the anti-GR antibody (STM IP) compared to the 35S::STM-GR sample mock-immunoprecipitated with no antibody (STM NoAB) and the 35S::STM-GR sample immunoprecipitated with the anti-GR antibody compared to a WT sample also immunoprecipitated with the anti-GR antibody. **(B)** Distribution of ChIP-seq peaks throughout the five chromosomes of the Arabidopsis genome. Chromosome number is shown in boxes on the righthand side, chromosome length extends from left to right. Black vertical lines indicate positions of ChIP-seq peaks. **(C)** Location of ChIP-seq peaks in relation to gene model for closest genes to ChIP-seq peaks in the Arabidopsis genome. **(D)** Histogram showing the distribution of peak positions relative to the transcriptional start site (TSS) of STM-bound genes. The full list of ChIP-seq peaks, positions and associated genes is shown in Supplementary Table S1.

The ChIP-seq results show that the majority of the 858 peaks (56.7%) lie within 1 kb upstream of the transcription start site (TSS) of target loci (Fig. 1C, D). Of these, 257 peaks (~30%) directly overlapped the TSS. Fewer peaks (10.7%) were located between 1 kb and 2 kb upstream of the TSS, while 7.2% of peaks located between 2 kb and 5 kb upstream of the TSS. Furthermore, 13.4% of peaks were identified within the gene body (5' and 3' UTRs, exons and introns), while 11.2% of peaks were located downstream of target loci to up to 3 kb distal from the 3' end of the gene body, and 0.8% were located in distal intergenic regions. These results show that the majority of putative STM-binding sites lie in the promoter region at, or upstream of, the TSS, consistent with previous analyses of KNOX proteins binding sites<sup>25,26,29</sup>.

### Identification of putative consensus *cis*-regulatory DNA motifs in ChIP-seq peaks

We next sought to identify putative *cis*-regulatory motifs in the 858 peaks identified in the ChIP-seq analysis. KNOTTED1-like homeodomain proteins have previously been shown to bind to *cis*-acting DNA elements containing a core TGAC motif<sup>33,34,42,55–58</sup>, while other homeodomain proteins such as ANTENNAPEDIA and WUSCHEL have shown a binding preference for TAAT or TGAA<sup>59–64</sup>.

We performed a consensus DNA-binding motif analysis using a matrix-based hidden Markov model on the sequences of the 858 peaks identified by ChIP-seq analysis. This generated three distinct consensus motifs containing one or more TGAC cores, one motif containing a TGAT dyad core and one motif containing both TGAC and TGAT cores (Fig. 2; Supplementary Data 1). The latter motif was represented in the majority (691) of the 858 ChIP-seq peaks, with others being present in ~200–250 genes each (Supplementary Table S2). Some ChIP-seq peaks contained more than one consensus motif. We compared the abundance of these motifs with random genomic sequences from Arabidopsis and synthetic DNA sequences based on the Arabidopsis genome nucleotide composition, and found significantly higher occurrence of the motifs in the ChIP-seq peaks compared to these control sequences, suggesting that these motifs are significantly overrepresented in the ChIP-seq dataset. Overall, this strongly suggests that the core STM binding motif is similar to that of other KNOX homeodomain transcription factors. However, we could not detect any of the five motifs in some of the genes identified (58) in the ChIP-seq analysis, including well-characterised STM target genes such as *CUC1*. This was attributable to our stringent *p*-value (<1e-5) for binding site prediction, with motifs present in the *CUC1* promoter failing to reach the required threshold for significance. Hence our approach identified a number of putative *cis*-regulatory elements but was unable to capture all biologically relevant motifs in target loci.

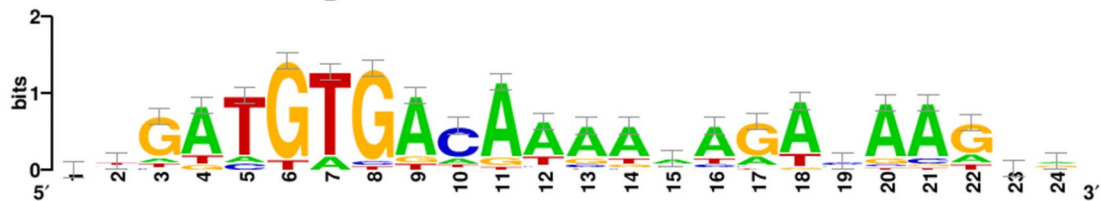
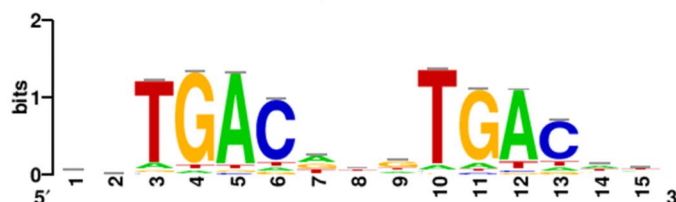
### Gene ontology enrichment analysis of genes containing putative STM-binding sites

To determine the functions of the putative target genes identified, we performed gene ontology (GO) enrichment analysis on the 859 genes identified as having putative STM-binding sites by ChIP-seq (Fig. 3 and Supplementary Table S3). Enriched GO categories for biological process (BP) included many related to developmental processes including shoot system development, pattern specification, flower and root development, post-embryonic plant morphogenesis, hormone synthesis and hormone-mediated signalling (including auxin-mediated signaling associated with phyllotaxis). Other biological processes included transcription, cell communication and signaling, biosynthetic processes and responses to various endogenous and exogenous stimuli. Enriched molecular function (MF) categories included DNA-binding transcription factor activity and protein binding, while cellular component (CC) enrichment was shown for the nucleus, plasma membrane and cell periphery. These enriched GO categories are broadly in line with those identified in previous studies for the functions of KNOX target genes<sup>25,26,29</sup>, and suggest that transcriptional regulators related to development comprise a key category of STM-regulated target genes.

### Correlation of genes containing putative STM-binding sites with STM transcriptionally-responsive target genes

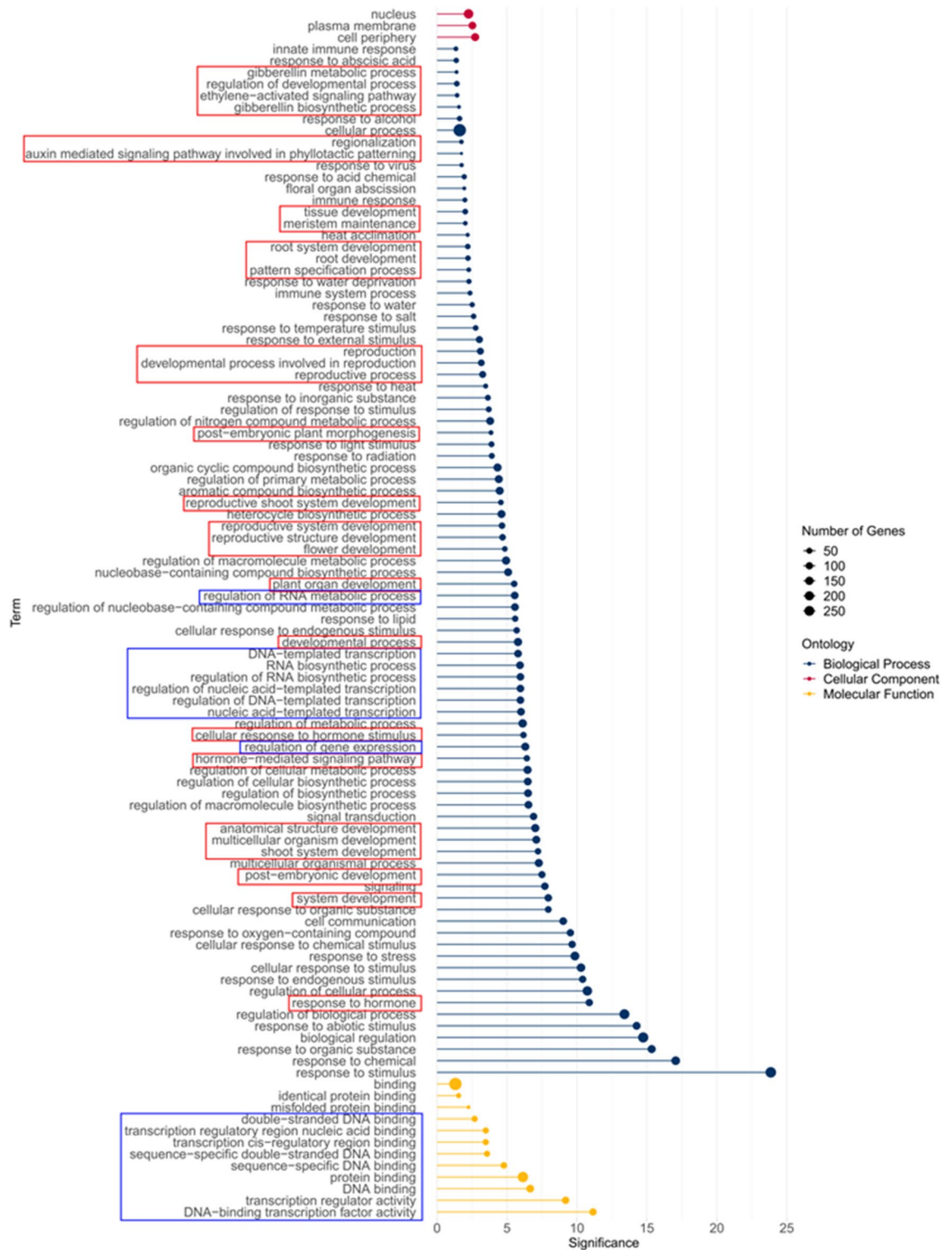
The data presented above show which loci are bound by STM, and include many known genes associated with meristem development (primarily transcription factors), meristem-organ boundary specification, lateral organ development, and the biosynthesis and signaling of the plant hormones auxin, cytokinin and GA (Supplementary Table S1). However, this analysis does not reveal which of these genes are also transcriptionally responsive to STM, and thus represent genuine 'functional' STM target genes. Indeed, several studies have shown that many TFs have a multitude of genomic binding sites that are not associated with transcriptional regulation - so-called non-functional binding sites<sup>65–67</sup>. We therefore compared our ChIP-seq data with published transcriptomics data that identified genes that show differential expression (i.e. are either up- or down-regulated) to altered levels of STM activity.

We previously identified STM-responsive genes in a time-course of *STM* expression using the DEX-inducible TGV system followed by mRNA differential expression analysis and global meta-analysis of the differentially expressed genes in inducible *STM*-overexpression and RNAi experiments, and we identified putative directly-regulated *STM* target genes using the DEX-inducible *STM*-GR line induced with DEX alone or with DEX and the protein synthesis inhibitor cycloheximide, followed by differential gene expression analysis<sup>29</sup>. We compared the loci identified in our ChIP-seq experiment with genes previously identified in the *STM* overexpression time-course/ meta-analysis and in the direct target experiment (Supplementary Table S4). Overall, we found that 432/859 genes identified in the ChIP-seq analysis showed a transcriptional response to altered levels of *STM* activity, with 428 of these identified in the time-course/ meta-analysis and an additional four genes identified only in the direct-target experiment. We classified genes as 'early-responding' if they initially showed differential expression in the 3 h direct-target experiment, the 8 h or 24 timepoints in the timecourse or the meta-analysis (74 genes; Supplementary Table S5), while genes that were shown to be differentially expressed only in the 72 h or 9-day timepoints were classified as 'late-responding' (358 genes; Supplementary Table S6). At these later timepoints, the phenotypic consequences of ectopic *STM* expression, such as impaired leaf differentiation, were

**Motif 1 – 691 genes****Motif 2 – 249 genes****Motif 3 – 246 genes****Motif 4 – 223 genes****Motif 5 – 205 genes**

**Fig. 2.** Identification of putative *cis*-regulatory motifs in STM target gene promoters. Motifs were identified by consensus DNA-binding motif analysis using a matrix-based hidden Markov model using the sequences of the 858 peaks as the input data. The number of genes containing each motif in their respective ChIP-seq peak is shown. The full list of genes associated with each motif is shown in Supplementary Table S2.





**Fig. 3.** Gene Ontology (GO) Enrichment Analysis of 859 genes identified in the ChIP-seq analysis. GO terms for Cellular Component, Biological Process and Molecular Function are shown. Terms associated with development and hormones are highlighted in red boxes, while genes associated with transcription are highlighted in blue boxes. The full list of enriched GO terms is shown in Supplementary Table S3.

apparent in newly emerging leaves (72 h) or throughout all the aerial tissues (9D;<sup>29</sup>). The ChIP-seq peaks from the STM-IP samples for a subset of target genes are shown in Fig. 4, with additional data from control samples in Supplementary Figure S1.

We once again performed GO enrichment analysis to investigate the functions of the genes identified in the early- and late-responding datasets. For the early-responding genes, most enriched GO terms related to development and transcriptional regulation, with BP terms such as specification of adaxial/abaxial axis, auxin-mediated signaling of phyllotactic patterning, maintenance of shoot apical meristem identity, post-embryonic morphogenesis and MF terms such as DNA-binding transcription factor activity showing enrichment, along with nucleus as the only enriched CC term (Supplementary Figure S2; Supplementary Table S7). For the late-responding genes, the MF terms protein-binding and catalytic activity were enriched, and enriched BP terms included response to diverse endogenous and exogenous stimuli including hormones and signaling, including hormone-mediated signaling. Enriched CC terms were the plasma membrane and cell periphery (Supplementary Figure S3; Supplementary Table S8). It is clear from these analyses that early-responding genes containing a putative STM-binding site were the most directly-related to shoot apical meristem development, primarily functioning at the level of transcriptional regulation, while the late-responding genes were involved in a more diverse array of functions connected with downstream processes such as cellular responses to hormones and other stimuli.

Given the enriched functions identified in the GO analysis, we next examined our ChIP and transcriptomics datasets to identify genes with known roles in meristem or lateral organ development, particularly those involved in the control of cellular differentiation and pluripotency or the regulation of plant hormone function. We found numerous genes with established roles in SAM development that were both bound by STM and showed an early transcriptional response to STM, with most of these showing sustained differential expression throughout the later stages of the time-course (Supplementary Table S5). Many of these encode transcription factors and include the meristem-associated *KNOX* genes *KNAT1/BP* and *KNAT2*, organ boundary-associated genes *CUP-SHAPED COTYLEDON1 (CUC1)* and *LIGHT-SENSITIVE HYPOCOTYLS1 (LSH1)*, *LSH3*, *LSH4* and *LSH5*, and the *AP2*-related genes *AIL7/PLT7* and *AIL6/PLT3* and *RAP2.6L*. TF encoding genes associated with lateral organ development included the MYB family gene *ASYMMETRIC LEAVES1 (AS1)* and the class-2 CIN-TCP gene *TCP4*, together with *BLADE-ON-PETIOLE1* and *BLADE-ON-PETIOLE2 (BOP1 and BOP2)*. Other TF encoding genes associated with shoot development included the homeobox genes *HB2*, *HB25* and *HB32*, *MYB17/LMI2*, the *ASYMMETRIC LEAVES2*-related gene *ASL9*, and the growth regulating factor gene *GRF7*. Furthermore, we identified two genes involved in chromatin modification: *SUVRI*, which encodes a histone lysine methyltransferase, and *CHR40/CLASSY4*, which encodes an SNF2-domain protein that is involved in tissue-specific DNA methylation. In addition to transcriptional regulators, we also identified several genes encoding enzymes involved in auxin and cytokinin metabolism. These included the *MES9* and *MES10* genes, which encode methyl indole-3-acetic acid esterases, and the adenine phosphoribosyl transferase *APT2* and the cytokinin riboside 5'-monophosphate phosphoribohydrolase *LONELY GUY8 (LOG8)*.

A number of genes bound by STM showed a transcriptional response to STM only at the later timepoints in the time-course (Supplementary Table S6), and many of these were associated with plant hormone function, such as the auxin response factors *ARF4*, *ARF6* and *ARF10*, the auxin conjugating enzyme *GH3.3*, the *PINOID* kinase and the carboxylesterase *CXE6*, which has potential methyl indole-3 acetic acid esterase activity. In addition, two other adenine phosphoribosyl transferase enzymes, *APT3* and *APT4*, were detected in the late-responding dataset, together with the cytokinin response factor *CRF2* and the gibberellin associated genes *GA20OX3*, *GASA1* and *GID1B*.

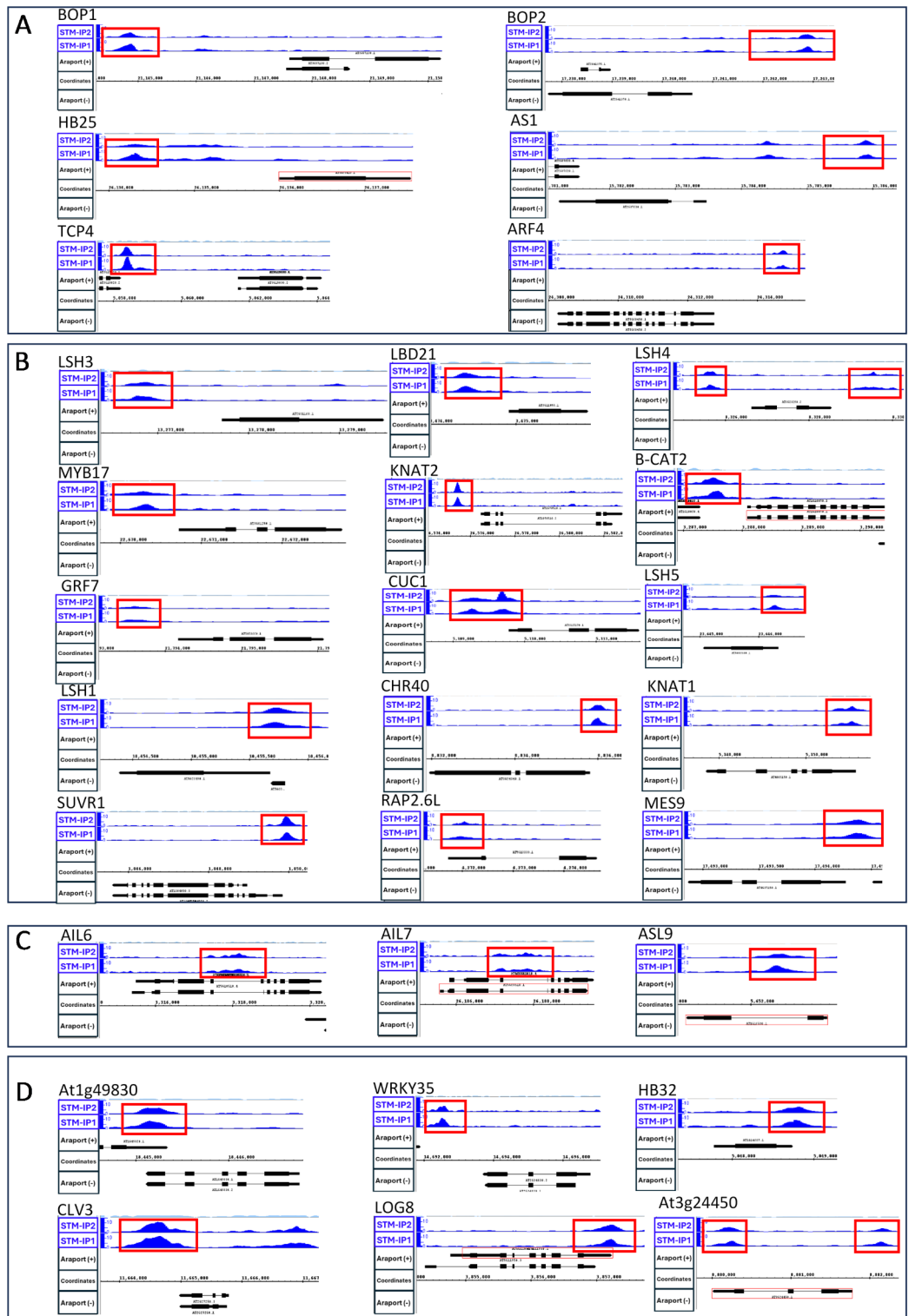
Previous studies have also linked *KNOX* gene function, particularly that of *KNAT1/BP*, to cell wall metabolism<sup>68–71</sup>, especially secondary cell wall formation and lignification. In agreement with this, we identified several cell wall-associated genes in the late-responding dataset, including cellulose biosynthesis genes *CSLA11*, *CESA5*, *CSLG2*, xyloglucan metabolism genes *XTH24* and *XTH31* and pectin methylesterases *PME1*, *PME3* and *PMEPCRf*. The cellulose synthase *CSL11A* and pectin methylesterase *PMEPCRf* were also identified as early-responding target genes.

Finally, in the late-responding dataset we also identified the cell cycle regulator *CDKB2;1*, which has a role in shoot meristem development<sup>72</sup>, and the meristem-associated transcription factors *HBI*, *ATH1* and *BELL-LIKE HOMEODOMAIN4/SAWTOOTH2 (BLH4/SAW2)*, the latter two of which encode homeodomain proteins that physically interact with *KNOX* proteins including STM<sup>73</sup>.

Overall these results indicate that STM directly regulates a number of important factors in the control of development and hormone function. Many other genes potentially involved in meristem development, organ development or hormone function were identified in the ChIP-seq analysis yet did not show a transcriptional response to STM according to our transcriptomics analysis, or for which no transcriptional response data were available.

### STM gene regulatory network construction using bayesian network inference

Our analysis has revealed that genes encoding transcriptional factors (TFs) are significantly enriched among the STM-bound and STM-responsive target genes, suggesting that STM functions largely through controlling the expression of other transcriptional regulators. To explore the potential regulatory interactions among the STM-regulated transcriptional regulators and reveal the topology of the STM GRN, we performed Bayesian network analysis<sup>74,75</sup> using the early responding STM-bound genes that encode transcriptional regulators as network components (nodes) and discretised data from ~2000 publicly available gene expression datasets to infer conditional dependency relationships to generate network edges. Though all these genes are likely to be directly regulated by STM, this approach allowed us to infer potential novel regulatory relationships and interplay between the various target genes and capture the potential functional sequence of these factors in



**Fig. 4.** ChIP-seq peaks for selected STM target genes. (A) Genes with STM binding site substantially upstream of target locus TSS. (B) Genes with STM binding site upstream or around the target gene TSS. (C) Genes with STM binding site within middle of gene body. (D) Genes with STM binding site at distal end or downstream of target gene. The full data tracks with all controls are shown in Supplementary Figure S1.

regulating meristem function, as determined by the degree of separation (i.e. the number of intermediate nodes) from STM in the network. A total of 33 TF-encoding genes were selected for network construction (including STM; Supplementary Table S9), and the transcriptional response to STM was confirmed by qRT-PCR for a subset of these genes (Supplementary Figure S4), with most showing differential expression at short- and/or



long-term timepoints following STM induction. Bayesian analysis was performed (50 iterations) to predict the frequency of conditional dependency relationships, indicated as edges between nodes (in both directions), and different frequency thresholds were explored to generate the consensus network, selecting > 30% confidence as the optimum threshold (i.e. conditional dependency relationships between two nodes being predicted in 30–100% of the 50 network iterations). Note that some genes formed mutual dependence relationships, which suggests feedback regulation. In these cases, the combined dependence frequencies for both directions should be considered.

The consensus network (Fig. 5) was able to capture several known regulatory relationships among the component genes, in the form of parent-child relationships, giving confidence to this approach. For example, *STM* and *TCP4* show mutual dependence, in agreement with the mutual transcriptional repression between these two genes<sup>15,29</sup>. Furthermore, *AS1* was shown to act as a parent to *KNAT1/BP*, which recapitulates the known transcriptional repression of *KNAT1/BP* by *AS1*<sup>8–10</sup>. Several members of the *LSH* gene family were connected to *CUC1* and *STM*, in agreement with their known role in comprising a meristem-organ boundary regulatory module<sup>76</sup>, together with *BOP2* which is also associated with regulation of *KNAT1/BP* and other *KNOX* genes<sup>46,48</sup>.

Our analysis reveals several genes that act as highly connected nodes, or hubs, within the network. As expected, *STM* itself is one of the most highly connected nodes, and shows direct connection to the boundary genes *LSH1*, *LSH4* and *LSH5*, which share potential feedback regulation with one another and themselves connect to *CUC1*, and several other genes involved in pluripotency, shoot formation and organogenesis (*AIL7*, *MYB17*, *BOP1*, *GRF7*, *HB32*, *HB25* and *TCP4*). *HB25*, a gene known to function in the promotion of shoot regeneration and development<sup>50</sup>, connects directly to several genes involved in multiple aspects of meristem function, such as pluripotency (*STM*, *AIL6*, *MYB17*, *RAP2.6L*), boundary formation (*CUC1* and *BOP2*) and organogenesis (*TCP4* and *AS1*). Moreover, *HB25* displays potential feedback relationships with *STM* and *MYB17*. Another key regulatory hub is *LBD21*, a member of the LOB-domain gene family which has no characterised function to date. *LBD21* connects directly to, and displays potential feedback with *TCP4*, and is also connected to *HB25*, *LSH1* and *LSH5*, *AIL6*, *RAP2.6L* and *GRF7*, which itself connects *STM* to *KNAT1/BP* and *KNAT2*. However, *LBD21* only showed a moderate transcriptional response to *STM*, so its importance must be treated with caution. Furthermore, the *AIL7* gene, which promotes pluripotency and controls phyllotactic patterning<sup>39–41</sup>, is also a highly-connected node in the network, with direct connections to *AS1*, the chromatin remodeller *CHR40* and the boundary gene *LSH3*, among others. Based on network analysis, we have also identified three genes (*HB7*, *HB32* and *At1g49830*, which encodes a bHLH TF) which may have important, hitherto unknown functions in the *LSH-CUC-BOP* meristem-organ boundary regulatory module. *HB32* connects *STM* to *CUC1*, and both *HB7* and *At1g49830* connect to *CUC1* and/or members of *LSH* gene family, suggesting possible involvement in this regulatory sub-network.

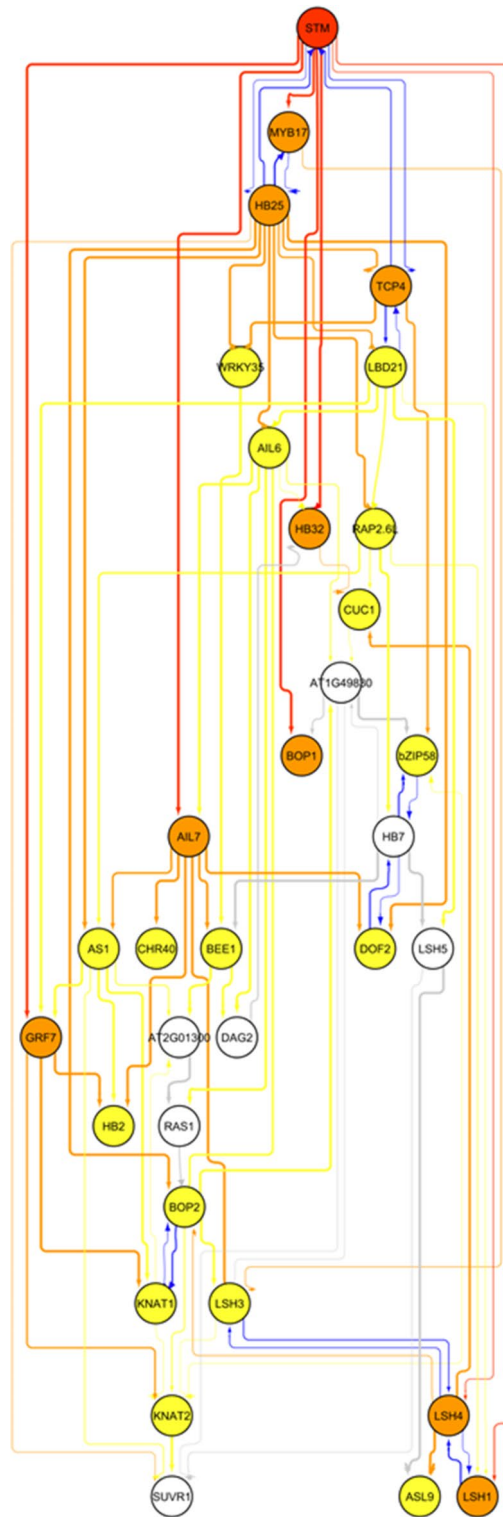
Overall, our Bayesian network analysis suggests a number of potential regulatory relationships between the various *STM* target transcriptional regulators and indicates that some may have more central roles than others in coordinating meristem function, based on their level of connectivity.

## Discussion

The class-1 *KNOX* genes encode TALE homeodomain proteins that are important regulators of shoot apical meristem development and function. A key function of the *KNOX* protein *STM* is to promote pluripotency and represses cellular differentiation in the cells of the SAM. Several studies have identified putative *KNOX* target genes in a variety of plant species, including rice, maize and Arabidopsis<sup>25,26,29</sup>. These studies have revealed that most *KNOX* target genes are associated with the control of hormone biosynthesis and response, cell differentiation and meristem organisation. In particular, *STM* has been shown to regulate a number of transcriptional regulators associated with meristem function<sup>29,42</sup>. Here, we expand this analysis to identify putative *STM* binding sites in the Arabidopsis genome using ChIP-seq analysis and correlate these sites with genes previously identified as transcriptionally responsive to altered *STM* levels. We also investigated the topology of the *STM* GRN using Bayesian network analysis. These combined approaches enabled us to develop an integrated picture of the regulatory interactions that take place among *STM*-regulated loci and allowed us to identify new genes involved in meristem function.

We utilised a robust approach for the identification of putative *STM* binding sites by comparing the *STM*-GR immunoprecipitated sample with both *STM*-GR no-antibody control and WT immunoprecipitated samples, which served as a control lacking the GR antigen. This approach was taken to eliminate noise and reduce the number of false-positive peak predictions, and substantially reduced the number of putative binding sites identified compared to the use of only a no antibody control. Hence, our analysis is conservative and there may be additional *STM* binding sites in the Arabidopsis genome which were excluded using this approach.

Our ChIP-seq analysis revealed 858 peaks corresponding to 859 distinct loci which contain one or more putative binding sites for *STM*. These binding sites were mostly located around the TSS or < 1 kb upstream of the TSS, though some binding sites were located further upstream and others were detected in, or downstream of, the gene body. This is in broad agreement with previous studies<sup>25</sup>. Assigning a locus to a particular peak was done on the basis of the closest gene to the peak, but was not always straightforward, as some peaks were detected in intergenic regions that were approximately equidistant from two adjacent loci. In such cases, we assigned the peak to whichever of the two genes showed a transcriptional response to increased *STM* levels using our previously published transcriptomics data<sup>29</sup>. If both were responsive or unresponsive to *STM*, then the peak was assigned to both loci (6 genes). Hence, we made the assumption that peaks were associated with the transcriptional control of the closest locus, but this does not take into account potential long-range effects. Moreover, identification of transcriptionally responsive genes was done using a rather limited dataset generated using Affymetrix ATH1 microarrays, which does not contain probe-sets for all transcripts from the Arabidopsis genome. Therefore, comparison of binding sites with global RNA-seq datasets should be performed to gain



**Fig. 5.** Bayesian network of STM target genes encoding transcriptional regulators. Nodes represent genes, and edges represent conditional dependency relationships. STM is shown in red, nodes directly connected to STM (1st order nodes) are shown in orange, nodes connected to these (2nd order nodes) are shown in yellow, and nodes connected to these (3rd order nodes) are shown in white. Edge connections use the same colour scheme. Edge thickness represents the connection threshold value (thinnest = 30% and thickest = 100%). Feedback relationship edges are highlighted in blue. Arrowheads on the edges indicate parent-to-child directionality.

a more comprehensive identification of STM-responsive genes. Though many genes were identified by ChIP-seq that did not show a transcriptional response to STM, or for which no data were available, we confine the following discussion primarily to those genes which were both bound by STM and showed differential expression in response to altered STM activity, as these are the most likely to represent functionally important STM targets.

Identification of putative KNOX target genes has led to the discovery of conserved *cis*-regulatory elements in target gene promoters. Canonical homeodomain proteins such as Antennapedia in *Drosophila melanogaster* typically bind to *cis*-regulatory elements containing a TAAT core in target gene promoters or enhancer elements<sup>59</sup>. Likewise in plants, the WOX homeodomain protein WUS also has affinity for TAAT sequences but has also been shown to bind the sequence TGAA and the G-box sequence TCACGTGA<sup>60–64</sup>. TALE homeodomain proteins preferentially bind to *cis*-regulatory elements containing a different core motif, such as the core TTAC motif bound by the yeast protein Mata2<sup>77</sup>. Previous studies have shown that class-1 KNOX proteins in plants have high affinity for *cis*-acting sequences that contain one or more TGAC cores. This was first shown for the HVH21 KNOX protein in barley<sup>56</sup>, and subsequently for NTH15 in tobacco<sup>33</sup> and KN1 in maize<sup>57</sup>. Further studies revealed that many KNOX-bound *cis*-regulatory elements comprised two TGAC cores, separated by a small number of nucleotides. For example, KN1 binds to a motif composed of two TGAC cores in the first intron of the *GA20ox1* gene<sup>55</sup> while in potato a similar motif was bound by POTH1 in the promoter region of *Ga20ox1*<sup>34</sup>. In rice, OSH1 bound a motif containing two TGAC cores in its own promoter and in that of other rice KNOX genes<sup>58</sup>, while in Arabidopsis, STM was shown to bind a motif comprising two TGAC cores in the promoter of the *CUC1* gene<sup>42</sup>.

We searched for conserved motifs within the sequences comprising peaks in the STM-ChIP experiment using a hidden Markov model-based approach. We were able to identify 5 distinct motifs within STM ChIP peaks, each containing one or more TGAC or TGAT cores. Of these, 3 were dyads comprising TGAC + TGAC, TGAT + TGAT or TGAT + TGAC motifs, while two contained a single TGAC core. Motif 4 was the longest motif, with several A or G residues present immediately after the TGAC core. The presence of motifs containing a TGAT core may indicate that KNOX proteins, or KNOX-BEL heterodimers, also have binding affinity for this sequence. Our motif prediction may have been too stringent to capture all biologically relevant *cis*-regulatory sequences, as we were unable to detect any of these motifs in the upstream promoter of *CUC1*, despite the presence of multiple TGAC sequences with the STM-bound region. However, the strong overrepresentation of TGAC motifs within the STM-bound sequences suggests that, as in other species, this motif forms the core of most STM-bound *cis*-regulatory elements STM target gene promoters or other regulatory regions. Future studies utilising in vitro binding assays, such as electrophoretic mobility shift assays, could help to understand the STM binding affinities for each of these *cis*-regulatory motifs.

Given that KNOX proteins have affinity for such a generic and widespread binding motif, it has been proposed that binding partners are involved in conferring target gene specificity. Indeed, KNOX proteins form heterodimeric complexes with members of the BLH (BEL1-like homeodomain) family of TALE homeodomain proteins, comprising 13 members in Arabidopsis<sup>78–80</sup>. Specificity in the regulation of target genes could therefore rely on the composition of these heterodimeric complexes, which also facilitate nuclear import, with different KNOX and BELL partners interacting to recognise different *cis*-regulatory elements using nucleotides outside of the TGAC core<sup>34</sup>. Furthermore, WUS has been shown to heterodimerise with STM and bind to the regulatory sequences of stem cell-signaling gene *CLV3*, with such STM binding also captured in our ChIP-seq data, highlighting the importance of such cooperative regulatory interactions<sup>61,81</sup>.

From a functional perspective, the STM-bound target genes identified here were in broad agreement with those found in maize and rice<sup>25,26</sup>. We found that many of the target genes which showed a transcriptional response to STM at early timepoints in our transcriptomics data are known regulators of SAM function. In particular, our data showed direct STM regulation of other KNOX genes (*KNAT1/BP* and *KNAT2*), in agreement with previous studies, but did not detect STM itself as an STM-bound target, whereas the rice ortholog OSH1 promotes its own expression as well as that of other KNOX genes<sup>58</sup>. Our network analysis also captured ASI-mediated regulation of *KNAT1*<sup>8–10</sup>, and revealed GRF7 as a potential novel factor involved in the regulation of *KNAT1*, *KNAT2* and the homeobox gene *HB2*.

Several genes involved in the formation of meristem-organ boundaries were identified, including *CUC1*<sup>42</sup>, which also activates *STM* expression during embryogenesis<sup>43–45</sup>, together with several members of the *LSH* gene family whose expression is also activated by *CUC1*<sup>76</sup>. These data support the known role for STM in organ boundary specification and previously established transcriptional regulatory relationships<sup>29,82</sup>. Our Bayesian network analysis identified direct connections and feedback relationships between several members of the *LSH* gene family, *STM* and *CUC1*. Though STM directly activates *CUC1* expression, the network only predicted a direct connection between these two factors at 20% confidence, which might result from the fact that *CUC1* expression is activated independently of STM during embryogenesis. Instead, *CUC1* showed that connection to *STM* is routed through *LSH4* and *HB32*, the latter of which may represent a novel component of the boundary regulatory module. Indeed, two additional genes (*At1g49830*, encoding a bHLH transcription factor, and the homeobox gene *HB7*) were identified that showed network connections to *CUC1* and/or the *LSH* genes and may represent additional new components of this regulatory module.

STM was also shown to directly regulate genes involved in pluripotency, shoot regeneration and phyllotaxis. These included three members of the AP2-related family of transcription factors. These were *RAP2.6L*, which has an established role in shoot development and regeneration<sup>83,84</sup>, and the *AINTEGUMENTA-like/PLETHORA* genes *AIL7/PLT7* and *AIL6/PLT3*, which are involved in promoting pluripotency in the SAM and in controlling phyllotactic patterning<sup>39–41</sup>. Hence, their functions overlap with those of *STM* suggesting they may be important downstream mediators of STM function in SAM development. In support of this, *AIL7*, and to a lesser extent *AIL6* and *RAP2.6L*, were highly connected hub nodes within the network, and showed connections to numerous

other genes with functions including boundary formation, *KNOX* gene regulation, hormone function and chromatin regulation.

Several genes involved in lateral organ formation were also identified, and many of these the functioned in both the repression of *KNOX* gene expression and the promotion of differentiation and organ polarity. *BLADE-ON-PETIOLE1* (*BOP1*) and *BOP2* were directly bound by STM and were transcriptionally induced in response to increased STM activity. *BOP1* and *BOP2* function to promote leaf adaxial identity and have been shown previously to repress the expression of *KNAT1/BP*, a relationship that is also captured in our Bayesian network analysis, together with *KNAT2* and *KNAT6*, while activating expression of *AS2*, the *LATERAL ORGAN BOUNDARIES* (*LOB*) gene and *ATH1*, which encodes a BELL domain protein that interacts with STM<sup>46–49,85</sup>. Furthermore, we have previously shown that STM downregulates several members of the class-2 *TCP* gene family<sup>29</sup>, which function in the repression of *KNOX* gene expression and the promotion of differentiation in leaf primordia. Here we show that STM binds upstream of the *TCP4* promoter, though not the promoters of the related genes *TCP3* and *TCP10*, suggesting only regulation of *TCP4* is direct. Indirect down-regulation could be mediated by promoting expression of repressors of CIN-TCP expression, such as the *miR319* gene family, for example<sup>86</sup>. Together with the known repression of *KNOX* gene expression by TCPs<sup>15</sup>, this provides a mutually antagonistic mechanism by which the *KNOX* and *TCP* expression domains can be delineated, ensuring robust separation of meristem and organ primordium identities during development. In support of this, our network analysis captured the potential mutual regulatory relationship between STM and *TCP4*.

Chromatin structure regulation has been shown to be important in the control of STM expression during floral development<sup>87,88</sup>. However, little is known about how STM might itself regulate chromatin at target loci. Here, we identify two chromatin regulatory factors as targets of STM: *SUVRI*, which encodes a histone lysine methyltransferase and *CHR40/CLASSY4*, which encodes an SNF2-domain protein that is involved in tissue-specific DNA methylation. *CHR40* expression has been shown to be localised to the SAM, suggesting that it has specific roles related to meristem function, and acts to control RNA-directed DNA methylation<sup>89,90</sup>. *SUVRI* forms a complex with other chromatin remodelling proteins and is involved in transcriptional gene silencing through changes in nucleosome positioning<sup>91</sup>. Hence, these two factors may be involved in the control of STM target gene accessibility to the transcriptional machinery, both at the level of DNA methylation and nucleosome positioning.

STM has previously been implicated in the control of phytohormone biosynthesis, perception and signaling, especially relating to cytokinin, auxin and GA. STM promotes expression of the *IPT7* gene<sup>32,36</sup>, which encodes an isopentyl transferase enzyme that is required for CK biosynthesis, in addition to the CK-receptor *AHK4*, and represses expression of cytokinin oxidase (*CKX*) genes<sup>22</sup>. We did not detect binding of STM to these loci, suggesting such regulation may be indirect. However, we did detect direct regulation of the cytokinin riboside 5'-monophosphate phosphoribohydrolase gene *LONELY GUY8* (*LOG8*), which is involved in converting cytokinin nucleotide conjugates to the active free-base form<sup>92</sup>, and adenine phosphoribosyl transferases *APT2-4*, which catalyse the reverse reaction<sup>93</sup>. We further show that STM also binds to and transcriptionally regulates several genes involved in the auxin pathway, with STM targets including several methyl indole-3-acetic acid esterases and carboxylesterases, which catalyse the conversion of methyl-IAA to the free IAA form. Additional auxin-associated factors included several genes encoding auxin response factors (ARFs), *SAUR71*, the auxin-conjugating enzyme *GH3.3* and the kinase *PINOID*. For GA associated genes, we detected STM binding to genes encoding a GA20 oxidase, a GA2 oxidase and a GA3 oxidase, though only *Ga20ox3* showed a transcriptional response. Paradoxically, this gene showed up-regulation in response to STM in the later timepoints, suggesting that this may be some kind of feedback regulation or compensation mechanism to restore GA homeostasis arising from the severely altered morphology of these plants. In agreement with this, we detected binding and up-regulation of the GA-responsive gene *GAS1* and the GA-receptor *GID1B* at later timepoints following STM up-regulation. We also identified an additional STM target gene, *BRANCHED-CHAIN AMINO ACID TRANSAMINASE2* (*BCAT2*), which involved in regulation of both *KNOX* gene expression and the biosynthesis of GA and cytokinin<sup>94,95</sup>.

Overall our combined analysis of genes that are both bound by STM and show a transcriptional response to increased levels of STM activity, together with our Bayesian network analysis, has revealed many directly-regulated STM target genes that include known transcriptional regulators and hormone-associated genes that have critical roles in controlling pluripotency and cellular differentiation associated meristematic activity and lateral organ formation. We also identified several genes with hitherto unknown regulatory roles that may be biologically important in the control of meristem function that make promising targets for future functional studies.

## Materials and methods

### Plant lines and growth conditions

The transgenic p35S::STM-GR line was provided as a gift from Rüdiger Simon<sup>20</sup>. WT Columbia-0 (Col-0) seedlings were used as the no-GR epitope control for ChIP. Seeds were germinated on GM medium (4.4 g/L MS salts, 0.5% MES buffer, 1.5% sucrose, 1% microagar) supplemented with 60 µM dexamethasone (DEX) and grown under 24 h continuous white light at 22°C for 11 days. Shoot tissue was harvested for ChIP experiments, and root tissue was discarded. All plant research was conducted within the Plant Growth Technology Hub in the Cardiff School of Biosciences in compliance with international and UK guidelines. No endangered species were used in this research.

### Chromatin immunoprecipitation (ChIP) and DNA sequencing

ChIP was performed according to Morohashi et al. (2009)<sup>96</sup> starting with 1 g of homogenised aerial tissue of 11 day-old p35S::STM-GR plants<sup>20</sup> and WT Col-0 plants (no GR-epitope control) grown on GM agar containing



60  $\mu\text{M}$  DEX. Chromatin fragmentation was performed using a Covaris M220 focused ultrasonicator. Protein A agarose beads (Merck) were used to pre-clear the digested chromatin followed by the anti-glucocorticoid receptor alpha antibody PA1-516 (Thermo Fisher) for immunoprecipitation (IP) or BSA for no-antibody mock-immunoprecipitation (NoAB) controls. Two biological replicates were performed for each sample type (as per ENCODE guidelines). DNA from IP and NoAB chromatin samples was recovered using a Qiagen MinElute DNA Purification Kit alongside an input sample consisting of fragmented chromatin that was not immunoprecipitated. DNA concentration was measured using a Qubit fluorometer. Paired-end DNA libraries were prepared with NEXTFlex Rapid DNA-Sequencing Kit 1.0 (BioScientific) using the low input protocol and quality was checked using a D1000 ScreenTape (Agilent). Sequencing was performed on the Illumina Next Seq 500 System by the Biosciences Genomics Research Hub within the School of Biosciences, Cardiff University.

### ChIP-seq data analysis

Low quality reads, sequencing artefacts, and adaptors were removed using fastp v0.20<sup>97</sup>. Sample quality reports were produced using FastQC (v0.11.9) and summarised with MultiQC v1.9<sup>98</sup>. Trimmed FASTQ reads were mapped to the Arabidopsis thaliana TAIR10 reference genome<sup>99</sup> using Bowtie 2 v2.4.1<sup>100</sup> with a maximum fragment length of 500 bp. ChIP signal distribution was visually assessed in IGB<sup>101</sup> using BigWig files generated with deepTools v3.3.0<sup>102</sup>. Peak calling was performed with MACS2 v2.2.4<sup>103</sup> for each IP and NoAB sample using their respective input sample as a background control. ChIPQC v1.26.0 within R v4.0.2<sup>104</sup> was used to assess peak calling quality. DiffBind v3.17<sup>105</sup> was used to identify statistically enriched peaks in each IP sample using their respective NoAB as a control and DESeq2 with an FDR < 0.05. BEDTools intersect v2.29.2<sup>106</sup> was used to discard all peaks present in the no GR-epitope control. The resulting putative binding sites were annotated using ChIPseeker v1.27.3<sup>107</sup> and the Araport11 gene annotation<sup>108</sup>.

### Gene set enrichment analysis (GSEA)

Gene set function was assessed by performing Gene Ontology term enrichment using g: Profiler g: GOST<sup>109</sup> with Ensembl version 109 and a g: SCS threshold of 0.05. GSEA plots were generated using custom R scripts within R version 4.3.2 ([https://github.com/TamaraLechon/Visualisation/blob/main/go\\_lollipop\\_charts.R](https://github.com/TamaraLechon/Visualisation/blob/main/go_lollipop_charts.R)).

### Binding motif analysis

Motif enrichment analysis was performed within the RSAT Plants Suite<sup>110</sup>. Sequences for each of the peaks were retrieved from their respective genomic coordinates. RSAT peak-motifs was used to perform motif discovery<sup>111,112</sup>. Positional analysis (position-analysis and local-word-analysis) was performed with a hidden Markov model of order 2 and no background model, restricted to 6 and 7 bp oligomers and only 3 hits per algorithm. Word analysis (oligo-analysis and dyad-analysis) was performed with a Bernoulli model and a background model built of random Arabidopsis thaliana genome sequences, restricted to 100 bp around the peak summit, 6 and 7 bp oligomers and only 3 hits per algorithm. The resulting motif matrices were clustered to group similar motifs together using RSAT matrix-clustering<sup>113</sup>.

### Bayesian network structural inference

Gene regulatory network analysis was performed using static Bayesian machine learning from discretised microarray expression data as described in Scofield et al. (2018)<sup>29</sup>. 2372 microarrays annotated for seedling or shoot tissue were retrieved and discretised using a custom R script within R version 4.3.2 ([https://github.com/TamaraLechon/Bayesian-network-inference/blob/main/microarray\\_discretisation.R](https://github.com/TamaraLechon/Bayesian-network-inference/blob/main/microarray_discretisation.R)). The discretised matrix was used as input data to infer a consensus gene regulatory network using Bayesian Network Inference with Java Objects (BANJO) v2.2.0 with the simulated annealing algorithm. Networks were run for 1 h, with a maximum parent count of 10, over a maximum of 10,000 restarts, initial simulated annealing temperature of 10,000, a cooling factor of 0.7, reannealing temperature of 800, a maximum of 2,500 accepted networks before cooling, a maximum of 10,000 proposed networks before cooling and a minimum of 500 accepted networks before reannealing. Fifty top-scoring networks were generated and a consensus network was produced via influence scores with a threshold of 30%. Networks were visualised using Cytoscape v3.7.2<sup>114</sup>.

### qRT-PCR analysis

Total RNA was isolated with Tripure (Roche) and cDNA synthesis was performed using the Ambion Retroscript kit or the Protoscript II kit (New England Biolabs). qRT-PCR was conducted on the Rotorgene 6000 RT-PCR machine (Qiagen) using qPCRBIO SyGreen Lo-Rox Mix (PCR Biosystems) and ACTIN2 as reference. Data was analysed using the  $\Delta\Delta\text{CT}$  method<sup>115</sup>. 35:STM-GR<sup>20,29</sup> or 35S:TGV; pTF:STM<sup>29</sup> lines treated with 60 $\mu\text{M}$  DEX were used for short-term (3 h) or longer-term (from germination) experiments respectively. For short-term experiments, target gene expression in DEX treated samples was compared to mock (DMSO) treated samples. For longer-term experiments, a DEX-treated empty vector control line (designated 237) was compared to 35S:TGV; pTF:STM lines, with two independent 35S:TGV; pTF:STM lines (designated S34 and S38) used for analysis. Averages and standard deviations (error bars) from multiple experiments (2–6 independent biological replicates) are shown in Figure S4, together with data from individual replicates. Primer sequences are available in Supplementary Table S10.

### Data availability

All sequencing data are deposited in the NCBI Sequence Read Archive under BioProject ID PRJNA1167216.

Received: 30 September 2024; Accepted: 18 December 2024



## References

- Nicolas, A. & Laufs, P. Meristem Initiation and *de novo* stem cell formation. *Front. Plant. Sci.* **13**, 891228 (2022).
- Hong, L. & Fletcher, J. C. Stem cells: engines of Plant Growth and Development. *Int. J. Mol. Sci.* **24** (19), 14889 (2023).
- Bustamante, M., Matus, J. T. & Riechmann, J. L. Genome-wide analyses for dissecting gene regulatory networks in the shoot apical meristem. *J. Exp. Bot.* **67** (6), 1639–1648 (2016).
- Yadav, R. K. et al. Plant stem cell maintenance involves direct transcriptional repression of differentiation program. *Mol. Syst. Biol.* **9**, 654. <https://doi.org/10.1038/msb.2013.8> (2013). PMID: 23549482; PMCID: PMC3658276.
- Bahafid, E. et al. W. The Arabidopsis SHORTROOT network coordinates shoot apical meristem development with auxin-dependent lateral organ initiation. *eLife* **12**, e83334 (2023).
- Hay, A. & Tsiantis, M. KNOX genes: versatile regulators of plant development and diversity. *Development* **137**, 3153–3165 (2010).
- Jia, P. et al. KNOTTED1-like homeobox (KNOX) transcription factors - hubs in a plethora of networks: a review. *Int. J. Biol. Macromol.* **253** (Pt 3), 126878 (2023).
- Byrne, M. E. et al. Asymmetric leaves1 mediates leaf patterning and stem cell function in Arabidopsis. *Nature* **408** (6815), 967–971 (2000).
- Ori, N., Eshed, Y., Chuck, G., Bowman, J. & Hake, S. Mechanisms that control *knox* gene expression in the Arabidopsis shoot. *Development* **127**, 5523–5532 (2000).
- Byrne, M. E., Simorowski, J. & Martienssen, R. A. ASYMMETRIC LEAVES1 reveals *knox* gene redundancy in Arabidopsis. *Development* **129**, 1957–1965 (2002).
- Kumaran, M. K., Bowman, J. L. & Sundaresan, V. YABBY polarity genes mediate the repression of KNOX homeobox genes in Arabidopsis. *Plant. Cell.* **14** (11), 2761–2770 (2002).
- Hay, A., Barkoulas, M. & Tsiantis, M. ASYMMETRIC LEAVES1 and auxin activities converge to repress BREVIPELCELLUS expression and promote leaf development in Arabidopsis. *Development* **133**, 3955–3961 (2006).
- Guo, M., Thomas, J., Collins, G. & Timmermans, M. C. Direct repression of KNOX loci by the ASYMMETRIC LEAVES1 complex of Arabidopsis. *Plant. Cell.* **20** (1), 48–58 (2008).
- Xu, L. & Shen, W. H. Polycomb silencing of KNOX genes confines shoot stem cell niches in Arabidopsis. *Curr. Biol.* **18** (24), 1966–1971 (2008).
- Li, Z., Li, B., Shen, W. H., Huang, H. & Dong, A. TCP transcription factors interact with AS2 in the repression of class-I KNOX genes in Arabidopsis thaliana. *Plant J.* **71**(1), 99–107 (2012). (2012).
- Yu, H. et al. TCP5 controls leaf margin development by regulating KNOX and BEL-like transcription factors in Arabidopsis. *J. Exp. Bot.* **72** (5), 1809–1821 (2021).
- Machida, Y. et al. ASYMMETRIC LEAVES2 (AS2): roles in plant morphogenesis, cell division, and pathogenesis. *J. Plant. Res.* **135** (1), 3–14 (2022).
- Long, J. A., Moan, E. L., Medford, J. I. & Barton, M. K. A member of the KNOTTED class of homeodomain proteins encoded by the *STM* gene of Arabidopsis. *Nature* **379**, 66–69 (1996).
- Endrizzi, K., Moussian, B., Haecker, A., Levin, J. Z. & Laux, T. The SHOOT MERISTEMLESS gene is required for maintenance of undifferentiated cells in Arabidopsis shoot and floral meristems and acts at a different regulatory level than the meristem genes WUSCHEL and ZWILLE. *Plant. J.* **10** (6), 967–979 (1996).
- Brand, U., Grunewald, M., Hobe, M. & Simon, R. Regulation of *CLV3* expression by two homeobox genes in Arabidopsis. *Plant. Physiol.* **129**, 565–575 (2002).
- Lenhard, M., Jurgens, G. & Laux, T. The *WUSCHEL* and *SHOOTMERISTEMLESS* genes fulfil complementary roles in Arabidopsis shoot meristem regulation. *Development.* **129**, 3195–206 (2002). (2002).
- Scofield, S., Dewitte, W., Nieuwland, J. & Murray, J. A. The Arabidopsis homeobox gene SHOOT MERISTEMLESS has cellular and meristem-organisational roles with differential requirements for cytokinin and *CYCD3* activity. *Plant. J.* **75**, 53–66 (2013).
- Challa, K. R. et al. Active suppression of leaflet emergence as a mechanism of simple leaf development. *Nat. Plants.* **7** (9), 1264–1275 (2021).
- Byrne, M. E., Imlay, E. & Ridza, N. N. B. Shaping leaves through TALE homeodomain transcription factors. *J. Exp. Bot.* **75** (11), 3220–3232 (2024).
- Bolduc, N. et al. Unravelling the KNOTTED1 regulatory network in maize meristems. *Genes Dev.* **26**, 1685–1690 (2012).
- Tsuda, K., Kurata, N., Ohyanagi, H. & Hake, S. Genome-wide study of KNOX regulatory network reveals brassinosteroid catabolic genes important for shoot meristem function in rice. *Plant Cell.* **26**, 3488–500 (2014). (2014).
- Yan, F., Hu, G., Ren, Z., Deng, W. & Li, Z. Ectopic expression a tomato KNOX Gene *Tkn4* affects the formation and the differentiation of meristems and vasculature. *Plant. Mol. Biol.* **89** (6), 589–605 (2015).
- Richardson, A., Rebocho, A. B., Coen, E. & Ectopic, K. N. O. X. Expression affects Plant Development by altering tissue cell polarity and identity. *Plant. Cell.* **28** (9), 2079–2096 (2016).
- Scofield, S. et al. Coordination of meristem and boundary functions by transcription factors in the SHOOT MERISTEMLESS regulatory network. *Development* **145** (9), 157081 (2018).
- Chung, Y. et al. Auxin Response Factors promote organogenesis by chromatin-mediated repression of the pluripotency gene SHOOTMERISTEMLESS. *Nat. Commun.* **10** (1), 886 (2019).
- Hay, A. et al. The gibberellin pathway mediates KNOTTED1-type homeobox function in plants with different body plans. *Curr. Biol.* **12**, 1557–1565 (2002). (2002).
- Jasinski, S. et al. KNOX action in Arabidopsis is mediated by coordinate regulation of cytokinin and gibberellin activities. *Curr. Biol.* **15**, 1560–1565 (2005).
- Sakamoto, T., Kamiya, N., Ueguchi-Tanaka, M., Iwahori, S. & Matsuoka, M. KNOX homeodomain protein directly suppresses the expression of a gibberellin biosynthetic gene in the tobacco shoot apical meristem. *Genes Dev.* **15**, 581–590 (2001).
- Chen, H., Banerjee, A. K. & Hannapel, D. J. The tandem complex of BEL and KNOX partners is required for transcriptional repression of *ga20ox1*. *Plant. J.* **38** (2), 276–284 (2004).
- Song, X., Zhao, Y., Wang, J. & Lu, M. Z. The transcription factor *KNAT2/6b* mediates changes in plant architecture in response to drought via down-regulating *GA20ox1* in *Populus alba* × *P. glandulosa*. *J. Exp. Bot.* **72** (15), 5625–5637 (2021).
- Yanai, O. et al. Arabidopsis KNOXI proteins activate cytokinin biosynthesis. *Curr. Biol.* **15**, 1566–1571 (2005).
- Riou-Khamlichi, C., Huntley, R., Jacquard, A. & Murray, J. A. Cytokinin activation of Arabidopsis cell division through a D-type cyclin. *Science* **283** (5407), 1541–1544 (1999).
- Dewitte, W. et al. Arabidopsis *CYCD3* D-type cyclins link cell proliferation and endocycles and are rate-limiting for cytokinin responses. *Proc. Natl. Acad. Sci. U S A.* **104** (36), 14537–14542 (2007).
- Prasad, K. et al. Arabidopsis PLETHORA transcription factors control phyllotaxis. *Curr. Biol.* **21**, 1123–1128 (2011).
- Mudunkothge, J. S. & Krizek, B. A. Three Arabidopsis *AIL/PLT* genes act in combination to regulate shoot apical meristem function. *Plant. J.* **71**, 108–121 (2012).
- Kareem, A. et al. Prasad, K. PLETHORA genes control regeneration by a two-step mechanism. *Curr. Biol.* **25**, 1017–1030 (2015).
- Spinelli, S. V., Martin, A. P., Viola, I. L., Gonzalez, D. H. & Palatnik, J. F. A mechanistic link between *STM* and *CUC1* during Arabidopsis development. *Plant. Physiol.* **156**, 1894–1904 (2011).

43. Aida, M., Ishida, T. & Tasaka, M. Shoot apical meristem and cotyledon formation during Arabidopsis embryogenesis: interaction among the *CUP-SHAPED COTYLEDON* and *SHOOT MERISTEMLESS* genes. *Development* **126**, 1563–1570 (1999).
44. Takada, S., Hibara, K., Ishida, T. & Tasaka, M. The *CUP-SHAPED COTYLEDON1* gene of Arabidopsis regulates shoot apical meristem formation. *Development* **128**, 1127–1135 (2001).
45. Aida, M. et al. Establishment of the embryonic shoot Meristem involves activation of two classes of genes with Opposing functions for Meristem activities. *Int. J. Mol. Sci.* **21** (16), 5864 (2020).
46. Ha, C. M., Jun, J. H., Nam, H. G. & Fletcher, J. C. BLADE-ON-PETIOLE 1 and 2 control Arabidopsis lateral organ fate through regulation of LOB domain and adaxial-abaxial polarity genes. *Plant. Cell.* **19** (6), 1809–1825 (2007).
47. Ha, C. M., Jun, J. H. & Fletcher, J. C. Control of Arabidopsis leaf morphogenesis through regulation of the YABBY and KNOX families of transcription factors. *Genetics* **186**, 197–206 (2010).
48. Khan, M. et al. Antagonistic interaction of BLADE-ON-PETIOLE1 and 2 with BREVIPEDICELLUS and PENNYWISE regulates Arabidopsis inflorescence architecture. *Plant. Physiol.* **158** (2), 946–960 (2012).
49. Khan, M. et al. Repression of lateral organ boundary genes by PENNYWISE and POUND-FOOLISH is essential for Meristem Maintenance and flowering in Arabidopsis. *Plant. Physiol.* **169** (3), 2166–2186 (2015).
50. Hanano, S. et al. An Artificial Conversion of roots into organs with shoot stem characteristics by inducing two transcription factors. *iScience* **23** (7), 101332 (2020).
51. Li, S. The Arabidopsis thaliana TCP transcription factors: a broadening horizon beyond development. *Plant. Signal. Behav.* **10**, e1044192 (2015).
52. Koyama, T., Furutani, M., Tasaka, M. & Ohme-Takagi, M. TCP transcription factors control the morphology of shoot lateral organs via negative regulation of the expression of boundary-specific genes in Arabidopsis. *Plant. Cell.* **19**, 473–484 (2007).
53. Koyama, T., Mitsuda, N., Seki, M., Shinozaki, K. & Ohme-Takagi, M. TCP transcription factors regulate the activities of ASYMMETRIC LEAVES1 and miR164, as well as the auxin response, during differentiation of leaves in Arabidopsis. *Plant. Cell.* **22**, 3574–3588 (2010).
54. Schommer, C., Debernardi, J. M., Bresso, E. G., Rodriguez, R. E. & Palatnik, J. F. Repression of cell proliferation by miR319-regulated TCP4. *Mol. Plant.* **7**, 1533–1544 (2014).
55. Bolduc, N. & Hake, S. The maize transcription factor KNOTTED1 directly regulates the gibberellin catabolism gene *ga2ox1*. *Plant. Cell.* **21**, 1647–1658 (2009).
56. Krusell, L., Rasmussen, I. & Gausing, K. DNA binding sites recognised in vitro by a knotted class 1 homeodomain protein encoded by the hooded gene, *k*, in barley (*Hordeum vulgare*). *FEBS Lett.* **408** (1), 25–29 (1997).
57. Smith, H. M., Boschke, I. & Hake, S. Selective interaction of plant homeodomain proteins mediates high DNA-binding affinity. *Proc. Natl. Acad. Sci. USA.* **99** (14), 9579–9584 (2002).
58. Tsuda, K., Ito, Y., Sato, Y. & Kurata, N. Positive autoregulation of a KNOX gene is essential for shoot apical meristem maintenance in rice. *Plant. Cell.* **23**, 4368–4381 (2011).
59. Ekker, S. C., von Kessler, D. P. & Beachy, P. A. Differential DNA sequence recognition is a determinant of specificity in homeotic gene action. *EMBO J.* **11** (11), 4059–4072 (1992).
60. Yadav, R. K. et al. WUSCHEL protein movement mediates stem cell homeostasis in the Arabidopsis shoot apex. *Genes Dev.* **25** (19), 225–230 (2011).
61. Perales, M. et al. Threshold-dependent transcriptional discrimination underlies stem cell homeostasis. *Proc. Natl. Acad. Sci. USA.* **113** (41), E6298–E6306 (2016).
62. O'Malley, R. C. et al. Cistrome and Epicistrome features shape the Regulatory DNA Landscape. *Cell* **165** (5), 1280–1292 (2016).
63. Busch, W. et al. Transcriptional control of a plant stem cell niche. *Dev. Cell.* **18** (5), 849–861 (2010).
64. Sloan, J. et al. Structural basis for the complex DNA binding behavior of the plant stem cell regulator WUSCHEL. *Nat. Commun.* **11** (1), 2223 (2020).
65. Spivakov, M. Spurious transcription factor binding: non-functional or genetically redundant? *BioEssays* **36** (8), 798–806 (2014).
66. Cusanovich, D. A., Pavlovic, B., Pritchard, J. K. & Gilad, Y. The functional consequences of variation in transcription factor binding. *PLoS Genet.* **10**(3), e1004226 (2014).
67. Todeschini, A. L., Georges, A. & Veitia, R. A. Transcription factors: specific DNA binding and specific gene regulation. *Trends Genet.* **30** (6), 211–219 (2014).
68. Mele, G., Ori, N., Sato, Y. & Hake, S. The knotted1-like homeobox gene BREVIPEDICELLUS regulates cell differentiation by modulating metabolic pathways. *Genes Dev.* **17**, 2088–2093 (2003).
69. Felipo-Benavent, A. et al. Regulation of xylem fiber differentiation by gibberellins through DELLA-KNAT1 interaction. *Development* **145** (23), 164962 (2018).
70. Cheng, X. et al. Silico Genome-Wide Analysis of the Pear (*Bretnschneiderioideri*) KNOX Family anfunctional characterization of *PbKNOX1*, an Arabidopsis BREVIPEDICELLUS Orthologue involved in Cell Wall and Lignin Biosynthesis. *Front. Genet.* **10**, 632 (2019).
71. Townsley, B. T., Sinha, N. R. & Kang, J. KNOX1 genes regulate lignin deposition and composition in monocots and dicots. *Front. Plant. Sci.* **4**, 121 (2013).
72. Andersen, S. U. et al. Requirement of B2-type cyclin-dependent kinases for meristem integrity in Arabidopsis thaliana. *Plant. Cell.* **20** (1), 88–100 (2008).
73. Hackbusch, J., Richter, K., Müller, J., Salamini, F. & Uhrig, J. F. A central role of Arabidopsis thaliana ovate family proteins in networking and subcellular localization of 3-aa loop extension homeodomain proteins. *Proc. Natl. Acad. Sci. USA.* **102** (13), 4908–4912 (2005).
74. Heckerman, D. & Wellman, M. P. Bayesian networks. *Commun. ACM.* **38**, 27–31 (1995).
75. Friedman Inferring cellular networks using probabilistic graphical models. *Science* **303**, 799–805 (2004).
76. Takeda, S. et al. CUP-SHAPED COTYLEDON1 transcription factor activates the expression of LSH4 and LSH3, two members of the ALOG gene family, in shoot organ boundary cells. *Plant. J.* **66** (6), 1066–1077 (2011).
77. Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D. & Pabo, C. O. Crystal structure of a MAT alpha 2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell* **67** (3), 517–528 (1991).
78. Cole, M., Nolte, C. & Werr, W. Nuclear import of the transcription factor SHOOT MERISTEMLESS depends on heterodimerization with BLH proteins expressed in discrete sub-domains of the shoot apical meristem of Arabidopsis thaliana. *Nucleic Acids Res.* **34** (4), 1281–1292 (2006).
79. Rutjens, B. et al. Shoot apical meristem function in Arabidopsis requires the combined activities of three BEL1-like homeodomain proteins. *Plant. J.* **58** (4), 641–654 (2009).
80. Hisanaga, T. et al. Deep evolutionary origin of gamete-directed zygote activation by KNOX/BELL transcription factors in green plants. *eLife.* **10**, e57090 (2021).
81. Su, Y. H. et al. Integration of pluripotency pathways regulates stem cell maintenance in the Arabidopsis shoot meristem. *Proc. Natl. Acad. Sci. USA.* **117** (36), 22561–22571 (2020).
82. Balkunde, R., Kitagawa, M., Xu, X. M., Wang, J. & Jackson, D. SHOOT MERISTEMLESS trafficking controls axillary meristem formation, meristem size and organ boundaries in Arabidopsis. *Plant. J.* **90** (3), 435–446 (2017).
83. Che, P., Lall, S., Nettleton, D. & Howell, S. H. Gene expression programs during shoot, root, and callus development in Arabidopsis tissue culture. *Plant. Physiol.* **141** (2), 620–637 (2006).

84. Yang, S., Poretska, O. & Sieberer, T. ALTERED MERISTEM PROGRAM1 restricts shoot Meristem Proliferation and Regeneration by limiting HD-ZIP III-Mediated expression of RAP2.6L. *Plant. Physiol.* **177** (4), 1580–1594 (2018).
85. Jun, J. H., Ha, C. M. & Fletcher, J. C. BLADE-ON-PETIOLE1 coordinates organ determinacy and axial polarity in Arabidopsis by directly activating ASYMMETRIC LEAVES2. *Plant. Cell.* **22** (1), 62–76 (2010).
86. Palatnik, J. F. et al. Control of leaf morphogenesis by microRNAs. *Nature* **425**, 257–263 (2003).
87. Müller-Xing, R. et al. Polycomb proteins control floral determinacy by H3K27me3-mediated repression of pluripotency genes in Arabidopsis thaliana. *J. Exp. Bot.* **73** (8), 2385–2402 (2022).
88. Manrique, S. et al. HISTONE DEACETYLASE19 controls ovule number determination and transmitting tract differentiation. *Plant Physiol.* **194** (4), 2117–2135 (2024).
89. Zhou, M., Palanca, A. M. S. & Law, J. A. Locus-specific control of the de novo DNA methylation pathway in Arabidopsis by the CLASSY family. *Nature genetics.* **50**(6), 865–873 (2018). (2018).
90. Zhou, M. et al. The CLASSY family controls tissue-specific DNA methylation patterns in Arabidopsis. *Nat. Commun.* **13** (1), 244 (2022).
91. Han, Y. F. et al. SUV2 is involved in transcriptional gene silencing by associating with SNF2-related chromatin-remodeling proteins in Arabidopsis. *Cell Res.* **24** (12), 1445–1465 (2014).
92. Kuroha, T. et al. Functional analyses of LONELY GUY cytokinin-activating enzymes reveal the importance of the direct activation pathway in Arabidopsis. *Plant. Cell.* **21** (10), 3152–3169 (2009).
93. Allen, M., Qin, W., Moreau, F. & Moffatt, B. Adenine phosphoribosyltransferase isoforms of Arabidopsis and their potential contributions to adenine and cytokinin metabolism. *Physiol. Plant.* **115** (1), 56–68 (2002).
94. Diebold, R., Schuster, J., Däschner, K. & Binder, S. The branched-chain amino acid transaminase gene family in Arabidopsis encodes plastid and mitochondrial proteins. *Plant. Physiol.* **129** (2), 540–550 (2002).
95. Gao, F., Wang, C., Wei, C. & Li, Y. A branched-chain aminotransferase may regulate hormone levels by affecting KNOX genes in plants. *Planta* **230** (4), 611–623 (2009).
96. Morohashi, K., Xie, Z. & Grotewold, E. Gene-specific and genome-wide ChIP approaches to study plant transcriptional networks. *Methods Mol. Biol.* **553**, 3–12 (2009).
97. Chen, S., Zhou, Y., Chen, Y. & Gu, J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34** (17), i884–i890 (2018).
98. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32** (19), 3047–3048. 2016 (2016).
99. Swarbreck, D. et al. The Arabidopsis Information Resource (TAIR): gene structure and function annotation. *Nucleic Acids Res.* **36**, D1009–D1014 (2007).
100. Langmead, B. & Salzberg, S. Fast gapped-read alignment with Bowtie 2. *Nat. Methods.* **9**, 357–359 (2012).
101. Freese, N. H., Norris, D. C. & Loraine, A. E. Integrated Genome Browser: visual analytics platform for genomics. *Bioinformatics* **32** (14), 2089–2095 (2016).
102. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44** (W1), W160–W165 (2016).
103. Zhang, Y. et al. Model-based Analysis of ChIP-Seq (MACS). *Genome Biology.* **9**, R137 (2008).
104. Carroll, T. S., Liang, Z., Salama, R., Stark, R. & de Santiago, I. Impact of artifact removal on ChIP quality metrics in ChIP-seq and ChIP-exo data. *Front. Genet.* **5**, 75 (2014).
105. Stark, R. & Brown, G. *DiffBind: differential binding analysis of ChIP-Seq peak data.* (2012). <http://bioconductor.org/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf>
106. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* **26**(6), 841–2 (2010). (2010).
107. Yu, G., Wang, L. G. & He, Q. Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31** (14), 2382–2383 (2015).
108. Cheng, C.-Y. et al. Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. *Plant. J.* **89** (4), 789–804 (2017). (2017).
109. Raudvere, U. et al. G:profiler: a web server for functional enrichment analysis and conversions of gene lists. *Nucleic Acids Res.* **47** (W1), W191–W198 (2019).
110. Santana-Garcia, W. et al. RSAT 2022: regulatory sequence analysis tools. *Nucleic Acids Res.* **50** (W1), W670–W676 (2022).
111. Thomas-Chollier, M. et al. A complete workflow for the analysis of full-size ChIP-seq (and similar) data sets using peak-motifs. *Nat. Protoc.* **7** (8), 1551–1568 (2012).
112. Thomas-Chollier, M. et al. RSAT peak-motifs: motif analysis in full-size ChIP-seq datasets. *Nucleic Acids Res.* **9**. <https://doi.org/10.1093/nar/gkr1104> (2011).
113. Castro-Mondragon, J. A., Jaeger, S., Thieffry, D., Thomas-Chollier, M. & van Helden, J. RSAT matrix-clustering: dynamic exploration and redundancy reduction of transcription factor binding motif collections. *Nucleic Acid Res.* **45**, 13e119 (2017).
114. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13** (11), 2498–2504 (2003).
115. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* **25**, 402–408 (2001).

## Acknowledgements

We wish to thank Rudiger Simon for the p35S::STM-GR transgenic line, and Angela Marchbank for performing the DNA sequencing within the Genome Research Hub in Cardiff School of Biosciences. The bioinformatic analysis was performed on the Cardiff School of Biosciences' Biocomputing Hub HPC infrastructure. Resources were funded by the Cardiff School of Biosciences. This work was funded by Royal Society research Grant RG150742 and BBSRC Grant BB/R008752/1, both awarded to SS.

## Author contributions

TL performed experiments, analysed data and assisted in the writing of the manuscript. NAK analysed data and provided analytical advice. JAHM helped design the research and provided critical feedback. SS designed the research, provided funding, performed experiments, analysed data and wrote the manuscript.

## Declarations

## Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-83985-1>.

**Correspondence** and requests for materials should be addressed to S.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024