

Plant Biology

Reviewed Preprint

Published from the original preprint after peer review and assessment by eLife.

[About eLife's process](#)**Reviewed preprint posted**

June 13, 2023 (this version)

Sent for peer review

April 4, 2023

Posted to bioRxiv

March 20, 2023

Cell type-specific control of secondary cell wall formation by Musashi-type translational regulators in Arabidopsis

Alicia Kairouani, Dominique Pontier, Claire Picart, Fabien Mounet, Yves Martinez, Lucie Le-Bot, Mathieu Fanuel, Philippe Hammann, Lucid Belmudes, Rémy Merret, Jacinthe Azevedo, Marie-Christine Carpentier, Dominique Gagliardi, Yohann Couté, Richard Sibout, Natacha Bies-Etheve , Thierry Lagrange 

Laboratoire Génome et Développement des Plantes, Université de Perpignan via Domitia, CNRS, UMR5096, Perpignan Cedex, France • Laboratoire de Recherche en Sciences Végétales, Université de Toulouse III, CNRS UPS UMR5546, Castanet-Tolosan, France • FRAIB-CNRS plateforme imagerie, Castanet-Tolosan, France • Biopolymères Interactions Assemblages, UR1268 BIA, INRAE, Nantes, France • PROBE research infrastructure, BIBS Facility, INRAE, Nantes, France • Plateforme Protéomique Strasbourg Esplanade de CNRS, Université de Strasbourg, Strasbourg, France • Université Grenoble Alpes, INSERM, UA13 BGE, CNRS, CEA, FR2048, Grenoble, France • Institut de Biologie Moléculaire des Plantes, BMP, CNRS, Université de Strasbourg, Strasbourg, France

 (https://en.wikipedia.org/wiki/Open_access)

 (<https://creativecommons.org/licenses/by/4.0/>)

Abstract

Deciphering the mechanism of secondary cell wall/SCW formation in plants is key to understanding their development and the molecular basis of biomass recalcitrance. Although transcriptional regulation is essential for SCW formation, little is known about the implication of post-transcriptional mechanisms in this process. Here we report that two *bonafide* RNA-binding proteins homologous to the animal translational regulator Musashi, MSIL2 and MSIL4, function redundantly to control SCW formation in Arabidopsis. MSIL2/4 interactomes are similar and enriched in proteins involved in mRNA binding and translational regulation. MSIL2/4 mutations alter SCW formation in the fibers, leading to a reduction in lignin deposition, and an increase of 4-*O*- glucuronoxylan methylation. In accordance, quantitative proteomics of stems reveal an overaccumulation of glucuronoxylan biosynthetic machinery, including GXM3, in the *msil2/4* mutant stem. We showed that MSIL4 immunoprecipitates *GXM* mRNAs, suggesting a novel aspect of SCW regulation, linking translational control to the regulation of SCW biosynthesis genes.

eLife assessment

Secondary cell walls support vascular plants and conduct water throughout the plant body, and are **important** resources for lignocellulosic feedstocks. Here the authors present **convincing** genetic and biochemical evidence that secondary cell wall synthesis, known already to be under complex transcriptional control, is also controlled post-transcriptionally by MUSASHI-like RNA-binding proteins. These **important** results point to a new mechanism for control of secondary cell wall synthesis, which will be interesting to cell biologists and biochemists studying and attempting to manipulate plant biomass.

Introduction

As the main constituent of the terrestrial plant biomass, lignocellulose represents a huge reservoir of fixed carbon and a renewable resource for bioproducts and energy that is essential for human usages¹. Lignocellulose biomass enclosed mainly within secondary cell wall (SCW) reinforces vessels for long-distance transport and provides physical properties to fibers allowing upright growth of the plants^{1, 2}. SCWs are composite material made of high molecular weight biopolymers, including cellulose, hemicelluloses, and lignin, as well as cell wall proteins whose composition and structure can differ markedly between dicots and monocots^{2, 3}. Cellulose is synthesized at the plasma membrane by the cellulose synthase complex and is composed of linear chains of β -(1–4) glucans that aggregate to form highly crystalline cellulose microfibrils^{2, 3}. Xylans, the main hemicellulosic polysaccharides in SCW, are synthesized in the Golgi apparatus by a complex biosynthetic machinery and comprise a group of polysaccharides that share a common backbone of β -(1,4)-linked xylose (Xyl) units, but differs by the presence of side chains, whose nature is dependent on tissues and species². In particular, glucuronoxylan, the dominant xylan in the SCW of dicots, is notably substituted at O-2 position with α -D-glucuronic acid (GlcA) or (4-O-methyl)- α -D-glucuronic acid (MeGlcA) groups, and with O-acetyl moieties at C-2 or C-3 positions². By interacting with the cellulose and lignin, xylans contribute to the strengthening of the SCW, and are considered to be one of the main factors contributing to the biomass recalcitrance to enzymatic hydrolysis^{3, 4}. Finally, lignin polymers are made up of H, G, and S units which are the results of the oxidation and the polymerisation of the *p*-coumaryl, coniferyl and synapyl alcohols, respectively. These monomers, also called monolignols are synthesized on the ER surface/cytoplasm prior transportation to the cell wall. Lignins confer stiffness, strength and hydrophobicity to the SCW^{2–4}.

The metabolic load imposed by the coordinated production of SCW biopolymers emphasizes the importance for plants of having a precise control over temporal and spatial expression of the corresponding SCW biosynthetic genes. Based on founding studies performed in *Arabidopsis* model, it is generally recognized that the SCW synthesis is primarily controlled at a transcriptional level by a multilayer and interconnected network of transcription factors/TFs^{5, 6}. Key players of this regulatory network are related NAC (NAM-ATAF1/2-CUC) and MYB-type TFs that induce, in a redundant and combinatorial manner, the expression of SCW biosynthetic genes by binding their promoter region⁵. In addition, a family of Class III homeodomain leucine zipper (HD-ZIP III) TFs, whose members control several aspects of *Arabidopsis* development, was also shown to interact with SCW regulatory network, contributing to xylem cell specification and SCW synthesis^{5, 6}. Although this transcriptional level of control can gain complexity in perennial plants, functional orthologs of the main

regulators were identified in many different species, suggesting that SCW deposition proceeds via a conserved mechanism in all vascular plants^{7, 8}. During the last decade, engineering the SCW regulatory network to deposit modified lignin and/or improve polysaccharide composition became a promising strategy to optimize biomass processability^{9, 10}. However, despite huge progresses in understanding the regulation of SCW synthesis, the engineering of SCW regulatory network to improve biomass yield and digestibility has proven difficult and often impairs plant growth because of altered vascular tissue development^{11, 12}.

Downstream of transcriptional regulations by TFs, RNA-binding proteins (RBPs) are essential post-transcriptional modulators of gene expression across all kingdoms of life¹³. However, although a large cohort of RBPs exists in plants¹⁴, there is little information about RBPs-mediated posttranscriptional regulation of SCW biosynthetic genes. So far, only few microRNAs (miRNAs) families have been implicated in the post-transcriptional regulation of genes involved in both regulatory and enzymatic aspects of SCW biosynthesis⁶. In particular, miR165/166 were shown to target HD-ZIPIII transcriptional regulators, having a direct impact on the transcriptional network associated with SCW biosynthesis⁶. In the other hand, miRNA397/857/408 were shown to affect lignin content *via* the targeting of laccase/LAC genes, which control lignin polymerization from monolignols precursors^{15, 16}. In addition, two related tandem CCH zinc-finger proteins, which exhibit both DNA and RNA binding abilities, have been proposed to modulate SCW formation by regulating the expression of genes associated with cell wall metabolism¹⁷. Despite these mounting evidences, our current knowledge of the role of post-transcriptional regulators in SCW synthesis is still very limited.

In this work, we report that two RRM-domain containing RNA-binding proteins homologous to the animal translational regulator Musashi, Musashi-like2/MSIL2 and Musashi-like4/MSIL4, function redundantly to control various aspects development, including the stiffness of the inflorescence stem in Arabidopsis. We show that RRM-dependent RNA-binding activity is essential for MSIL2/4 functions *in vivo*, and that MSIL2/4 interactomes are similar, being enriched in proteins involved in 3'-UTR binding and translational regulation. MSIL2/4 mutations alter the formation of SCW in the interfascicular fiber cells, leading to a reduction in lignin deposition, and a change in the decoration pattern of glucuronoxylan that is associated with an increase of 4-*O*-methylation of GlcA substituent. In accordance, quantitative mass-spectrometry-based protein analysis reveal an overaccumulation of glucuronoxylan biosynthetic machinery, including the GlucuronoXylan Methyltransferase3/GXM3, in the *msil2/4* mutant stem. We show that MSIL4 immunoprecipitates GXM3 mRNA *in vivo*, likely regulating its expression at a translational level. Our results demonstrate that MSILs regulate SCW synthesis in interfascicular fiber cells and point to a novel aspect of SCW regulation linking translational repression to regulation of SCW biosynthesis genes.

Results

Musashi-like MSIL2/4 proteins redundantly control specific Arabidopsis development processes

While many studies have implicated the MSIs in the control of gene expression during cellular proliferation, cell fate determination and cancer in animals^{18, 19}, yet little is known about the presence and activity of MSI-type proteins in plants. In this work, we describe an RNA-binding protein family in Arabidopsis, whose members hereafter named as *MUSASHI-Like1* to *4/MSIL1-4*, share both sequence, domain organization, and model-based structural similarities with animal MSIs (Fig. 1a, b). Phylogenetic analysis revealed that Arabidopsis

indeed harbors a clade of seven *MSIL*-type genes with a prominent sub-clade harboring the *MSIL1-4* genes (Fig. 1b). Notably, a neighboring sub-clade contains two genes, *RBGD2* and *RBGD4* (Fig. 1b), that encode for heat-inducible RBPs implicated in the response to heat stress in Arabidopsis²⁰. *MSIL1-4* proteins share a common domain architecture consisting of two N-terminal RNA recognition motifs (RRM1 and RRM2) that are followed by a poorly conserved, intrinsically unstructured, carboxy-terminal region (Fig. 1a; and Supplementary Fig. 1a). The survey of the gene expression ARAPORT11 databases (<https://araport.org/>; <http://fgcz-pep2pro.uzh.ch>) revealed that the *MSIL1-4* genes are widely expressed in Arabidopsis, with *MSIL1* showing the lowest levels of expression (Fig. 1c). Database searches further indicated that *MSIL* orthologs are widely distributed in land plants, from bryophytes to angiosperms (Supplementary Fig. 1b). To gain a mechanistic understanding of *MSIL* function in Arabidopsis, we determined the subcellular localization of these proteins by performing fluorescence microscopy on GFP-tagged versions of *MSIL2* and *MSIL4* (*MSIL2G* and *MSIL4G*). Both proteins showed a diffuse cytoplasmic distribution in root cells of stable Arabidopsis transgenic plants that contrasts with the nucleus and cytoplasmic distributions of the free GFP protein control (Fig. 1d).

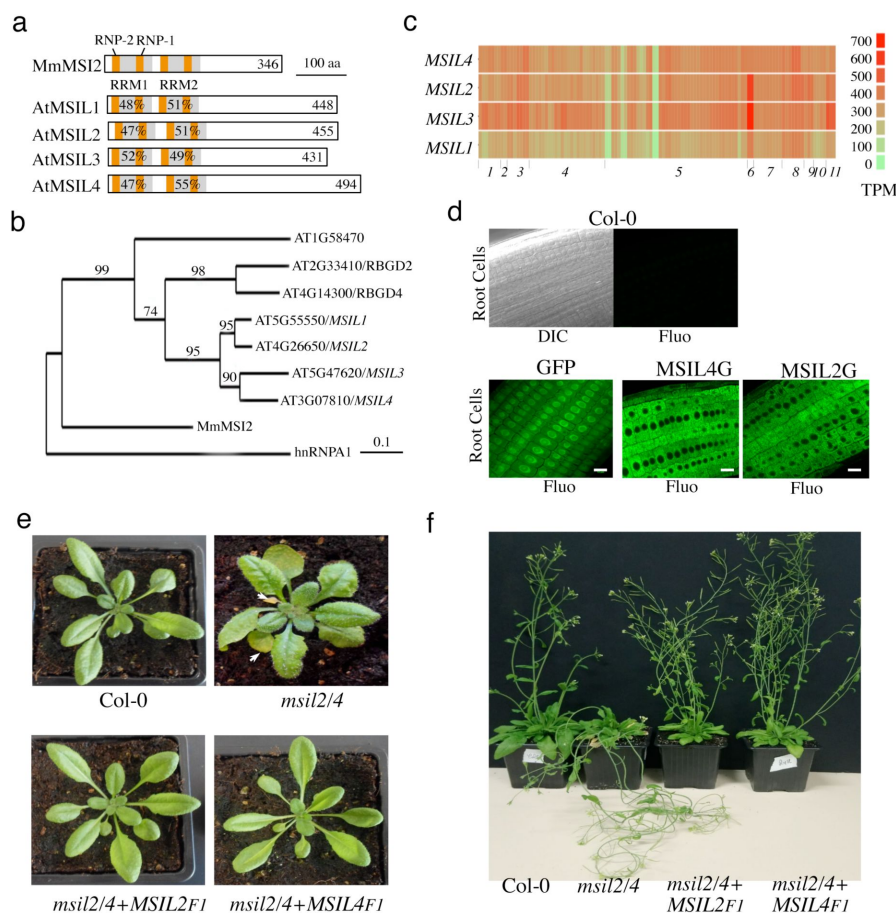


Fig. 1

Musashi-like *MSIL2/4* proteins redundantly control development in Arabidopsis.

a) Schematic representation of the Musashi-like (*MSIL*) protein family in *Arabidopsis thaliana*. The invariant RNP1 and RNP2 motifs within the conserved RRM domains are indicated. Numbers refer to amino acid identities between the Mouse Musashi/MSI RRM1/2 domains and the corresponding domains in Arabidopsis *MSIL* homologs. b) Evolutionary relationships between *MSIL*s and related RNA-binding proteins. The scale bar indicates the rate of evolutionary change expressed as number of amino acid substitutions per site. c) RNA-seq expression map of *MSIL* genes extracted from ARAPORT11. Shading is a log₂ scale of transcripts per million (TPM). d) General overview of the roots of 6-days-old wild type Arabidopsis

seedlings (Col-0) or transgenic seedlings expressing either a free GFP protein or GFP-tagged versions of *MSIL2/MSIL2G* and *MSIL4/MSIL4G*. Scale bar, 10 μm. e) Photographs of representative rosettes of Col-0, *msil2/4*, *msil2/4-MSIL2F1* and *msil2/4-MSIL4F1* plants. f) Photographs of representative inflorescence stems of Col-0, *msil2/4*, *msil2/4-MSIL2F1* and *msil2/4-MSIL4F1* plants. Abbreviations: Mus musculus Musashi2 (MmMsi2); human heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1).

To further investigate the function of the MSILs in Arabidopsis, we raised specific antibodies against non-conserved regions of MSIL1-4 and characterized T-DNA insertion *msil* mutant lines that lack full-length *MSIL* mRNA and MSIL proteins as judged by RT-PCR and western blot experiments, respectively (Supplementary Fig. 2a-c). However, none of the four characterized single *msil* mutants showed any obvious phenotype (Supplementary Fig. 3a and data not shown). To assess the potential functional redundancy between *MSIL* genes, we crossed the single *msil* mutants together and characterized the six possible double mutant combinations (Supplementary Fig. 3a). Only the *msil2-1/msil4-1* (*msil2/4*) double mutant showed discernable developmental abnormal phenotypes, including enlarged and curled leaves, early leaf senescence and a pendant inflorescence stem (Fig. 1e, f; and Supplementary Fig. 3a, b). To assess whether MSIL1 and MSIL3 also contribute to Arabidopsis development in the *msil2/4* double mutant background, we generated a *msil1-1/msil2-1/msil3-1/msil4-1* (*msil1/2/3/4*) quadruple mutant that shows no particular or aggravated developmental phenotypes with respect to the *msil2/4* double mutant (Supplementary Fig. 3a, c, d). To test the functional redundancy of MSIL2 and MSIL4 proteins *in vivo*, we transformed the *msil2/4* double mutant with constructs coding either for a Flag/HA-tagged version of MSIL2 (MSIL2F) or MSIL4 (MSIL4F), and selected two independent transgenic plants expressing similar and near physiological levels of the tagged protein, as evaluated by western blot using our home-made antibodies (Supplementary Fig. 3e). The ectopic expression of either MSIL2F or MSIL4F proteins was able to rescue all the phenotypes associated with the MSIL2/4 defect (Fig. 1e, f; and Supplementary Fig. 3b). This confirms that MSIL2F and MSIL4F are functional proteins and more importantly that MSIL2 and MSIL4 redundantly control various aspects of development in Arabidopsis.

RRM-dependent RNA-binding activity is essential for MSIL2/4 functions *in vivo*

Animal MSIs interact with target transcripts through their RRM domains and mutations in these domains have been shown to impair MSI activity *in vivo*^{21, 22}. MSILs are very similar to metazoan MSIs over their RRM sequences (Fig. 1a; and Supplementary Fig. 4a), a closeness of sequences also observed when homology models of the MSIL4 RRM motifs were generated using SWISS-MODEL server²³ (Fig. 2a). Homology-based structure prediction proposed that these domains adopt the characteristic RRM fold with a four-stranded β -sheet structure bearing the conserved phenylalanine residues involved in the specific recognition of RNA bases²⁴ (Fig. 2a; and Supplementary Fig. 4a). This supports the idea that MSILs have retained the ability to interact with RNA, a notion further supported by the identification of MSIL2/4 proteins in the experimentally determined Arabidopsis mRNA-binding proteome¹⁴. To examine the requirement for the interaction of MSIL with RNA in developmental control, we first generated an RNA-binding mutant form of MSIL4, MSIL4^{RRM}, by introducing phenylalanine to aspartate (F $\rightarrow$ D) mutations in both RRM domains (Fig. 2b). Protein-RNA interaction assays using RNA homopolymers immobilized on agarose beads confirmed that the MSIL4, but not MSIL4^{RRM}, exhibited intrinsic RNA binding activity *in vitro* (Fig. 2c). To assess the impact of these mutations *in vivo*, the MSIL4G and MSIL4G^{RRM} constructs were expressed under the control of the *MSIL4* endogenous promoter into the *msil2/4* mutant, and two independent transformants expressing near physiological levels of the WT (*MSIL4G*-3 and *MSIL4G*-8), and mutant (*MSIL4G*^{RRM}-3 and *MSIL4G*^{RRM}-10) proteins were selected for further complementation analysis (Supplementary Fig. 4b). MSIL4G, but not MSIL4G^{RRM} proteins were able to rescue the developmental defects incurred by *MSIL2/4* mutations (Fig. 2d; and Supplementary Fig. 5a, b), indicating that these proteins are likely to exert their functions by interacting with RNA targets *in vivo*.

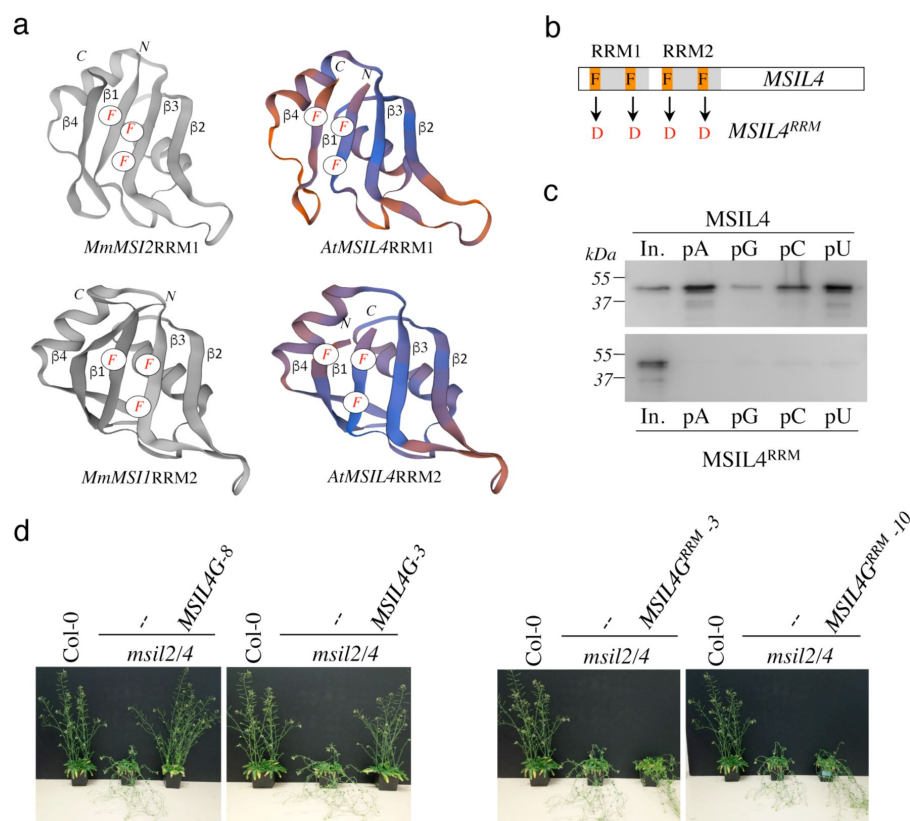


Fig. 2

RRM-dependent RNA-binding activity is essential for MSIL function *in planta*.

Experimentally determined structures of the RRM1 domain of MmMSI2 (PDB ID: 6C8U) and the RRM2 domain of MmMSI1 (PDB ID: 5x3y), and models of the RRM domains of MSIL4 generated using the homology-modeling server SWISS-MODEL. The α -helices and β -sheets of RRM domain as well as the phenylalanine residues that contact RNA bases are highlighted. **b)** Schematic representation of the mutations introduced in the RRM domains of MSIL4. **c)** Binding assays of MSIL4 RRM domains on ssRNA homopolymers *in vitro*. Single-stranded polyA (pA), polyG (pG), polyC (pC) or polyU (pU) RNAs were subjected to binding

with the His-tagged recombinant RRM domains from MSIL4 (upper part) or its mutated version (MSIL4^{RRM}, lower part). An anti-His antibody was used for detection. In. represents the input fraction. **d)** RRM-dependent RNA-binding activity is essential for MSIL2/4 function in stem. Photographs of representative inflorescence stems of Col-0 and *msil2/4* mutants complemented or not with a WT version of MSIL4 (MSIL4G-8 and 3) or an rrm mutant (MSIL4G^{RRM}-3 and 10).

MSIL2/4 protein interactomes are enriched in proteins involved in 3'-UTR binding and translational regulation

Animal MSIs modulate mRNA expression mostly at a translational level through binding to conserved motifs in the 3'-UTR of the target mRNA and further interactions with various mRNA binding partners^{18, 25, 26}. To further characterize the components of the MSIL2/4 network in Arabidopsis, we performed affinity purification coupled to LC-MS/MS, as described in Scheer et al.²⁷, using the complemented *MSIL2F1* and *MSIL4F1* lines. Consistent with their functional redundancy, MSIL2F1 and MSIL4F1 exhibit a similar protein interaction network, with 27 proteins significantly enriched in both MS analyses (Fig. 3a, b; and Supplementary Table 1). Examination of the PANTHER database (<http://www.pantherdb.org>) identified that MSIL2F1/4F1 are part of a protein-protein interaction network that was principally enriched in 4 categories of GO molecular functions, namely poly(A) binding protein (GO:0008143), mRNA 3-UTR binding (GO:0003730), single-stranded RNA binding (GO:0003727), and translation initiation factor activity (GO:0003743), highlighting the relationship between MSIL2/4 and the mRNA binding and regulatory proteome (Fig. 3c).

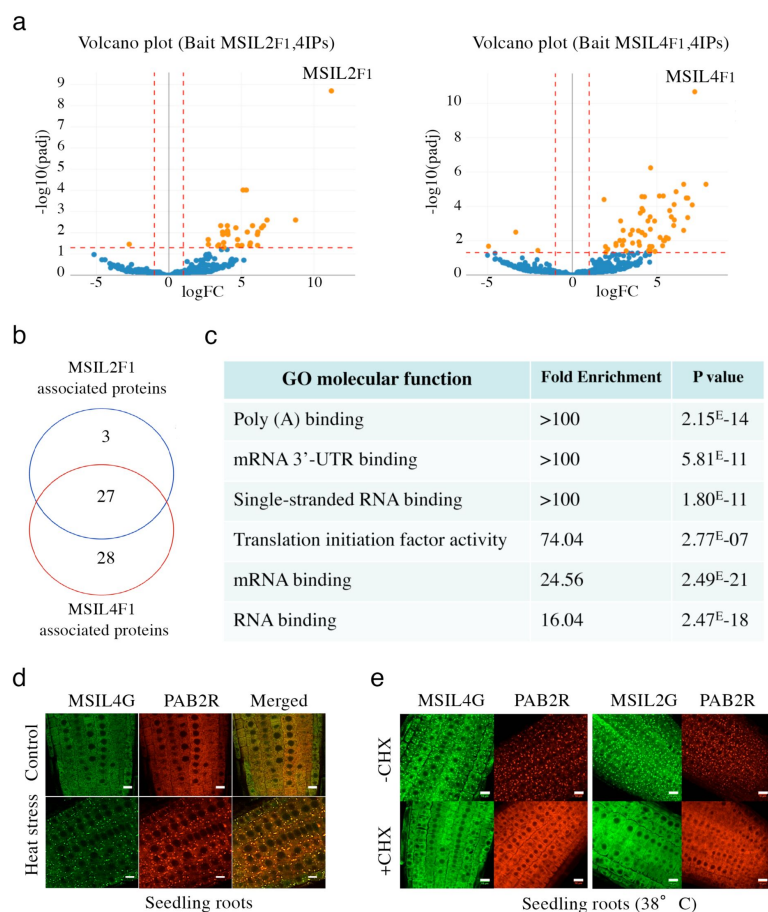


Fig. 3

MSIL2/4 protein interactomes are enriched in proteins involved in 3'-UTR binding and translation regulation.

a) The semi-volcano plots show the enrichment of proteins co-purified with MSIL2F1 and MSIL4F1 as compared with control IPs. Y- and X-axis display adjusted p-values and fold changes, respectively. The dashed lines indicate the threshold above which proteins are significantly enriched/depleted (fold change > 2; adjP < 0.05). b) Venn diagram showing the overlap between the MSIL2 and the MSIL4 interactomes. c) PANTHER-based classification of the GO molecular function that are overrepresented among the MSIL-interacting proteins. d) Confocal monitoring of the colocalization of MSIL2G, MSIL4G and PAB2R in root tips of 8-d-old seedlings were monitored after 30 min of exposure to 38°C (heat stress) or 20°C for control treatment. e) Cycloheximide treatment inhibits the formation of MSILG and PAB2R cytosolic foci. The transgenic plants expressing the stress granule marker PAB2R was used as a control. Root of 7-day-old

seedlings expressing either MSIL4G, MSIL2G, or PAB2R were monitored after 1 hour of exposure to 38°C in the presence of DMSO control treatment (DMSO) or in the presence of cycloheximide inhibitor (+CHX). Scale bar, 10 µm.

Many of the proteins present in the MSIL2/4 interactome, such as PABs, PUMs, UBP1 are involved in translation regulation and are targeted to stress granule (SGs) components, a membraneless cytoplasmic organelle that formed consecutive to the global inhibition of translation caused by heat-stress^{28, 29}. To assess the potential functional link between MSIL and those proteins, we crossed the MSIL2G/4G plants with a functional PAB2-RFP/PAB2R line²⁹, and assessed the localization of the fusion proteins before and after heat-stress in both parent and crossed lines. Fluorescence analysis revealed that MSIL2G/4G, like PAB2R, show a dispersed cytoplasmic signal under normal condition, and are recruited to cytoplasmic granules that overlap with PAB2R-containing SG upon heat stress (Fig. 3d; and Supplementary Fig. 6). Treatment with cycloheximide (CHX), a translation inhibitor that prevent SG assembly by trapping mRNAs at polysomes³⁰, confirmed that those foci correspond to *bonafide* SGs (Fig. 3e). These results were consistent with the RNA-binding protein-based signature and the nature of MSIL interactomes, further supporting a functional relationship between MSILs and the translation machinery in Arabidopsis.

MSIL2/4 proteins regulate the molecular architecture of SCW in Arabidopsis fibers

Pendant stem inflorescence is a common phenotypic trait of Arabidopsis mutants defective in SCW biosynthesis, that results from a lack of mechanical support and rigidity^{6, 31}. Given the current lack of knowledge of the role of post-transcriptional mechanisms in the control of SCW biosynthesis in plants, we henceforth focused our analysis on the pendant stem phenotype of the *msil2/4* double mutant. Survey of the Arabidopsis inflorescence stem tissue-specific transcriptome database³² indicated that the *MSIL2/4* genes are highly expressed in the SCW-forming xylary fiber (F) and xylem vessel (X) cells (Fig. 4a). Toluidine blue O staining and microscopic analysis of Col-0 and *msil2/4* mutant stem cross-sections revealed no major changes in the architecture of both the interfascicular and xylary fibers in the mutant stem. Notably, xylem cells deformation or collapse, which are a characteristic of mutants deficient in SCW synthesis, were not observed in the *msil2/4* mutant (Fig. 4b). However, the examination of the cell wall thickness in cross sections of stems revealed that the interfascicular fiber cells display a significantly thinner SCW in *msil2/4* mutant (Fig. 4c). UV autofluorescence and phloroglucinol staining of the Col-0, *msil2/4*, *msil2/4+MSIL2F1* and *msil2/4+MSIL4F1* stem sections revealed a reduced lignin signal in the interfascicular fibers of the *msil2/4* mutant compared to Col-0 or complemented plants (Fig. 4d; and supplementary Fig. 7a). In agreement with these observations, a significant decrease of 30% in AcBr lignin content was observed in the *msil2/4* mutant compared to the Col-0 control plants (Fig. 4e). Despite a tendency to decrease, the difference of lignin content was not significant in the complemented lines *msil2/4+MSIL2F1* and *msil2/4+MSIL4F1* compared to Col-0 (Fig. 4e). To further investigate potential changes in the lignin structure and composition in the Col-0 and *msil2/4* mutant, we evaluated the impact of MSIL2/4 deficiency on lignin composition by thioacidolysis, a degradation method allowing the determination of the relative amounts of H, G and S lignin units linked by beta-O-4 linkages. Total monomer yield did not significantly differ between the col-0 and the *msil2/4* mutant (Supplementary Fig. 7b), suggesting that the structure of the lignin polymer is not merely affected in the mutant. Further examination of β -glucan composition using the fluorescent dye calcofluor white³³ revealed an increased staining in the interfascicular fiber cells of *msil2/4* mutant, supporting changes in the composition/accessibility of cellulose (Fig. 4f; and Supplementary Fig. 7c).

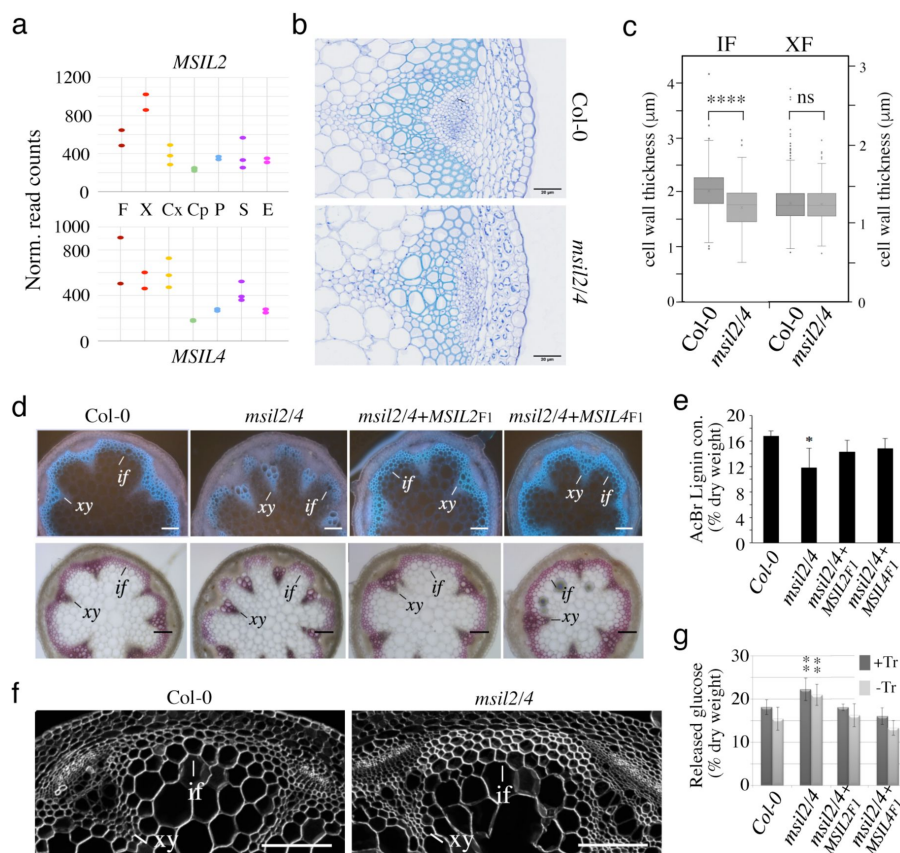


Fig. 4

MSIL proteins regulate the molecular architecture of SCW in Arabidopsis fibers.

a) Expression profiles of the *MSIL2* and *MSIL4* genes in the inflorescence stem tissues as retrieved from the Arabidopsis inflorescence stem tissue-specific transcriptome database⁴⁰ (<https://arabidopsis-stem.cos.uni-heidelberg.de/>). The results of two replicates are shown. F, fibers; X, xylem vessels; Cx, Cambium (xylem side); Cp, Cambium (Phloem side); P, Phloem; S, Starch sheath; E, Epidermis. b) Cross-sections of *Col-0* and *msil2/4* stems showing vascular bundle and interfascicular fiber cells stained with Toluidine blue. c) Measurements of the SCW thickness of xylary fiber and interfascicular fiber cells. Values are means ($n > 400$) $\pm$ SEM. Data were analyzed by

unpaired Student's test. Asterisks indicate significant differences relative to *Col-0*; **** $p < 0.0001$. d) Top: lignin autofluorescence under ultraviolet (UV) using confocal microscopy in *Col-0*, *msil2/4* mutant, and complemented *msil2/4-MSIL2F1* and *msil2/4-MSIL4F1* plant stem sections. xy: xylem fibers; if: interfascicular fibers. Bottom: Phloroglucinol-HCl staining of *Col-0*, *msil2/4*, and complemented *msil2/4-MSIL2F1* and *msil2/4-MSIL4F1* plant stem sections. Scale bar, 100 μ m. e) Histograms represent acetyl bromide lignin content, expressed as % of dry weight. Lignin content was measured in the bottom section of mature stems (3 biological replicates per line). Significant differences between *Col-0*, *msil2/4+MSIL2F1*, *msil2/4+MSIL4F1*, and *msil2/4* plant lines are indicated by asterisks*, P -value < 0.01 according to Student's t test ($n = 6$ to 10). f) Confocal microscopy of Calcofluor-white staining of cross sections of *Col-0* and *msil2/4* mutant inflorescence stems. Scale bar, 100 μ m. g) Glucose yield after incubation of *Col-0* and *msil2/4* lignocellulosic material with cellulases either without pre-treatment (-Tr) or after sodium hydroxide pre-treatment (+Tr).

To extend this analysis, we then subjected purified *Col-0* and *msil2/4* SCW materials to Fourier-Transform Infrared (FTIR) spectroscopy³⁴. PLS-DA analyses performed on FT-IR absorbance data clearly separated *msil2/4* mutant lines from *Col-0* plants, using the first two components (explaining 74% and 10% of total variability, respectively) (Supplementary Fig. 7d, left panel). The large-scale chemotyping of CW composition in stem by FTIR spectroscopy suggested that not only lignin deposition was impaired in *msil2/4* mutants but also polysaccharides composition (Supplementary Fig. 7d, right panel). The lowest absorbance values in *msil2/4* lines were observed in the region 1800-1500 cm^{-1} , mostly associated to the absorbance of lignin related compounds. On the contrary, a higher absorbance was detected in *msil2/4* samples within 1050-900 cm^{-1} region, mostly associated to polysaccharides compounds³⁴. Because the optimal interaction between the SCW lignin and polysaccharide polymers is believed to contribute to SCW recalcitrance, we used saccharification as a proxy

to compare cell wall sugars accessibility between Col-0, *msil2/4* and complemented mutant lines (Fig. 4g). A higher amount of reducing sugars from native cell wall (without pre-treatment/-Tr) was released using *msil2/4* samples compared to Col-0 samples. After cell wall loosening using alkali pre-treatment/+Tr, an even more important amounts of sugars (+5%) was released from double mutants CW enzymatic digestion (Fig. 4g). Together, our data indicate that MSIL2/4 regulate the molecular architecture of SCW in the interfascicular fibers, contributing significantly to the setting of biomass recalcitrance in Arabidopsis.

The accumulation and activity of the glucuronoxylan decoration machinery are mis-regulated in *msil2/4* mutant

To elucidate the impact of the MSIL2/4 on the expression of SCW-related biosynthesis genes in the Arabidopsis inflorescence stem, we performed mRNA sequencing (mRNA-seq) on both wild-type and *msil2/4* mutant backgrounds. In total, we found 234 genes to be significantly differentially regulated (DEG) in the *msil2/4* mutant background (with stringent thresholds having a logarithm of fold change >1.5 or <-1.5 and false discovery rate (FDR ≤ 0.05), including 156 up-regulated and 78 down-regulated mRNAs (Supplementary Fig. 8a; and Supplementary Table 2). Interestingly, genes encoding regulatory or enzymatic components of the SCW biosynthesis pathway were absent from the DEG list, and their expression was not significantly changed in *msil2/4* mutant (Supplementary Fig. 8a). Gene Ontology analysis of DEGs revealed significantly enriched signaling pathways (FDR ≤ 0.05) involved in plant defense or responses to biotic and abiotic stresses (Supplementary Fig. 8b), encoding products that are preferentially targeted to the extracellular region and the cell wall (Supplementary Fig. 8c). In this regard, the constitutive activation of defense pathways has been well documented in mutant plants defective in cell wall formation, including the SCW³⁵.

The absence of SCW biosynthesis genes in the DEG list suggests that MSIL2/4 could regulate the SCW formation at a post-transcriptional/translational level, a hallmark of MSI function in animal models^{18, 19}. To investigate proteome-level changes that could potentially occur upon MSIL2/4 mutations, we performed a comparative proteomic analysis of Col-0 and *msil2/4* inflorescence stems using MS-based quantitative proteomics, and represented ≈4300 reliably identified and quantified proteins in a volcano plot according to their statistical p-value and their relative difference of abundance. This analysis revealed 267 proteins with significant abundances in the *msil2/4* mutant versus Col-0 inflorescence stems (fold change ≥2, Benjamini-Hochberg FDR < 1%), including 156 up-regulated, and 111 down-regulated proteins (Fig. 5a; and Supplementary Table 3). While no specific GO terms were enriched in the set of down-regulated proteins, the inspection of the up-regulated proteins revealed a strong enrichment in enzymes involved in the glucuronoxylan biosynthetic pathway (Fig. 5a, b; and supplementary Table 3). Notably, the upregulated glucuronoxylan-related proteins included glycosyltransferases required for the xylan backbone synthesis (IRX9/10/14), as well as enzymes involved in its substitution (ESK1, GUX1, and GXM3)² (Fig. 5a, c). Visual inspection of the RNA-seq aligned reads of the up-regulated glucuronoxylan-related genes using the Integrative Genomics Viewer (IGV) and quantitative RT-PCR analysis of the glucuronoxylan-related decoration genes confirmed that the observed proteomic changes in *msil2/4* were not associated with significant variations in mRNA levels (Supplementary Fig. 9a, b).

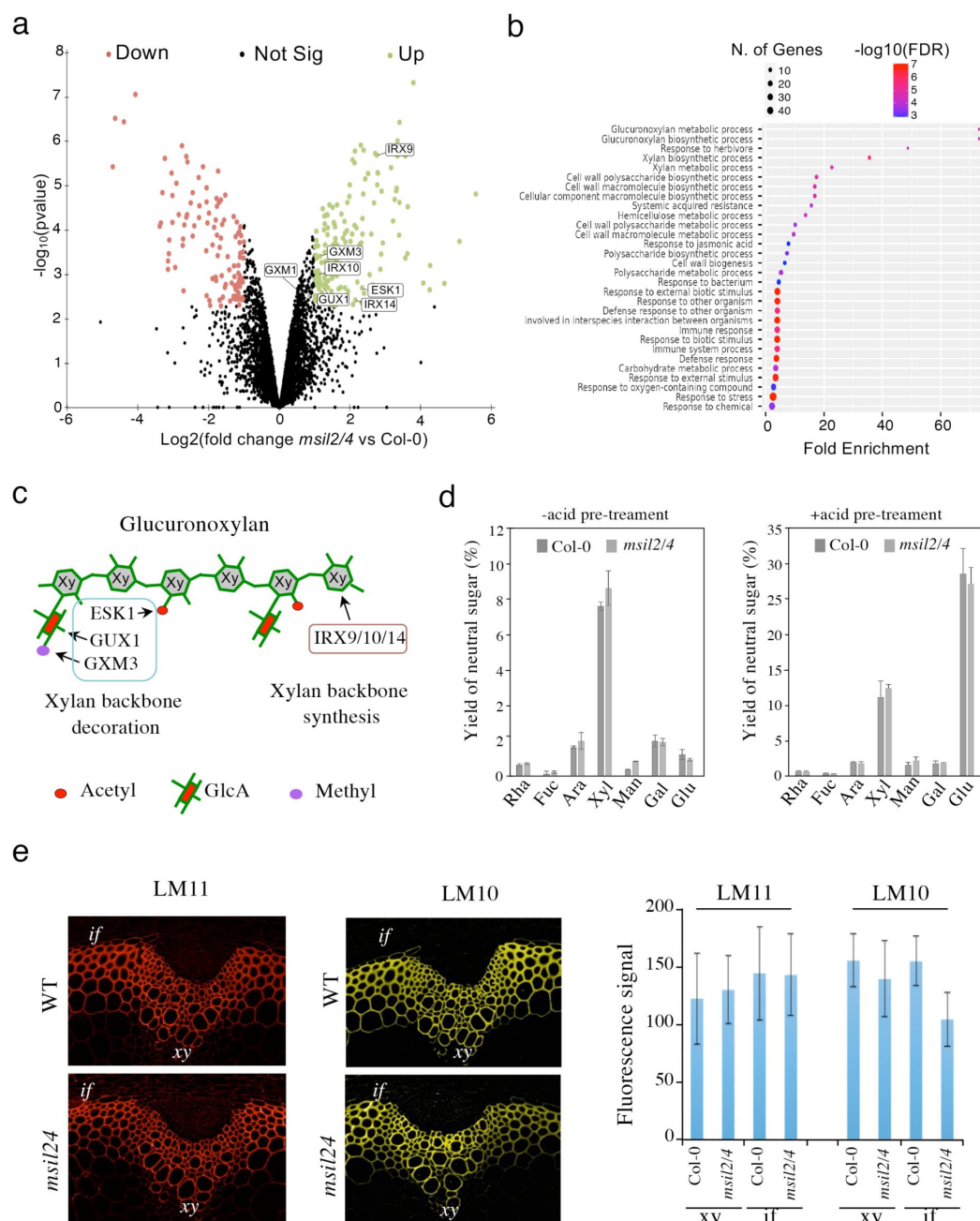


Fig. 5

The accumulation and activity of the glucuronoxylan decoration machinery are altered in *msil2/4* mutant.

a) MS-based quantitative comparison of *Col-0* and *msil2/4* inflorescence stem proteomes. Volcano plot displaying the differential abundance of proteins in *Col-0* and *msil2/4* cells analyzed by MS-based label-free quantitative proteomics. The volcano plot represents the $-\log_{10}(\text{limma p-value})$ on y axis plotted against the $\log_2(\text{Fold Change } msil2/4 \text{ vs } Col-0)$ on x axis for each quantified protein. Green and red dots represent proteins found significantly enriched respectively in *msil2/4* and *Col-0* Arabidopsis stems ($\log_2(\text{Fold Change}) \geq 1$ and $-\log_{10}(\text{p-value}) \geq 2.29$, leading to a Benjamini-Hochberg FDR = 1.01 %). The up-regulated proteins involved the glucuronoxylan biosynthetic pathways are indicated. b) Gene Ontology (GO) analysis of the proteins that are significantly up-regulated in the *msil2/4* mutant. The gene ontology analysis of DEGs was performed using ShinyGO v0.76 software. Lollipop diagrams provide information about GO fold

enrichment, significance (FDR in log10), and number of genes in each pathway. c) Schematic model of glucuronoxylan substitution patterns in Arabidopsis and the enzymes involved. IRX9/10/14, glycosyltransferases involved in the synthesis of xylan (Xy) backbone. ESK1, eskimo1. GUX1, glucuronic acid substitution of xylan1; GXM3, glucuronoxylan methyltransferase3 are involved in the glucuronoxylan decoration. d) Neutral monosaccharide composition of alcohol insoluble residue (AIR) extracted from inflorescence stems of wild-type and *msil2/4* mutant plants that have been pre- hydrolyzed or not with acid. Rha, rhamnose; Fuc, fucose; Ara, arabinose; Xyl, xylose; Man, mannose; Gal, galactose; Glc, glucose. e) Left: Immunofluorescence labeling of transverse sections of Col-0 and *msil2/4* stems with the LM10 and LM11 antixylan antibodies. Xy, xylem fibers; if, interfascicular fibers. Right: Quantification of the fluorescence was done using ImageJ software and processed according to Supplementary table 4.

The outcomes of our histological and proteomic analyses suggest potential variations in polysaccharide content and/or composition in the SCW of *msil2/4* mutant. To address this point, we first performed glycome profiling on either untreated (-Tr) or H₂SO₄-treated (+Tr) stem samples. The levels of glucose (Glc) and xylose (Xyl), which are the building blocks of cellulose and xylan polysaccharides in the SCW were not substantially affected in the *msil2/4* mutant (Fig. 5d). Similarly, no significant changes were observed in the levels of the primary cell wall-related monosaccharides, galactose, rhamnose, arabinose, mannose, fucose and galacturonic acid (Fig. 5d), indicating that the MSIL2/4 mutations did not significantly altered cellulose or glucuronoxylan levels in the SCW.

To assess possible changes in glucuronoxylan decoration in *msil2/4* mutant, we performed cytochemical analysis on Col-0 and *msil2/4* stem sections using two antixylan antibodies (LM10 and LM11), whose binding specificities depend on the degree of glucuronoxylan substitution³⁶. LM11, which binds with similar affinity both unsubstituted and substituted glucuronoxylan, shows a similar signal in the interfascicular and xylary fibers of Col-0 and *msil2/4* mutant stems, confirming that the glucuronoxylan content was not significantly modified in *msil2/4* compared to Col-0 (Fig. 5e). In contrast, LM10, which binds less efficiently the substituted glucuronoxylan³⁶, shows a significant signal decrease in the interfascicular fibers of *msil2/4* mutant (Fig. 5e; and Supplementary Table 4). Together, our data indicate that the MSIL2/4 mutations impact the accumulation and activity of the glucuronoxylan decoration machinery in the interfascicular fiber cells of Arabidopsis stem.

MSILs restrain the degree of 4-O-methylation of glucuronoxylan in Arabidopsis

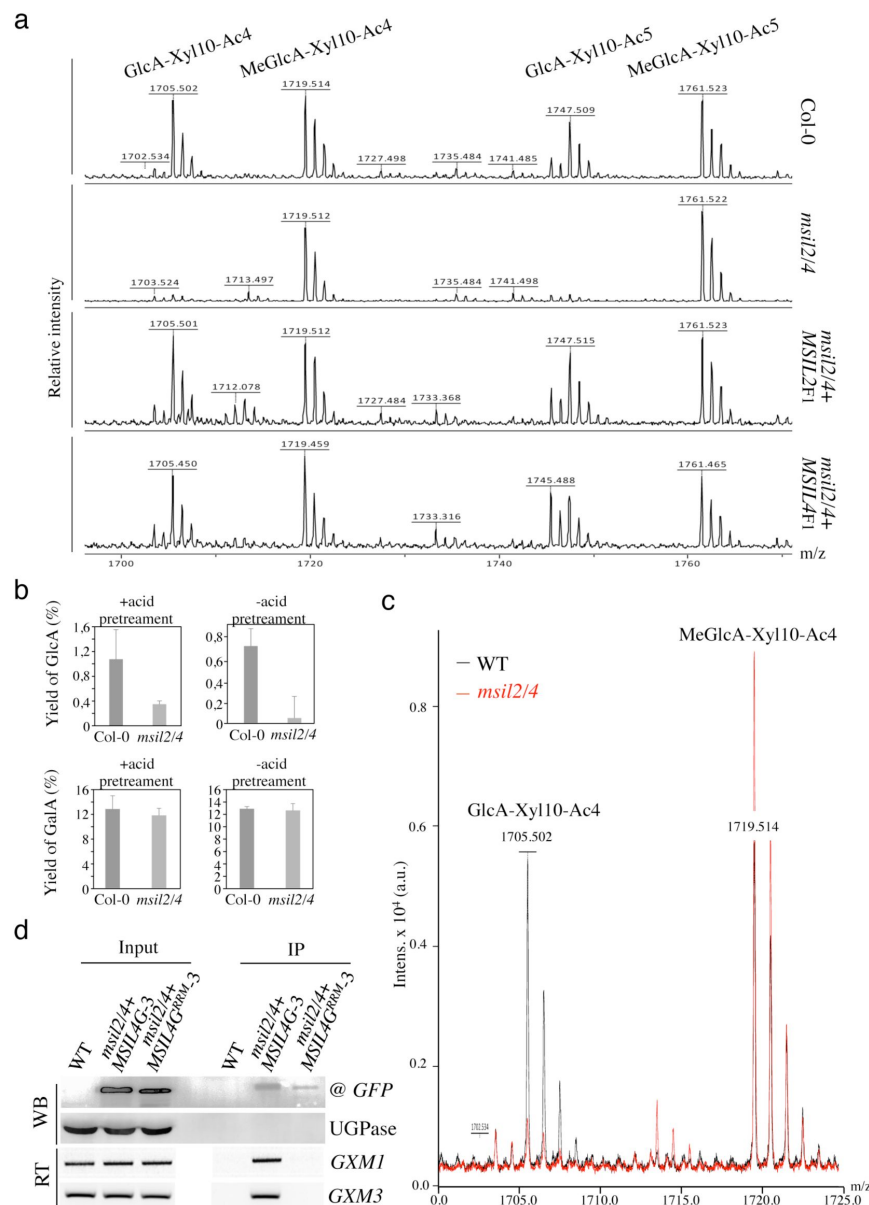
In Arabidopsis, the glucuronoxylan backbone is decorated predominantly by acetyl, GlcA, and MeGlcA groups, that are deposited in a regular and controlled manner by a specific enzymatic machinery^{2, 37, 38}. Interestingly, the ESK1, GUX2, and GXM3 enzymes, which are involved in the deposition of the acetyl and GlcA groups, as well as the methylation of GlcA group, respectively, are over-accumulated in the *msil2/4* stem proteomics, providing a possible explanation for the modification in glucuronoxylan decoration observed in the interfascicular fiber cells of the *msil2/4* mutant. To investigate further the variations in glucuronoxylan decoration incurred by the *msil2/4* mutation, we analyzed the patterns of xylanase-released SCW oligosaccharide using matrix-assisted laser desorption ionization-time-of-flight mass spectrometry (MALDI-TOF)³⁹. Consistent with previously observations^{40, 41}, MALDI-TOF analysis of Col-0 glucuronoxylan oligosaccharides showed the release of xylo-oligosaccharide pairs, evenly spaced by their degrees of polymerization and acetylation, bearing a GlcA (m/z 1705.5/1747.5) or methylated GlcA (m/z 1719.5/1761.5) substitution group (Fig. 6a). Interestingly, the level of GlcA branched xylo-oligomers was strongly decreased in the *msil2/4* double mutant, and this whatever the complexity of the oligomers considered (Fig. 6a; and Supplementary Fig. 10a). Importantly, the molecular

phenotypes observed in the *msil2/4* mutant were rescued in both *msil2/4*+*MSIL2F1* and *msil2/4*+*MSIL4F1* complemented lines, confirming the specificity of the mutation (Fig. 6a). In agreement with the MALDI-TOF outcome, quantification of the acidic sugar levels in untreated (-Tr), or H₂SO₄-treated (+Tr) WT and *msil2/4* stem samples confirmed that the release of GlcA moiety was strongly reduced in *msil2/4* double mutant, as opposed to the galacturonic acid/GalA group located on pectins (Fig. 6b).

Fig. 6

MSILs restrain the degree of 4-O-methylation of glucuroxylan in Arabidopsis inflorescence stem.

a) MALDI-TOF mass spectra of xylooligosaccharides generated by xylanase digestion of xylan from Col-0, *msil2/4*, and complemented *msil2/4*-*MSIL2F1* and *msil2/4*-*MSIL4F1* inflorescence stem materials. The ions at m/z 1705/1747 and 1719/1761 correspond to acetylated xylo-decapolysaccharides bearing a GlcA residu (GlcA-Xyl10-Ac4//GlcA-Xyl10-Ac5) or a methylated GlcA residue (MeGlcA-Xyl10-Ac4//MeGlcA-Xyl10-Ac5). b) Acidic monosaccharide composition of alcohol insoluble residue (AIR) extracted from inflorescence stems of wild-type and *msil2/4* mutant plants that have been pre- hydrolyzed or not with acid. GalA, galacturonic acid; GlcA, glucuronic acid. c) Relative changes in unmethylated/methylated GlcA de-capolysaccharide ratio in Col-0 (black) and *msil2/4* mutant (red) as controlled by the addition of external spike-in control corresponding to a pentaacetyl-chitopentaose. d) GFP-based RNA-IP assays. Top panel (WB): western blots performed using antiGFP or antiUGPase antibodies on protein fractions from inputs



or antiGFP immunoprecipitates from WT and the *msil2/4* mutant plants expressing the WT (*MSIL4G-3*) or RRM mutant (*MSIL4G^{RRM-3}*) *MSIL4*-GFP fusions. Bottom panel (RT) : Corresponding RT-PCR using *GXM1* or *GXM3* specific primers on RNA fractions.

It has been previously proposed that the intermediate level of glucuronoxylan methylation observed in Arabidopsis is due to an intrinsic rate-limiting activity of GXM enzymes in vivo⁴². We noticed that in addition to GXM3, a second member of the GXM family, GXM1⁴², tends also to be over-expressed in *msil2/4* stem although at a fold change of 1.77 slightly

below the fixed significance threshold in the proteomic analysis (Fig. 5a; and Supplementary Table 3). This observations make us wonder whether the apparent decrease of GlcA branched xylooligosaccharides in *msil2/4* could be due to an overmethylation of the GlcA substituent. To address this question, we performed the xylanase digestion and MALDI-TOF analysis in presence of an added spike-in oligosaccharide control (Pentaacetyl-chitopentaose) for normalization of the MALDI-TOF data (Supplementary Fig. 10a). Upon normalization, the MALDI-TOF data indicated that the decrease of the peaks corresponding to the GlcA-branched xylooligosaccharides was correlated with a increase in corresponding meGlcA branched xylooligosaccharides, (Fig. 6c; and Supplementary Fig. 10b; and Supplementary Table 5), suggesting that the MSIL2/4 mutations unleash the activity of the GXM1/3 in stem. RNA immunoprecipitation assays further confirmed that the *GXM3* and *GXM1* mRNAs are efficiently pulled-down from MSIL4G-3, but not MSIL4G^{RRM}-3 plant extracts (Fig. 6d), suggesting that the MSIL2/4 interaction network controls the 4-O-methylation of glucuronoxylan by repressing the activity of the GXM1/3 mRNAs in stem.

Discussion

RNA-binding proteins (RBPs) are essential components of the gene regulation machinery that govern the fate and expression of cellular RNA at post-transcriptional levels, including processing, splicing, base modification, translation and degradation¹³. Sequence-based bioinformatic, reverse and forward genetic, and affinity-based proteomic analyses have converged to show that plants, like animals and fungi, harbor a large diversity of RBPs whose functions in gene regulation and plant development are ill-defined¹⁴. In this study, we have identified an RRM motif-containing RBP family, hereafter named as MSIL, whose members share sequence, structural and functional similarities with the animal translational regulator Musashi/MSI. We found that two of the four MSIL members, MSIL2 and MSIL4, function redundantly to control various aspects of Arabidopsis development, including leaf senescence, morphology of rosette leaves, and the rigidity of the inflorescence stem. We provide evidence that the RRM domains of the MSIL2/4 proteins exhibit RNA-binding ability and are essential for the activity of these proteins *in vivo*, an observation that is corroborated by the presence of the MSIL2/4 proteins in an experimentally determined Arabidopsis mRNA-binding proteome¹⁴. Further supporting a role for MSIL2/4 in mRNA-related transactions, affinity purification-mass spectrometry analysis revealed that these proteins share a common network of protein-protein interactions with well-known mRNA-binding proteins, including the PAB2/4/8, RBP45/47, and the translation factor EIF4G. Notably, the animal polyA-binding protein/PABP (homolog of the plant PABs) was previously shown to be a functional partner of MSI that is essential for its activity in translation repression^{18, 26}. Interestingly, the MSIL2/4-associated proteins also include the phylogenetically related RBGD2/4 proteins, two constitutively expressed RBPs that are involved in heat resistance in Arabidopsis²⁰. The comparison of the MSIL2/4 and RBGD2/4 interactomes indicates only a partial overlap, suggesting a certain level of subfunctionalization. In accordance with this observation, the MSILs and RBGDs clearly diversify in function, since the *rbgd2/4* double mutant exhibits no developmental phenotypes under normal growth conditions.

A penetrant developmental phenotype associated with the MSIL2/4 mutations is a loss in stem rigidity that we could trace back to a defect in SCW formation, an important reservoir of fixed carbon and a renewable resource of major environmental and economic importance¹. Yet, with the exception of few microRNAs involved in the control of lignin polymer synthesis^{6, 15, 16}, no information about RBPs-mediated posttranscriptional regulation of SCW biosynthetic genes was previously reported. In this study, we show that MSIL2 and MSIL4, function redundantly to promote SCW formation in the interfascicular

fibers and that their ability to bind RNA is essential to fulfill this function. Our conclusions are supported by independent approaches which highlight changes of SCW thickness, lignin content, and xylan decoration profile in the interfascicular fiber cells of the *msil2/4* mutant. The fiber-specific effect of the MSIL2/4 mutations on SCW formation, is further supported by our evidence showing that the *msil2/4* mutant grows normally and lacks the xylem collapses usually observed in more pleiotropic SCW deficient mutants. Previous studies have shown that specific master transcriptional regulators, such as NST1 and NST3, control cell-lineage specific formation of SCW formation in Arabidopsis^{5, 31}. Our study extend this observation by showing that RBP-mediated post-transcriptional mechanisms are also involved in cell-type specific control of SCW formation in Arabidopsis. Notably, the presumed role of MSIL2/4 in interfascicular fiber cells contrasts with the rather constitutive expression of the *MSIL2/4* genes, suggesting that their activity depends on cell-context. This is in accordance with the reported cell-type specific activities of ubiquitous animal MSI2 in cell renewal, cell differentiation and cancer control^{18, 19}. Understanding how the activity of the *MSIL2/4* genes is controlled by the cellular context will be essential from a basic point of view, providing a unique model for studying RBP regulation in specific cell type in plants.

Our data indicate that the accumulation of gene transcripts of SCW biosynthesis pathway enzymes and regulators are unaffected in *msil2/4*, suggesting that MSIL2/4 proteins are likely to act at a translational level to regulate the SCW formation in interfascicular fiber cells, an observation that is supported by the nature of the MSIL2/4 interactome. In this regard, the stem proteomics gives us a glimpse into the changes in protein abundance that could account for the observed *msil2/4* SCW phenotype. Indeed, our MS data reveal that the glucuronoxylan biosynthesis machinery, including enzymes involved both in the synthesis and substitution of the xylan backbone, is specifically over-accumulated in the *msil2/4* double mutant. The molecular specificity of the observed glucuronoxylan phenotype with respect to the changes in glucuronoxylan enzymatic machinery content raises several interesting questions about the impact of the MSIL2/4 mutations on the SCW formation. First, we noticed that the outcome of our proteomic data contrasts with the observed normal level of glucuronoxylan deposition in the interfascicular fibers SCW, suggesting that independent regulatory mechanisms act downstream of MSIL2/4 to control final glucuronoxylan content in the SCW. One possible explanation for this observation could be that the activities of the GUX, ESK, IRXs glucuronoxylan biosynthesis enzymes, in contrast to that of GXM3, are not rate-limiting for glucuronoxylan synthesis *in vivo*. Alternatively, and in agreement with studies that have reported that glucuronoxylan and cellulose contents are correlated *in vivo* in xylan-deficient mutants⁴³, one could imagine that cellulose is the primary determinant of the content in glucuronoxylan in the SCW, and that the glucuronoxylan produced in excess in the golgi compartment of the *msil2/4* mutant is targeted for degradation upon secretion if unbound to cellulose. In contrast, the concomitance between the glucuronoxylan hypermethylation and the GXM3 overaccumulation phenotypes observed in *msil2/4* clearly support the previous assumption that the GXM-mediated methylation is a rate-limiting step in glucuronoxylan that can be overcome by GXM protein overexpression⁴². Our data support the idea that MSIL2/4 play a specific role in the developmental control of the GXM protein accumulation/activity, actively restraining the level of glucuronoxylan methylation in the Arabidopsis interfascicular fiber cells. We currently posit that MSIL2/4 could control the level of the GXM enzymes by directly regulating the translation of the GXM mRNAs or by indirectly regulating the translation of an yet unknown translational regulator of the glucuronoxylan biosynthesis pathway (XTRe) (Fig. 7). Future biochemical and functional studies will be necessary to precisely understand the mechanism by which MSIL2/4 proteins impact gene expression during SCW synthesis in Arabidopsis.

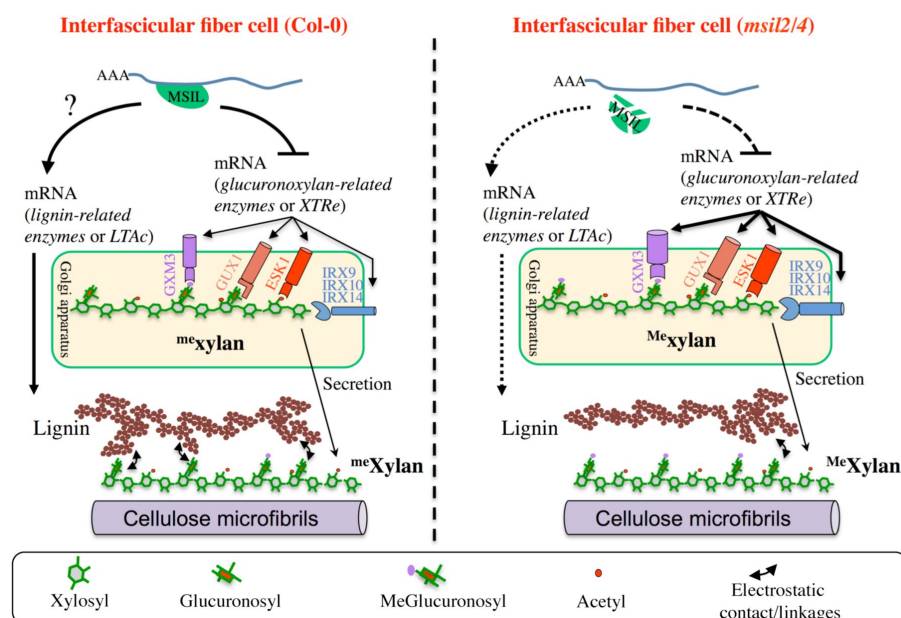


Fig. 7

Model of MSIL-dependent control of glucuronoxylan methylation in Arabidopsis and its consequence for SCW architecture.

In the interfacial fiber cells of Col-0, MSIL2/4 restrain the translation of the glucuronoxylan biosynthesis machinery, including the rate-limiting GXM3 enzyme. This activity would keep the level of glucuronoxylan methylation at an intermediate level, therefore providing a biochemical environment that favors the interactions between the glucuronoxylan and lignin polymers.

In the interfacial fiber cells of *msil2/4* mutant, the translation of the glucuronoxylan biosynthesis machinery, including GXM3, is increased, leading to the deposition of an over-methylated form of glucuronoxylan that has detrimental effects on glucuronoxylan-lignin interaction and the SCW formation. In a non exclusive manner, MSIL2/4 could also have a positive role in lignin synthesis, whose defect in *msil2/4* would lead to a decrease in lignin content. The model was inspired from Grantham et al.³⁷.

Our data indicate that, in addition to the reported change in glucuronoxylan methylation, the MSIL2/4 mutations is also associated with a specific decrease of lignin content in the SCW of the interfacial fiber cells, whose origin remains unclear. On the basis of the previous observations, we envision two possible mechanisms by which the MSIL2/4 mutations could impact the lignin deposition in the fiber cells. In the first model, we propose that MSIL2/4 positively control the lignin synthesis pathway in the interfacial cells by promoting the translation of lignin biosynthesis genes or that of a putative lignin translational activator (LTAc) (Fig. 7). In this regard, although no GO enrichment could be observed for lignin biosynthesis related genes in the proteomic analysis, two lignin biosynthetic enzymes, CAD4 and PAL2, are significantly down-accumulated in *msil2/4* mutant (Supplementary Table 3), an observation that could account for the observed decrease in lignin content. However, previous published biochemical and genetic analyses do not support this conclusion, as the single *cad4* and *pal2* knock-out mutants display no significant alteration of lignin content due to genetic redundancy^{9, 44}. Moreover, thioacidolysis analyses showed that the ratio or degree of polymerization of the different lignin monomers remain unchanged in the *msil2/4* mutant, reinforcing the idea that the lignin biosynthesis pathway is functional in this mutant, leaving open the question of the specificity of MSIL2/4 action on the lignin pathway (Fig. 7). An alternative explanation for the decrease in lignin phenotype observed in the *msil2/4* mutant would be that it results from an indirect cascade effect primarily triggered by the *msil2/4*-dependent changes in glucuronoxylan methylation. In this regard, it has been proposed that the 4-O-methylation of GlcA substituent results in the xylan being more hydrophobic, possibly impacting its non-covalent interaction with lignin and the capacity of lignin to assemble in the SCW^{41, 45, 46}. Meanwhile, the hypermethylation of the carbon-4 hydroxyl of the GlcA side chains could affect the capacity of glucuronoxylan to form ether linkages to lignin (Fig. 7). Indeed, a study

recently reported the existence of an ether linkage between a carbon-6 hydroxyl of a mannosyl residue of a galactoglucomannan hemicellulose and lignin in pine wood⁴⁷. Although xylan has been proposed to bind covalently to lignin via ether bonds involving xylose or Arabinose residues, it remains unclear to date whether glucuronoxylan can form GlcA-dependent ether linkages with lignin, and the functional relevance of such linkages remains to be demonstrated⁴⁸.

Methods

Plant material

Arabidopsis thaliana Col-0 and mutants *msil1-1* (GABI_462A04), *msil2-1* (Salk_066670), *msil3-1* (Salk_002477) and *msil4-1* (Salk_094167) obtained from the SALK Institute genomic analysis laboratory were used in this study. PAB2-RFP/PAB2R line has been provided by Cécile Bousquet-Antonelli (LGDP, Université Perpignan). The primers used to genotype lines are listed in Supplementary Table 4. Double mutants and quadruple mutants were generated by crossing plants. Plants were either grown in soil or cultivated *in vitro* on plates containing 2.20 g/l synthetic Murashige and Skoog (MS) (Duchefa) medium and 0.5g/l MES, pH 5.7, and 7 g/l agar. To breakdown dormancy, seeds were incubated for 48 h at 6°C in the dark. Germination and development were performed in growth chambers, at 20°C, 60–75% hygrometry with a 16-h light/8-h dark photoperiod (100 $\mu\text{E m}^{-2} \text{s}^{-1}$ light [fluorescent bulbs with white 6500K spectrum, purchased from Sylvania]). For *in vitro* culture, the seeds were surface-sterilized before been sown on plates, incubated for 48 h at 6°C in the dark, and placed in a growth cabinet at 20°C with a 16-h-d/8-h-dark cycle and 120 $\mu\text{E m}^{-2} \text{s}^{-1}$ light (LEDs with white 4500K spectrum, purchased from Vegeled).

Transgenic lines construction

MSIL2F-and MSIL4F-tagged versions were produced from Phusion-generated genomic DNA PCR products (promoter and coding region) using primers TL976(SalI)-TL924(PstI) for MSIL2F and TL974(SalI)-TL975(BamHI) for MSIL4F. PCR products were cloned in the CTL235 binary vector containing a Flag-HA tag in front of a NosT terminator and the hygromycin resistance gene driven by a 35S promoter. To obtain the MSIL2G and MSIL4G constructs used for fluorescence microscopy, the same PCR products were cloned in the CTL579 binary vector in front of a EGFP tag. The obtained plasmids were then used to transform *mg12/4* plants via *Agrobacterium* transformation. MSIL4- GFP/MSIL4G-tagged version used for complementation and RIP corresponds to the fusion of a genomic PCR product containing *MSIL4* promoter (primers TL3527(HindIII)-TL3528(SalI)) fused with a second PCR cDNA fragment (primers TL3529(SalI)-TL3530(BamHI)) cloned initially in pGEM T Easy (Promega). After sequencing the fusion DNA has been cloned in the binary vector CTL579 containing GFP cDNA sequence. The MSIL4G-tagged construct was finally introduced in *msil2/4* plants via *Agrobacterium* transformation. The rrm mutated MSIL4G Δ r version was produced as MSIL4G except that the cDNA fragment has been gene-synthesized (Genecust) to modify RNA binding ability by having aspartic instead of phenylalanine at position 9, 49 and 51 in RRM1 domain and in position 111, 151 and 153 in RRM2 domain. Transgenic seeds selection was performed in the presence of 30 $\mu\text{g/l}$ hygromycin and the progenies were screened by PCR.

Protein extraction and immunodetection

Total plant extracts (up to 100 mg) were ground in liquid nitrogen and proteins were extracted according to Hurkman and Tanaka⁴⁹. Before migration, SDS–PAGE loading buffer

was added and Coomassie staining was used to calibrate loadings. Proteins were separated on SDS/ PAGE gels and blotted onto Immobilon-P PVDF membrane (Cat. No. IPVH00010; MerckMillipore). Protein blot analysis was performed using the Immobilon Western Chemiluminescent HRP Substrate (Cat. No. WBKLS0500; MerckMillipore). Specific MSIL antibodies were raised in rabbits against defined epitopes (Epitope MSIL1 (SCDGTSSSTFGYNRIPS), MSIL2 (RLQEYFGKYGDLVE), MSIL3 (GYGVKPEVRYSPAVGN) and MSIL4 (TWRSPTEGAPF) (Eurogentec).

Protein production and purification

The MSIL4 RRM binding domains were amplified using primers TL4128 and TL4129 either from MSIL4G construct for the wild type version or from MSIL4GΔr for the mutated version. PCR products were then cloned in pET41 to obtain His-tagged RRM domains. Tagged proteins were produced by induction in *Escherichia coli* BL21 cells. Cultures of 200 mL of an ampicillin-resistant (100 mg mL⁻¹) colony were grown at 37°C and induced by 100mM isopropyl-b-D-galactopyranoside in the exponential phase (optical density 0.5 at 500 nm). After induction, bacteria were harvested by centrifugation and pellets were resuspended in 5 ml of binding buffer (His-bind buffer kit (Millipore), proteins were extracted using a Constant cell disrupter system (Constant Systems) with a disruption pressure of 2.35kbar and then purified using His-bind buffer kit (Millipore) following manufacturer's recommendations.

RNA binding assays

Binding assays were performed using homopolymer RNA conjugated to CNBr-activated Sepharose beads (Sigma **C9142**). Either poly(A), poly(G), poly(C) or poly (U) from Mercks (respectively P9403, P4903, P4404 and P9528) were used following Sigma's recommendations. 0.7 to 1 mg of RNA were used for 200 mg sepharose beads. 0.45 mg of purified protein (MSIL4 RRM binding domain or mutated version) in 0.5 ml binding buffer (100 mM NaCl, 10 mM Tris-HCl, pH 7.5, 2.5 mM MgCl₂, 2.5 % Triton X-100) were incubated with 5 mg RNA conjugated Sepharose beads for 1 hour. After incubation, beads were washed 5 times for 10 min each in 0.5 ml of binding buffer. Beads were then pelleted at 12,000 rpm, resuspended in SDS sample buffer, boiled for 20 min, and pelleted again. Supernatants were transferred to fresh tubes and loaded on a 12% SDS-PAGE minigel for electrophoresis (Bio-Rad) and protein blot analysis.

RNA immunoprecipitation

RIP was performed as described previously in Merret et al.⁵⁰ using 400 mg of stem powder as starting material.

Expression analysis

Total RNA was isolated using the TRI re-agent (Cat. No. TR-118; Euromedex) according to the manufacturer's instructions. Genomic DNA was then digested out using the RQ1 RNase-Free DNase (Cat. No. M6101; Promega). cDNAs were obtained from 400 ng of RQ1-treated RNA using 1 µl of GoScript Reverse Transcriptase (Cat. No. A5003; Promega) in a 20-µl final volume reaction using random primers (Cat. No. C1181; Promega) in the presence of 20 units of RNasin (Recombinant Ribonuclease Inhibitor, Cat. No. C2511; Promega) and dNTP mix at a final concentration of 0.5 mM of each dNTP (Cat. No. U1511; Promega). Semi-quantitative RT-PCR amplifications were performed on 1 µl of cDNA in a 12.5-µl reaction volume to start with. The amplification of EF1α transcripts was used to equilibrate. qRT-PCR was performed

on a Light Cycler 480 II machine (Roche Diagnostics) by using the Takyon No ROX SYBR MasterMix blue dTTP kit (Cat. No. UF-NSMT-B0701; Eurogentec). Each amplification reaction was set up in a 10- μ l reaction containing each primer at 300 nM and 1 μ l of RT template in the case of total RNA with a thermal profile of 95°C for 10 min and 40 amplification cycles of 95°C, 15 s; 60°C, 60 s. Relative transcript accumulation was calculated using the $\Delta\Delta$ Ct methodology⁵¹, using ACTIN2 as internal control. Average $\Delta\Delta$ Ct represents three experimental replicates with standard errors. PolyA⁺ RNAs were purified from RNA obtained by the TRI Reagent extraction, using the PolyATtract mRNA Isolation System III (Cat. No. Z5310; Promega) according to the manufacturer's instructions. cDNAs were then synthesized according to the above protocol using 50 ng of polyA⁺ RNA as starting material. Semiquantitative and quantitative RT-PCRs were performed as described above with a 1/25 dilution of cDNAs from polyA⁺ RNA as template for qRT-PCR. The primers used for RT and qRT-PCR are listed in Supplementary Table 6.

RNA Sequencing

Total RNA were extracted from the basal stem section of mature Col-0 and *msil2/4* mutant plants using the RNeasy Plant Mini Kit (Qiagen, Cat.#74904) and treated onto the column with the RNase-Free DNase Set (Qiagen, Cat.#79254). Two replicates were performed per plant lines. PolyA purification, library preparation using the Illumina TruSeq stranded mRNA kit and library quality controls were performed by the BioEnvironment Illumina Platform. The mRNA sequencing of Col-0 and *msil2/4* mRNAs was done with single-reads, 1x 75pb and 30–35 $\times$ 10⁶ reads per library.

RNASeq analysis

For each library, 30–35 $\times$ 10⁶ reads were obtained with 85–90% of the bases displaying a Q-score $\geq$ 30, with a mean Q-score of 38 as assessed with Fastqc (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Reads were mapped against the TAIR 10 genome using gtf annotation file and default parameters of TopHat2 v2.0.7⁵². Assembly and transcript quantification were performed with Cufflinks v2.2.1⁵³. Finally, the low expressed transcripts -less than 1 RPKM (reads per kilobase per million mapped reads) in one of the libraries were filtered out. The differential analysis was conducted by Cuffdiff software, belonging to the Trapnell suite. Foldchange (FC) was determined as the ratio between the normalized read counts between mutant and wild-type.

Fluorescence detection and tissue colorations

GFP detections were performed on *in vitro* grown one week seedling roots fixed in fixation buffer (50mM Pipes, 5mM EGTA, 5mM MgSO₄, pH7 added with 2% para-formaldehyde and 0.2% Triton). After 5 min vacuum infiltration and 10 min agitation at room temperature, roots were washed twice with fixation buffer and stored at 4°C until confocal observation. Confocal imaging was done using an Zeiss Axio Observer Z1 LSM 700. Roots were observed in water and for GFP detection samples were excited at 488 nm and the signal captured at 495–555 nm. For RFP detection, excitation was performed at 555 nm and the signal captures at 600–700 nm. Images were analysed with ImageJ. At least 15 samples per line were analyzed.

Stem coloration were performed from 5 cm length basal stem pieces conserved in ethanol 80%. Stems were rehydrated first in ethanol 50% and later in water and then embedded in agarose 5% before being transversally cut using a vibratome (Leica vt 1000s) to obtain 120 micron thickness sections stored in water. The presence of lignin was determined by staining with phloroglucinol for 30 seconds, giving a red coloration in the presence of lignin

cinnamaldehyde groups. Sections were observed using an inverted microscope (Leitz DMRIE, Leica Micro-systems, Wetzlar, Germany) and images were registered using a CCD camera (Color Collview, Photonic Science, Milham, UK).

For cellulose detection, the brightener Calcofluor White ST [4,4'-bis(anilino-6-bis(2-hydroxyethyl)-amino-s-triazine-ylamino)-2,2'-stilbene disulfonic acid; Cyanamide Co., Bound Brook, N.J.] was used at a final concentration of 0.1%. The samples were incubated with calcofluor for 3 min, washed three times with water before being observed. For detection, excitation was performed at 405 nm and the signal captures at 434–496 nm. For auto-fluorescence analysis stem sections were excited at 405 nm and the signal captures at 410–470 nm for blue. Pictures have been analyzed with LAS X (Leica Application Suite X).

Biochemical analyses of SCWs

Arabidopsis stems (5 cm sections at the bottom part of the stems) were harvested from Col0 and *msil2/4* double mutant as four independent biological replicates. Soluble extractives were eliminated as described by Ployet et al.⁵⁴. Briefly, stems were freeze dried for 48h, ground using a Mixer Mill MM 400 (Retsch), and extracted by hot solvents (successively water, ethanol, ethanol/toluene (1:1 v/v) and acetone) to obtain extractive-free stem residues (ESR). The determination of acetyl bromide (AcBr) lignin was performed on 10mg of ESR as described by Ployet et al.⁵⁵. All analyses were done in triplicate. Fourier transform infrared spectroscopy (FT-IR) analysis was performed on 100–200 mg of ESR. Spectra were recorded from ten technical replicates, in the range 400–4000 cm⁻¹ with a 4 cm⁻¹ resolution and 32 scans per spectrum using an attenuated total reflection (ATR) Nicolet 6700 FT-IR spectrometer (Thermo Fisher, Illkirch-Graffenstaden, France) equipped with a deuterated-triglycine sulfate (DTGS) detector. Spectra analyses were performed as described in Dai et al.⁵⁶. Saccharification yield was estimated with or without alkali pretreatment, as described by Van Acker et al.⁵⁷. Reducing sugar concentration was assessed with dinitrosalicylic acid (DNS) reagent using 10µl of the supernatant after 9 h of reaction. Enzyme activity was assessed at 0.25 filter paper units (FPU) ml⁻¹ in our conditions.

Mass spectrometry (MS)-based proteomic analyses

Proteins from total inflorescence stem extracts of three biological replicates of Col-0 and *msil2/4* mutant plants were solubilized in Laemmli buffer and heated for 10 min at 95°C. They were then stacked in the top of a 4–12% NuPAGE gel (Invitrogen), stained with Coomassie blue R-250 (Bio-Rad) before in-gel digestion using modified trypsin (Promega, sequencing grade) as previously described⁵⁸. The resulting peptides were analyzed by online nanoliquid chromatography coupled to MS/MS (Ultimate 3000 RSLCnano and Orbitrap Exploris 480, Thermo Fisher Scientific) using a 180-min gradient. For this purpose, the peptides were sampled on a precolumn (300 µm x 5 mm PepMap C18, Thermo Scientific) and separated in a 75 µm x 250 mm C18 column (Aurora Generation 2, 1.6µm, IonOpticks). The MS and MS/MS data were acquired using Xcalibur 4.4 (Thermo Fisher Scientific).

Peptides and proteins were identified by Mascot (version 2.8.0, Matrix Science) through concomitant searches against the Viridiplantae database (from Uniprot, Arabidopsis thaliana (thale-cress) taxonomy, 136447 sequences) and a homemade database containing the sequences of classical contaminant proteins found in proteomic analyses (human keratins, trypsin..., 126 sequences). Trypsin/P was chosen as the enzyme and two missed cleavages were allowed. Precursor and fragment mass error tolerances were set at respectively at 10 and 20 ppm. Peptide modifications allowed during the search were: Carbamidomethyl (C, fixed), Acetyl (Protein N-term, variable) and Oxidation (M, variable). The Proline software⁵⁹ (version 2.2.0) was used for the compilation, grouping, and filtering of the results

(conservation of rank 1 peptides, peptide length ≥ 6 amino acids, false discovery rate of peptide-spectrum-match identifications $< 1\%$ ⁶⁰, and minimum of one specific peptide per identified protein group). MS data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository⁶¹. Proline was then used to perform a MS1 label-free quantification of the identified protein groups based on razor and specific peptides.

Statistical analysis was performed using the ProStaR software⁶² based on the quantitative data obtained with the four biological replicates analyzed per condition. Proteins identified in the contaminant database, proteins identified by MS/MS in less than two replicates of one condition, and proteins quantified in less than three replicates of one condition were discarded. After log2 transformation, abundance values were normalized using the variance stabilizing normalization (vsd) method, before missing value imputation (SLSA algorithm for partially observed values in the condition and DetQuantile algorithm for totally absent values in the condition). Statistical testing was conducted with limma, whereby differentially expressed proteins were selected using a log2(Fold Change) cut-off of 1 and a p-value cut-off of 0.00513, allowing to reach a false discovery rate close to 1% according to the Benjamini-Hochberg estimator. Proteins found differentially abundant but identified by MS/MS in less than two replicates, and detected in less than three replicates, in the condition in which they were found to be more abundant were invalidated (p-value = 1).

MALDI-TOF MS analyses of the patterns of xylanase-released SCW oligosaccharide

10 mg of dried alcohol insoluble cell wall residues were incubated with 10 U mL⁻¹ of endo-1,4-D-xylanase (E-XYLNP) from MEGAZYME previously desalted with centrifugal filters (Amicon Ultra-0.5 ml) in 1 ml during 12 hrs at 40°C. After boiling during 5 min to inactivate enzyme activity, samples were filtered with 50 kD centrifugal filters (Amicon Ultra-0.5 ml). As positive control, 2 mg of wheat AX (MEGAZYME) were treated in a similar manner. The samples were then analyzed by matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) MS using DMA/DHB (N,N-dimethylaniline/2,5-dihydroxybenzoic acid) matrix³⁹. The samples (1 μ L) were deposited and then covered by the matrix (1 μ L) on a polished steel MALDI target plate. MALDI measurements were then performed on a rapiflex MALDI-TOF spectrometer (Bruker Daltonics, Bremen, Germany) equipped with a Smartbeam 3D laser (355 nm, 10000 Hz) and controlled using the Flex Control 4.0 software package. The mass spectrometer was operated with positive polarity in reflectron mode. Spectra were acquired in the range of 560–3200 *m/z*. The detected species correspond to ions in the form of sodium adducts.

Thioacidolysis

This method is adapted from Lapierre et al.⁶³. Briefly, 10 mg of alcohol insoluble cell wall residues were incubated in 3 ml of dioxane with ethanethiol (10%), BF₃ etherate (2.5%) containing 0.1% of heinecosane C21 diluted in CH₂Cl₂ at 100°C during 4 hours. 3 ml of NaHCO₃ (0.2 M) were added after cooling and mixed prior the addition of 0.1 mL of HCl (6 M). The tubes were vortexed after addition of 3 mL of dichloromethane and the whole lower organic phase was collected in a new tube before concentration under nitrogen atmosphere to approximately 0.5 ml. Then, 10 μ L were trimethylsilylated (TMS) with 100 μ L of N,O-bis(trimethylsilyl) trifluoroacetamide and 10 μ L of ACS-grade pyridine. The trimethylsilylated samples were injected (1 μ L) onto an Agilent 5973 Gas Chromatography–Mass Spectrometry system. Specific ion chromatograms reconstructed at *m/z* 239, 269 and 299 were used to quantify H, G and S lignin monomers respectively and compared to the internal standard at *m/z* 57, 71, 85).

Cell wall sugar analysis

Neutral monosaccharide sugar content was determined by gas chromatography after acid hydrolysis and conversion of monomers into alditol acetates as described in Hoebler et al.⁶⁴ and Blakeney et al.⁶⁵ Gas Chromatography was performed on a DB 225 capillary column (J&W Scientific, Folsom, CA, USA; temperature 205 °C, carrier gas H₂). Calibration was made with standard sugar solution and inositol as internal standard.

For acid sugar analysis, the uronic acid content was automatically quantified after hydrolysis of polysaccharides in concentrated sulfuric acid (18 M) containing sodium tetraborate (12.5 mM) or not, followed by m-hydroxydiphenyl colorimetric determination^{66, 67}. Galacturonic and Glucuronic acid were used as standard for quantification.

Data availability

The high-throughput sequencing data generated in this study are accessible at NCBI's Gene Expression Omnibus (GEO) via GEO Series accession number GSE223732. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD040207 (proteomic identification of MSIL-interacting partners), and PXD040020 (proteomic analysis of total inflorescence stem extracts).

Acknowledgements

We thank Aurélie Le Ru for the histological staining and cell wall thickness measurements, Hua Cassan-Wang for helping to the preparation of stem sections for staining, and Michele Laudie for preparation of the mRNA libraries. This work was supported by the Centre National de la Recherche Scientifique (CNRS), and grants ANR-08-BLAN-0206 and 12-BSV6- 0010 from Agence National de la Recherche (ANR) to TL. This study was supported by the “Ecole Universitaire de Recherche (EUR) TULIP-GS (ANR-18-EURE-0019), and the EPIPLANT Groupement de Recherche (CNRS, France). The proteomic experiments performed by LB and YC were partially supported by Agence Nationale de la Recherche under projects ProFI (Proteomics French Infrastructure, ANR-10-INBS-08) and GRAL, a program from the Chemistry Biology Health (CBH) Graduate School of University Grenoble Alpes (ANR-17-EURE-0003). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the paper.

Author contributions

Study design: T.L., F.M., R.S., M.F., DP., and N.B.E.; mutant characterization and molecular biology: A.K., D.P., C.P., R.M., and J.A.; RNA binding analysis: N.B.E.; MSIL co-purification experiments, mass spectrometry and statistical analysis of co-IP data: D.G. and P.H.; Total inflorescence stem proteomic and statistical analysis: L.B. and Y.C.; Polysaccharide biochemical analysis: L.L-B. and R.S.; Lignin biochemical and FTIR analysis: F.M.; immunohistochemical staining: Y.M.; MALDI-TOF analysis and statistical analysis: R.S. and M.F.; writing: T.L., F.M., R.S., M.F., D.P., and N.B.E.; visualisation: T.L.; supervision: T.L. and N.B.E.

Competing interests

The authors declare no competing interests.

References

1. Meents M.J. , Watanabe Y. , Samuels A.L (2018) **The cell biology of secondary cell wall biosynthesis** *Ann Bot* **121**:1107–1125
2. Scheller H.V. , Ulvskov P. (2010) **Hemicelluloses** *Annu Rev Plant Biol.* **61**:263–289
3. Yokoyama R. , Nishitani K (2004) **Genomic basis for cell-wall diversity in plants. A comparative approach to gene families in rice and Arabidopsis** *Plant cell Physiol* **45**:1111–1121
4. Kumar M. , Campbell L. , Turner S (2016) **Secondary cell walls: biosynthesis and manipulation** *J Exp Bot* **67**:515–531
5. Taylor-Teeple M. , et al. (2015) **An Arabidopsis gene regulatory network for secondary cell wall synthesis** *Nature* **517**:571–575
6. Zhang J. , Xie M. , Tuskan G.A. , Muchero W. , Chen J.-G (2018) **Recent advances in the transcriptional regulation of secondary cell wall biosynthesis in the woody plants** *Front Plant Sci* **9**
7. Myburg A.A. , et al. (2014) **The genome of Eucalyptus grandis** *Nature* **510**:356–362
8. Zhong R. , Ye Z.-H (2014) **Complexity of the transcriptional network controlling secondary cell wall biosynthesis** *Plant Sci* **229**:193–207
9. Sibout R. , Eudes A. , Mouille G. , Pollet B. , Lapierre C. , Jouanin L. , Séguin A (2005) **CINNAMYL ALCOHOL DEHYDROGENASE-C and -D are the primary genes involved in lignin biosynthesis in the floral stem of Arabidopsis** *Plant Cell* **17**:2059–2076
10. Whitehead C. , et al. (2018) **A glycosyl transferase family 43 protein involved in xylan biosynthesis is associated with straw digestibility in Brachypodium distachyon** *New Phytol* **218**:974–985
11. Bonawitz N.D. , Chapple C (2013) **Can genetic engineering of lignin deposition be accomplished without an unacceptable yield penalty?** *Curr Opin Biotechnol* **24**:336–343
12. Wang Y. , et al. (2016) **Genetic modification of plant cell walls to enhance biomass yield and biofuel production in bioenergy crops** *Biotechnol Adv* **34**:997–1017

13. Singh G. , Pratt G. , Yeo G.W. , Moore M.J (2015) **The clothes make the mRNA: past and present trends in mRNP fashion** *Ann Rev Biochem* **84**:325–354
14. Reichel M. , et al. (2016) **In planta determination of the mRNA-binding proteome of Arabidopsis etiolated seedlings** *Plant Cell* **28**:2435–2452
15. Wang C.Y. , et al. (2014) **MiR397b regulates both lignin content and seed number in Arabidopsis via modulating a laccase involved in lignin biosynthesis** *Plant Biotechnol J* **12**:1132–1142
16. Zhao Y. , et al. (2015) **MicroRNA857 is involved in the regulation of secondary growth of vascular tissues in Arabidopsis** *Plant physiol* **169**:2539–2552
17. Chai G. , et al. (2015) **Arabidopsis C3H14 and C3H15 have overlapping roles in the regulation of secondary wall thickening and anther development** *J Exp Bot* **66**:2595–2609
18. Fox R.G. , Park F.D. , Koechlein C.S. , Kritzik M. , Reya T (2015) **Musashi signaling in stem cells and cancers** *Annu Rev Cell Dev Biol* **31**:249–267
19. Kudinov A.E. , Karanicolas J. , Golemis E.A. , Bumber Y (2017) **Musashi RNA-binding proteins as cancer drivers and novel therapeutic targets** *Clin Cancer Res* **23**:2143–2153
20. Zhu S. (2022) **Liquid-liquid phase separation of RBGD2/4 is required for heat stress resistance in Arabidopsis** *Dev Cell* **57**:583–597
21. Chavali P.L. , et al. (2017) **Neurodevelopmental protein Musashi-1 interacts with the Zika genome and promotes viral replication** *Science* **357**:83–88
22. Nguyen D.T.T. , et al. (2020) **HyperTRIBE uncovers increased MUSASHI-2 RNA binding activity and differential regulation in leukemic stem cells** *Nat Commun* **11**
23. Biasani M. , et al. (2014) **SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information** *Nucleic Acids Res* **42**:W252–258
24. Daubner G.M. , Cléry A. , Allain F.H (2013) **RRM-RNA recognition: NMR or crystallography... and new findings** *Curr Opin Struct Biol* **23**:100–108
25. Weill L. , Belloc E. , Castellazzi C.L. , Mendez R (2017) **Musashi 1 regulates the timing and extent of meiotic mRNA translational activation by promoting the use of specific CPEs** *Nat Struct Mol Biol* **24**:672–681
26. Cragle C.E. , et al. (2019) **Musashi interaction with poly(A)-binding protein is required for activation of target mRNA translation** *J Biol Chem* **294**:10969–10986
27. Scheer H. , et al. (2021) **The TUTase URT1 connects decapping activators and prevents the accumulation of excessively deadenylated mRNAs to avoid siRNA biogenesis** *Nat Commun* **12**

28. Chantarachot T. , Bailey-Serres J (2018) **Polysomes, stress granules, and processing bodies: a dynamic triumvirate controlling cytoplasmic mRNA fate and function** *Plant Physiol* **176**:254–269
29. Scutenaire J. , et al. (2018) **The YTH domain protein ECT2 is an m6A reader required for normal trichome branching in Arabidopsis** *Plant Cell* **30**:986–1005
30. Mahboubi H. , Stochaj U (2017) **Cytoplasmic stress granules: Dynamic modulators of cell signaling and disease** *Biochim Biophys Acta Mol Basis Dis* **1863**:884–895
31. Mitsuda N. , et al. (2007) **NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in Woody tissues of Arabidopsis** *Plant Cell* **19**:270–280
32. Shi D. , et al. (2021) **Tissue-specific transcriptome profiling of the Arabidopsis inflorescence stem reveals local cellular signatures** *Plant Cell* **33**:200–223
33. Herth W. , Schnepf E (1980) **The fluorochrome, calcofluor white, binds oriented to structural polysaccharide fibrils** *Protoplasma* **105**:129–133
34. Alonso-Simón A. , et al. (2011) **The use of FTIR spectroscopy to monitor modifications in plant cell wall architecture caused by cellulose biosynthesis inhibitors** *Plant Signal Behav* **6**:1104–1110
35. Miedes E. , Vanholme R. , Boerjan W. , Molina A (2014) **The role of the secondary cell wall in plant resistance to pathogens** *Front Plant Sci* **5**
36. McCartney L. , et al. (2006) **Differential recognition of plant cell walls by microbial xylan-specific carbohydrate-binding modules** *Proc Natl Acad Sci USA* **103**:4765–4770
37. Grantham N.J. , et al. (2017) **An even pattern of xylan substitution is critical for interaction with cellulose in plant cell walls** *Nat Plants* **11**:859–865
38. Wierzbicki M.P. , Maloney V. , Mizrahi E. , Myburg A.A (2019) **Xylan in the middle: understanding xylan biosynthesis and its metabolic dependencies toward improving wood fiber for industrial processing** *Front Plant Sci* **25**
39. Ropartz D. , et al. (2011) **Performance evaluation on a wide set of matrix-assisted laser desorption ionization matrices for the detection of oligosaccharides in a high-throughput mass spectrometric screening of carbohydrate depolymerizing enzymes** *Rapid Commun Mass Spectrom* **25**:2059–2070
40. Mortimer J. C. , et al. (2010) **Absence of branches from xylan in Arabidopsis gux mutants reveals potential for simplification of lignocellulosic biomass** *Proc Natl Acad Sci USA* **107**:17409–17414
41. Urbanowicz B.R. , et al. (2012) **4-O-methylation of glucuronic acid in Arabidopsis glucuronoxylan is catalyzed by a domain of unknown function family 579 protein** *Proc Natl Acad Sci USA* **109**:14253–14258

42. Yuan Y. , Teng Q. , Lee C. , Zhong R. , Ye Z.-H (2014) **Modification of the degree of 4-O-methylation of secondary cell wall glucuronoxylan** *Plant Sci* :219–220
43. Wu A.-M. , et al. (2009) **The Arabidopsis IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation** *Plant J* **57**:718–731
44. Rohde A. , et al. (2004) **Molecular phenotyping of the pal1 and pal2 mutants of Arabidopsis thaliana reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism** *Plant cell* **16**:2749–2771
45. Busse-Wicher M. , et al. (2016) **Evolution of xylan substitution patterns in gymnosperms and angiosperms: implications for xylan interaction with cellulose** *Plant Physiol* **171**:2418–2431
46. Kang X. , et al. (2019) **Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR** *Nat Commun* **10**
47. Nishimura H. , Kamiya A. , Nagata T. , Katahira M. , Watanabe T (2018) **Direct evidence for α ether linkage between lignin and carbohydrates in wood cell walls** *Sci Rep* **8**
48. Terret O.M. , Dupree P (2019) **Covalent interactions between lignin and hemicelluloses in plant secondary cell walls** *Curr Opin Biotechnol* **56**:97–104
49. Hurkman W.J. , Tanaka C.K (1986) **Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis** *Plant Physiol* **81**:802–806
50. Merret R. , et al. (2017) **Heat shock protein HSP101 affects the release of ribosomal protein mRNAs for recovery after heat shock** *Plant Physiol* **174**:1216–1225
51. Livak K.J. , Schmittgen T.D (2001) **Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method** *Methods* **25**:402–408
52. Langmead B. , Salzberg S.L (2012) **Fast gapped-read alignment with Bowtie 2** *Nat. Methods* **9**:357–359
53. Kim D. , et al. (2013) **TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions** *Genome Biol* **14**
54. Ployet R. , et al. (2018) **Long cold exposure induces transcriptional and biochemical remodelling of xylem secondary cell wall in Eucalyptus** *Tree Physiol* **38**:409–422
55. Ployet R. , et al. (2019) **A systems biology view of wood formation in Eucalyptus grandis trees submitted to different potassium and water regimes** *New Phytol* **223**:766–782
56. Dai Y. , et al. (2020) **Implementing the CRISPR/Cas9 technology in Eucalyptus hairy roots using wood-related genes** *Int J Mol Sci* **21**

57. Van Acker R. , Vanholme R. , Piens K (2016) **Saccharification protocol for small-scale lignocellulosic biomass samples to test processing of cellulose into glucose** *Bio-Protocol* **6**
58. Casabona M.G. , Vandenbrouck Y. , Atree I. , Couté Y (2013) **Proteomic characterization of *Pseudomonas aeruginosa* PAO1 inner membrane** *Proteomics* **13**:2419–2423
59. Bouyssie D. , et al. (2020) **Proline: an efficient and user-friendly software suite for large-scale proteomics** *Bioinformatics* **36**:3148–3155
60. Couté Y. , Bruley C. , Burger T (2020) **Beyond target-decoy competition: stable validation of peptide and protein identifications in mass spectrometry-based discovery proteomics** *Anal Chem* **92**:14898–14906
61. Perez-Riverol Y. , et al. (2019) **The PRIDE database and related tools and resources in 2019: improving support for quantification data** *Nucleic Acids Res* **47**:D442–D450
62. Wieczorek S. , et al. (2017) **DAPAR & ProStaR: software to perform statistical analyses in quantitative discovery proteomics** *Bioinformatics* **33**:135–136
63. Lapierre C. , Pollet B. , Rolando C (1995) **New insights into the molecular architecture of hardwood lignins by chemical degradative methods** *Res Chem Intermed* **21**:397–412
64. Hoebler C , Barry J.-L. , David A. , Delort-Laval J (1989) **Rapid acid hydrolysis of plant cell wall polysaccharides and simplified quantitative determination of their neutral monosaccharides by gas-liquid chromatography** *J Agric Food Chem* **37**:360–365
65. Blakeney A.B. , Harris P.J. , Henry R.J. , Stone B.A (1983) **A simple and rapid preparation of alditol acetates for monosaccharide analysis** *Carbohydr Res* **113**:291–299
66. Blumenkrantz N. , Asboe-Hansen G (1973) **New method for quantitative determination of uronic acids** *Analytical Biochemistry* **54**:484–489
67. Thibault J.F (1979) **Automated-method for the determination of pectic substances** *Lebensmittel-Wissenschaft Und-Technologie* **12**:247–251

Author information

Alicia Kairouani

Laboratoire Génome et Développement des Plantes, Université de Perpignan via Domitia, CNRS, UMR5096, Perpignan Cedex, France

Dominique Pontier

Laboratoire Génome et Développement des Plantes, Université de Perpignan via Domitia, CNRS, UMR5096, Perpignan Cedex, France

Claire Picart

Laboratoire Génome et Développement des Plantes, Université de Perpignan via Domitia, CNRS, UMR5096, Perpignan Cedex, France

Fabien Mounet

Laboratoire de Recherche en Sciences Végétales, Université de Toulouse III, CNRS UPS UMR5546, Castanet-Tolosan, France

Yves Martinez

FRAIB-CNRS plateforme imagerie, Castanet-Tolosan, France

Lucie Le-Bot

Biopolymères Interactions Assemblages, UR1268 BIA, INRAE, Nantes, France

Mathieu Fanuel

Biopolymères Interactions Assemblages, UR1268 BIA, INRAE, Nantes, France, PROBE research infrastructure, BIBS Facility, INRAE, Nantes, France

Philippe Hammann

Plateforme Protéomique Strasbourg Esplanade de CNRS, Université de Strasbourg, Strasbourg, France

Lucid Belmudes

Université Grenoble Alpes, INSERM, UA13 BGE, CNRS, CEA, FR2048, Grenoble, France

Rémy Merret

Laboratoire Génome et Développement des Plantes, Université de Perpignan via Domitia, CNRS, UMR5096, Perpignan Cedex, France
ORCID iD: [0000-0002-3790-1115](https://orcid.org/0000-0002-3790-1115)

Jacinthe Azevedo

Laboratoire Génome et Développement des Plantes, Université de Perpignan via Domitia, CNRS, UMR5096, Perpignan Cedex, France

Marie-Christine Carpentier

Laboratoire Génome et Développement des Plantes, Université de Perpignan via Domitia, CNRS, UMR5096, Perpignan Cedex, France

Dominique Gagliardi

Institut de Biologie Moléculaire des Plantes, BMP, CNRS, Université de Strasbourg, Strasbourg, France

Yohann Couté

Université Grenoble Alpes, INSERM, UA13 BGE, CNRS, CEA, FR2048, Grenoble, France
ORCID iD: [0000-0003-3896-6196](https://orcid.org/0000-0003-3896-6196)

Richard Sibout

Biopolymères Interactions Assemblages, UR1268 BIA, INRAE, Nantes, France

Natacha Bies-Etheve

Laboratoire Génome et Développement des Plantes, Université de Perpignan via Domitia, CNRS, UMR5096, Perpignan Cedex, France

For correspondence: natacha.etheve@univ-perp.fr

Thierry Lagrange

Laboratoire Génome et Développement des Plantes, Université de Perpignan via Domitia, CNRS, UMR5096, Perpignan Cedex, France

For correspondence: lagrange@univ-perp.fr

ORCID iD: [0000-0003-3090-0918](https://orcid.org/0000-0003-3090-0918)

Editors

Reviewing Editor

Heather McFarlane

University of Toronto, Canada

Senior Editor

Jürgen Kleine-Vehn

University of Freiburg, Germany

Reviewer #1 (Public Review):

Secondary cell walls support vascular plants and conduct water throughout the plant body, but are also important resources for lignocellulosic feedstocks. Secondary cell wall synthesis is under complex transcriptional control, presumably because it must only be initiated after cell growth is complete. Here, the authors found that two Musashi-type RNA-binding proteins, MSIL2 and MSIL4 are redundantly required for secondary cell wall development in Arabidopsis. The plant phenotypes could be complemented by the wild-type version of either protein, but not by a MSIL4 version that carries mutations in the conserved RNA-binding domains, and the authors localized MSIL2 & 4 to stress granules, implicating the RNA-binding function of MSIL4 in the cell wall phenotype. Upon closer inspection, the secondary cell wall phenotypes included changes in vasculature morphology, and minor changes to lignin and hemicellulose (glucuronoxylan). While there were no changes to likely cell wall target genes in the transcriptome of *msil2msil4* plants, proteomics experiments found glucuronoxylan biosynthesis components were upregulated in the mutants, and they detected an increase in substituted xylan via several methods. Finally, they documented MSIL4 binding to RNA encoding one of these targets, suggesting that MSIL2 and MSIL4 act to post-transcriptionally regulate glucuronoxylan modification. Altogether, this is a new mechanism by which cell wall composition could be regulated.

Overall, the manuscript is well-written, the data are generally high-quality, and the authors typically use several independent methods to support each claim. However, several important questions remain unanswered by this work in its current state and the model presented in Figure 7 is quite speculative. For example, the link between the striking plant phenotype and GXM misregulation is unclear since GXM overexpression doesn't alter plant phenotypes or lignin content (Yuan et al 2014 Plant Science), so misregulation of GXMs in *msil2msil4* mutants clearly is not the whole story. It also remains to be determined why one particular secondary cell wall synthesis enzyme is regulated likely post-transcriptionally, while so much of the pathway is regulated at the transcriptional level. There are likely other targets for MSIL2- and MSIL4-mediated regulation since it seems that MSIL2 and MSIL4 are expressed in tissues that are not synthesizing secondary cell walls.

Reviewer #2 (Public Review):

This work explored the biological functions of a small family of RNA-binding proteins that was previously studied in animals, but was uncharacterized in plants. Combinatorial T-DNA insertional mutants disrupting the expression of the four Mushashi-like (MSIL) genes in *Arabidopsis* revealed that only the *msil2 msil4* double mutant visibly alters plant development. The *msil2/4* plants produced stems that could not stand upright. Transgene complementation, site-directed mutagenesis of MSIL4 conserved RNA-binding motifs, and *in vitro* RNA binding assays support the conclusion that the loss of MSIL2 and MSIL4 function is responsible for the observed morphological defects. MSIL2/4 interact with proteins associated with mRNA 3'UTR binding and translational regulation.

The authors present compelling biochemical evidence that Mushashi-like2 (MSIL2) and MSIL4 jointly regulate secondary cell wall biosynthesis in the *Arabidopsis* stem. Quantitative analyses of proteins and transcripts in *msil2/4* stems uncovered upregulation of several xylan-related enzymes (despite WT-like RNA levels). Consistent with MALDI-TOF data for released xylan oligosaccharides, the authors propose a model in which MSIL2/4 negatively regulate the translation of GXM (glucuronoxylan methyltransferase), a presumed rate-limiting step. The molecular links between overmethylated xyans and the observed stem defects (which include subtle reductions in lignin and increases beta-glucan polymer distribution) warrants further investigation in future studies. Similarly, as the authors point out, it is intriguing that the loss of the broadly expressed MSIL2/4 genes only significantly affects specific cell types in the stem.

Reviewer #3 (Public Review):

The manuscript by Kairouani et al. investigates the function of a small family of plant RNA binding proteins with similarity to the well-studied Musashi protein in animals, and, therefore, called MUSASHI-LIKE1-4 (MSL1-4). Studies on the biological importance of post-transcriptional control of gene expression via RNA-binding proteins in plants are not numerous, and advances in this important field are much needed. The thorough work presented in this manuscript is such an advance.

The central observations of the paper are:

- Knockout of any MSL gene alone does not produce a phenotype.
It is of note that basic characterization of knockout mutations is really well done - for example, the authors have taken care to raise specific antibodies to each of the MSL proteins and use them to demonstrate that each of the T-DNA insertion mutants used actually does knock out protein production from the corresponding gene.
- Knockout of MSL2/4 (but no other double mutant) produces a clear leaf phenotype, and a remarkable stem phenotype in which the mutants collapse as they are unable to support upright growth
- The phenotypes of knockout mutants persist in point mutants defective in RNA-binding, indicating that RNA-binding is required for biological activity. Consistent with this, and associate physically with other RNA-binding proteins and translation factors.
- MSL proteins are cytoplasmic
- The *msl2/4* mutants present multiple defects in secondary cell wall composition and structure, probably explaining their inability to grow upright. I did not examine the cell wall

analyses in detail as I am no specialist in this field.

- Msl2/4 mutants show transcriptomic changes with at large two big categories of differentially expressed genes compared to wild type.

(1) Genes related to cell wall metabolism

(2) Genes associated with defense against herbivores and pathogens

- Two of the mRNAs encoding cell wall factors with significant upregulation in msl2/4 mutants compared to wild type also associate physically with MSL4 as judged by RNA-immunoprecipitation-RT-PCR assays, and this physical association is abrogated in the RNA-binding deficient MSL4 mutant.

Altogether, the study shows clear biological relevance of the MSL family of RNA-binding proteins, and provides good arguments that the underlying mechanism is control of mRNAs encoding enzymes involved in secondary cell wall metabolism (although concluding on translational control in the abstract is perhaps saying too much - post-transcriptional control will do given the evidence presented). One observation reported in the study makes it vulnerable to alternative interpretation, however, and I think this should be explicitly treated in the discussion:

The fact that immune responses are switched on in msl2/4 mutants could also mean that MSL2/4 have biological functions unrelated to cell wall metabolism in wild type plants, and that cell wall defects arise solely as an indirect effect of immune activation (that is known to involve changes in expression of many cell wall-modifying enzymes and components such as pectin methylesterases, xyloglucan endotransglycosylases, arabinogalactan proteins etc. Indeed, the literature is rich in examples of gene functions that have been misinterpreted on the basis of knockout studies because constitutive defense activation mediated by immune receptors was not taken into account (see for example Lolle et al., 2017, *Cell Host & Microbe* 21, 518-529).

With the evidence presented here, I am actually close to being convinced that the primary defect of msl2/msl4 mutants is directly related to altered cell wall metabolism, and that defense responses arise as a consequence of that, not the other way round. But I do not think that the reverse scenario can be formally excluded with the evidence at hand, and a discussion listing arguments in favor of the direct effect proposed here would be appropriate. Elements that the authors could consider to include would be the isolation of a cellulose synthase mutant as a constitutive expressor of jasmonic acid responses (cev1) as a clear example that a primary defect in cell wall metabolism can produce defense activation as secondary effect. The interaction of MSL4 with GXM1/3 mRNAs is also helpful to argue for a direct effect, and it would strengthen the argument if more examples of this kind could be included.