An updated sulfate transporter phylogeny uncovers a perennial-specific subgroup associated with lignification

Samantha M. Surber[1,2], Chen Hsieh[3], Scott A. Harding[4], Chung-Jui Tsai[1,2,3,4,5]

[1] Department of Plant Biology, University of Georgia, Athens, GA 30602

[2] Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN 37831

[3] Institute of Bioinformatics, University of Georgia, Athens, GA 30602

[4] Warnell School of Forestry, University of Georgia, Athens, GA 30602

[5] Department of Genetics, University of Georgia, Athens, GA 30602

Running title: SULTR phylogeny and lignin-specific PtaSULTR3.2

Keywords: SULTR, Populus, sulfur assimilation, lignin biosynthesis, xylem

Corresponding author: Chung-Jui Tsai (cjtsai@uga.edu), 120 E. Green Street, Athens, GA 30602, USA

ABSTRACT

Sulfate-proton co-transporters (SULTRs) mediate sulfate uptake, transport, storage, and assimilation within plants. The SULTR family has historically been categorized into four groups, SULTR1 to SULTR4. While the roles of SULTR groups 1, 2, and 4 are well characterized, there is much less consensus regarding the functions of the large SULTR3 group. Here, we present an updated SULTR phylogenetic analysis of angiosperms, including multiple early divergent lineages. Our results suggest that the enigmatic SULTR3 group consists of four distinct subfamilies which predate the emergence of angiosperms. This finding justifies reclassification of the SULTR family from four to seven subfamilies. Key evidence supporting this expansion includes subfamily-specific variations in gene structure and amino acid substitutions in the substrate-binding pocket. Gene expression analysis of *Populus tremula* × *alba* revealed divergent tissue expression preferences among *SULTR* subfamilies and between genome duplicates. We observed partitioned expression in vascular tissues among the four *SULTR3* subfamilies, with *PtaSULTR3.4a* and *PtaSULTR3.2a* preferentially expressed in primary and secondary xylem, respectively. Gene coexpression analysis revealed coordinated expression of *PtaSULTR3.4a* with genes associated with phosphate starvation responses and various nutrient transport functions. In contrast, *PtaSULTR3.2a* showed strong coexpression with lignification genes and their upstream transcription regulators. The transcript abundance of *PtaSULTR3.2a* was down-regulated in response to genetic perturbation or tension wood formation which resulted in reduced lignin accrual. *PtaSULTR3.2a* belongs to a eudicot-specific branch of the SULTR3.1 subfamily found only in perennial species, supporting dedicated roles in lignifying tissues. Together, our findings provide a refined phylogenetic inference of the SULTR family and suggest potential neofunctionalization of the expanded SULTR3 subfamilies during the evolution of vascular and perennial plants.

INTRODUCTION

Sulfur is an essential plant macronutrient taken up from the soil as inorganic sulfate (Kopriva et al. 2019). Uptake and plantwide distribution are mediated by a family of sulfate-proton co-transporters (SULTRs) historically classified into four groups, referred to as SULTR1 to SULTR4 (Takahashi et al. 2000). This Arabidopsis (*Arabidopsis thaliana*)-based nomenclature has been followed in many subsequent phylogenetic analyses (Takahashi et al. 2012; Gallardo et al. 2014; Ding et al. 2016; Xun et al. 2021; Zhao et al. 2022; Chen et al. 2024). Two recent studies classified the SULTR family into five groups but with inconsistent nomenclature and memberships (Heidari et al. 2023; Puresmaeli et al. 2023). Meanwhile, molecular characterization continues to shed light on all SULTR groups as evidenced by their varied subcellular localization, transport kinetics, tissue expression, and stress responsiveness (Takahashi 2019).

Among the different groups, SULTR4 transporters are tonoplast-localized and facilitate sulfate efflux from the vacuole (Kataoka et al. 2004b). SULTR1 and SULTR2 members are plasma membrane-localized with SULTR1 being high-affinity transporters involved in sulfate uptake by the roots (Takahashi et al. 2000; Yoshimoto et al. 2002), and SULTR2 being low-affinity transporters expressed primarily in the root cylinder for root-to-shoot sulfate transport (Takahashi et al. 1997; Takahashi et al. 2000). SULTR3 is the largest and most diverse group, with five genes in Arabidopsis—*AtSULTR3.1*, *AtSULTR3.2*, *AtSULTR3.3*, *AtSULTR3.4,* and *AtSULTR3.5* (Zuber et al. 2010; Takahashi et al. 2012). The subcellular distribution and function of SULTR3 members are less precisely known but recent reports suggest localization to the plasma membrane (Kataoka et al. 2004a; Yamaji et al. 2017; Ding et al. 2020), endoplasmic reticulum (Zhao et al. 2016), and/or chloroplast (Cao et al. 2013; Chen et al. 2019).

SULTR3 members are enigmatic with respect to their sulfate transport capabilities. AtSULTR3.1, AtSULTR3.2, and AtSULTR3.3 did not show sulfate uptake activity in yeast (Takahashi et al. 2000). To date, successful restoration of yeast mutant growth on sulfate-containing media by SULTR3 has only been reported for LjSST1 (SULTR3.5) of lotus (*Lotus japonicus*) and MhSULTR3.1a of apple (*Malus hupehensis*), although the uptake kinetics were not determined (Krusell et al. 2005; Xun et al. 2021). Cao et al. (2013) used individual *sultr3* mutants to show that all Arabidopsis SULTR3 members, except AtSULTR3.5, contribute substantially to sulfate uptake by isolated chloroplasts. Although AtSULTR3.5 alone does not support sulfate uptake by isolated Arabidopsis chloroplasts or in yeast, its co-expression with *AtSULTR2.1* improved growth of yeast mutants over those expressing *AtSULTR2.1* alone (Kataoka et al. 2004a). This

95 cooperation of AtSULTR3.5 and AtSULTR2.1 in the vasculature contributes to root-to-shoot sulfate
96 transport in Arabidopsis (Kataoka et al. 2004a). In rice (*Oryza sativa*), two independent *low phytic acid*
97 (*lpa*) mutant alleles were both mapped to *OsSULTR3.3*, even though no *in vitro* transport activity of
98 OsSULTR3.3 for sulfate, phosphate, inositol, or inositol triphosphate could be demonstrated in yeasts or
99 *Xenopus* oocytes (Zhao et al. 2016). Similarly, SULTR3.4 was independently characterized as a SULTR-like
100 phosphorus distribution transporter (SPDT) involved in phosphate allocation from xylem to developing
101 tissues in rice, barley (*Hordeum vulgare*), and Arabidopsis (Yamaji et al. 2017; Ding et al. 2020; Gu et al.
102 2022). OsSULTR3.4, HvSULTR3.4, and AtSULTR3.4 all function as an influx transporter for phosphate, but
103 not sulfate, in oocytes, and their knockout mutants showed impaired phosphorus distribution to
104 developing organs (Yamaji et al. 2017; Ding et al. 2020; Gu et al. 2022). These studies highlight some of
105 the ongoing challenges in elucidating the evolution and function of the diverse SULTR3 group.
106
107 *SULTR* gene expression is sensitive to abiotic stressors, including drought and salt, in Arabidopsis (Cao et
108 al. 2014), barrelclover (*Medicago truncatula*) (Gallardo et al. 2014), poplar (*Populus trichocarpa*)
109 (Malcheska et al. 2017), maize (*Zea mays*) (Huang et al. 2018), wheat (*Triticum turgidum* L. ssp.
110 *durum*)(Puresmaeli et al. 2023), and soybean (*Glycine max*) (Zhou et al. 2024). The responsiveness of
111 *SULTR* genes to drought and other abiotic stresses varies between studies in part because expression is
112 modulated not only by plant sulfur status (Vidmar et al. 1999; Takahashi et al. 2000; Rouached et al.
113 2008; Maruyama-Nakashita et al. 2015), but also by stress hormones like ABA (abscisic acid) and MeJA
114 (methyl jasmonic acid) (Cao et al. 2014; Xun et al. 2021). Notably, four of five *AtSULTR3* genes were
115 identified among the 124 transporter transcripts enriched in guard cells (Bauer et al. 2013). Moreover,
116 quintuple *sultr3* mutants showed impaired sulfate uptake for stress-induced ABA synthesis in Arabidopsis
117 (Chen et al. 2019). This supports sulfate as a xylem-borne signal that precedes localized ABA biosynthesis
118 and stomatal closure in response to drought (Ernst et al. 2010; Malcheska et al. 2017; Batool et al. 2018).
119
120 Studies on gene expression and the seasonal dynamics of wood and xylem sap sulfate abundance in
121 *Populus tremula* × *alba* have provided additional insights into the function of SULTR3 in stems. The
122 transcripts of *PtaSULTR3.1b*, *PtaSULTR3.2a*, and *PtaSULTR3.3a* peaked during active growth in the
123 summer (Dürr et al. 2010; Ko et al. 2011; Malcheska et al. 2013). The hydraulic transport network for
124 long-distance and lateral movements of solutes and water are more complex in large stature trees
125 (Pfautsch et al. 2015). In addition, the formation of large woody sinks depends heavily on sulfur
126 assimilation to supply one-carbon (C1) units for methyl-intensive lignification (Rajinikanth et al. 2007).

Here, we report an updated SULTR phylogeny using 20 angiosperm species, including several woody perennials where SULTR function has been little explored. The inclusion of basal lineages revealed a much more ancient origin of the expanded SULTR3 group than previously recognized. Multiple lines of evidence, including gene structure, substrate binding pocket sequence substitutions, and gene expression, support the functional divergence of SULTR3. We also discuss the potential association of some SULTR3 members with lignification in *Populus*.

METHODS

Sequence collection and annotation

SULTR protein sequences from 22 species were obtained from Phytozome v13 (https://phytozome-next.jgi.doe.gov/)(Goodstein et al. 2012), and the list, including gene models, genome version, and species, is provided in Supplemental Table S1. Sequences were aligned using the multiple sequence alignment tool, T-Coffee (Notredame et al. 2000) for inspection. Sequences with large gaps, misalignment, and/or gene structure misannotation were manually curated, sometimes in consultation with independently annotated genome versions at NCBI. Exon-intron annotation was retrieved from Phytozome for spikemoss (*Selaginella moellendorffii*), moss (*Physcomitrium patens*), pineapple (*Ananas comosus*), blue columbine (*Aquilegia coerulea*), Arabidopsis (*Arabidopsis thaliana*), poplar (*Populus trichocarpa*), rice (*Oryza sativa*), and sorghum (*Sorghum bicolor*) and the structure was illustrated using Gene Structure Draw Tool (https://www.bioinformatics.uni-muenster.de/tools/strdraw/).

Phylogenetic reconstruction

Preliminary alignment of all 262 protein sequences performed with T-Coffee on EMBL-EBI (Madeira et al. 2024) was used in a model test using ModelFinder (Kalyaanamoorthy et al. 2017) implemented in IQ-TREE v1.6.12 (Nguyen et al. 2015). This revealed a general JTT model for amino acid substitution rate (Jones et al. 1992), gamma rate of 4 (G4), allowing for a proportion of invariable sites (I), and empirical frequencies (F) [JTT+G4+I+F]. The model JTT+G4+I+F was used for maximum likelihood inference for phylogenetic tree construction with IQ-TREE v1.6.12 and 1000 ultrafast bootstrap cycles. Trees were visualized on TreeViewer (https://treeviewer.org/) (Bianchini and Sánchez-Baracaldo 2024) and edited in

MEGA-X (Kumar et al. 2018) for presentation. The full tree (Figure 1) is unrooted, while the SULTR3-specific tree (Figure 2) is rooted on the basal spikemoss SULTR3 sequences.

Substrate binding site alignment and annotation

All poplar and Arabidopsis SULTR protein sequences were aligned as above. The structure-resolved AtSULTR4.1 (AT5G13550) protein sequence was used as the reference to annotate key residues involved in substrate and co-substrate binding and interactions according to Wang et al. (2021).

Gene expression analysis

RNA-Seq datasets from published *Populus tremula* × *alba* INRA 717-1B4 studies (Swamy et al. 2015; Xue et al. 2016; Tsai et al. 2020; Zhou et al. 2023) were obtained from the NCBI Short Read Archive (SRA) using the `nfcore/fetchngs` pipeline (v1.11.0) (Ewels et al. 2020). Quality control and read trimming were performed using Cutadapt v4.4 (Martin 2011) with parameters -q 30, -m 30, and --trim-n. Reads were mapped to the *P. tremula* subgenome (*Populus tremula* × *alba* HAP1 v5.1) of the hybrid (Zhou et al. 2023) available on Phytozome v13. For expression quantification, kallisto v0.48.0 (Bray et al. 2016) was used, with a *k*-mer size of 21 for indexing. Single-end reads were quantified with an estimated average fragment length of 200 and a standard deviation of 20. Paired-end reads were quantified without the stranded option. *SULTR* gene expression values were extracted and statistical significance of treatment responses was performed using PRISM v10.1.0 (GraphPad Software, Boston, MA). Heatmap and bar graphs were generated using Excel.

Gene coexpression analysis was performed using 179 RNA-seq datasets from *Populus tremula* × *alba* stem vascular tissues. The full list of datasets is provided in Supplemental Table 2. Data were similarly processed as above to obtain expression values. Genes with low expression (TPM < 5 in all samples) were filtered out and only those with a coefficient of variance (CV) > 50% were retained. Pairwise Pearson correlation coefficients (PCC) were calculated for all remaining genes across the 179 samples. Genes with a PCC $\geq$ 0.7 in relation to *PtaSULTR3.2a* or *PtaSULTR3.4a* were extracted, along with the correlation matrices, for network visualization in Cytoscape v3.10.3 using the Prefuse Force-Directed OpenCL layout (Shannon et al. 2003). Gene Ontology (GO) enrichment was performed using ShinyGO v0.77 (Ge et al. 2020) implemented at http://aspendb.uga.edu/ShinyGO/.

RESULTS

Differential expansion of SULTR subfamilies

A total of 262 SULTR protein sequences from 20 angiosperms and two early diverging land plant lineages were retrieved from Phytozome v13 (Goodstein et al. 2012) (https://phytozome-next.jgi.doe.gov/) (Table 1). These SULTR proteins are typically annotated based on their Arabidopsis orthologs into four groups, SULTR1 to SULTR4 (Takahashi et al. 2000; Takahashi et al. 2012). The number of sequences per group varies due to differential gene expansion among the species studied. The SULTR4 subfamily consists of one or two sequences in each sampled taxon, while the other three groups show varied expansions. SULTR3 has the highest number of sequences, followed by SULTR1, whereas SULTR2 exhibits a modest expansion, observed only in some eudicot species (Table 1).

We constructed a SULTR phylogenetic tree using maximum likelihood inference with IQ-TREE (Nguyen et al. 2015) (Figure 1). The SULTR members formed distinct clusters in a topology consistent with previous reports (Takahashi et al. 2012; Ding et al. 2016; Zhao et al. 2022; Chen et al. 2024). This topology suggests the existence of two ancient clades, based on the placement of SULTRs from the bryophyte moss (*Physcomitrella patens*), as also suggested by Takahashi et al. (2012). The smaller Ancient Clade I (AC-I) comprises the tonoplast SULTR4, whereas the larger Ancient Clade II (AC-II) encompasses the other three SULTR groups (Figure 1). Within AC-II, the ancient vascular plant spikemoss (*Selaginella moellendorffii*) contains two distinct groups that are basal to two subclades: one represented by SULTR1-SULTR2 (referred to as AC-IIA) and the other by SULTR3 (AC-IIB). This suggests that the divergence of the two subclades coincides with the evolution of vascular plants (Figure 1).

SULTR1 and SULTR2 originate from ancient local duplication

AC-IIA consists of two strongly supported nodes represented by the SULTR1 and SULTR2 subfamilies. Both subfamilies are evolutionarily conserved in angiosperms, including the basal lineages Amborella (*Amborella trichopoda*) and tulip tree (*Liriodendron tulipifera*). Interestingly, in both genomes, *SULTR1* and *SULTR2* are found in close proximity with a divergent (head-to-head) orientation. This head-to-head pattern has been maintained in nearly all sampled angiosperm taxa, despite lineage-specific variations, including additional whole-genome or tandem duplications (see below, Supplemental Table 1). This suggests that *SULTR1* and *SULTR2* originated from a local duplication event before the advent of

220 angiosperms. The data further indicates that the proximally located and oppositely oriented *SULTR1* and
221 *SULTR2* copies are likely the founding members of each subfamily.
222
223 Monocot SULTR1s formed two clusters, each represented by all sampled taxa, including the basal lineage
224 *Ananas comosus* (pineapple) (Figure 1). This suggests their divergence early during monocot evolution.
225 Eudicot SULTR1s also formed two distinct clusters. The smaller cluster has limited species representation,
226 and there is only a single SULTR1 in the basal eudicot blue columbine (*Aquilegia coerulea*)(Figure 1). This
227 suggests that the expansion of SULTR1 in eudicots occurred later and only in certain taxa. In the smaller
228 eudicot SULTR1 cluster, four of the eight species—eucalyptus (*Eucalyptus grandis*), lettuce (*Lactuca*
229 *sativa*), monkey flower (*Mimulus guttatus*), and oak (*Quercus rubra*)—harbor tandem duplicates
230 (Supplemental Table 1). This suggests a tendency for lineage-specific local amplification of this SULTR1
231 subgroup.
232
233 The SULTR2 subfamily is smaller than SULTR1, with 1-2 copies of the two basal angiosperm lineages and
234 monocots (Figure 1, Table 1). However, the SULTR2 subfamily size is more variable among eudicots,
235 ranging from a complete absence in blue columbine to four to eight copies in the legumes (Table 1). In
236 the latter case, we detected whole genome, tandem, and other duplication events both before and after
237 the split of barrelclover and soybean, leading to a disproportionately expanded SULTR2 subfamily in
238 these lineages (Supplemental Table 1).
239
240 **SULTR3 encompasses distinct subfamilies predating angiosperm evolution**
241 AC-IIB encompasses the entire SULTR3 group, with spikemoss members at the basal position (Figure 1).
242 This supports the vascular-specific origin of SULTR3s as previously suggested (Takahashi et al. 2012). Like
243 AC-IIA, AC-IIB is divided into two well-supported nodes, each containing two distinct branches. The first
244 node, hereafter AC-IIB-1, includes the SULTR3.1/3.2 and SULTR3.5 branches, while the second node, AC-
245 IIB-2, comprises the SULTR3.3 and SULTR3.4 branches (Figure 1). Each of these four branches contains
246 one or two members from Amborella and tulip tree, which are sister to all flowering plants (Table 1). This
247 indicates that, like SULTR1 and SULTR2, the four SULTR3 branches evolved before the emergence of
248 angiosperms.
249
250 To better understand the evolutionary history of AC-IIB, we constructed a phylogenetic tree using only
251 SULTR3 sequences (Figure 2). The overall topology mirrors that of the AC-IIB group in Figure 1, with

252 angiosperm members clustering into four branches predating the emergence of basal angiosperms.
253 Three of the four branches correspond to the SULTR3.3, SULTR3.4, and SULTR3.5 subfamilies, each with
254 distinct monocot and dicot groups consistent with their early divergence during angiosperm evolution
255 (Figure 2). Most species are represented by a single copy, while others harbor two to four paralogs (Table
256 1).
257
258 The SULTR3.1 branch (also called SULTR3.1/3.2 branch in previous papers, see below) has substantially
259 expanded in both monocots and eudicots after their split. Within the monocots, we identified two
260 strongly supported subgroups for grass SULTR3.1s, rooted with the lone member of pineapple (Figure 2),
261 indicating their origin from a grass-specific duplication event (Figure 2). We also observed two distinct
262 subgroups for eudicots, including blue columbine, indicating their divergence predates the common
263 ancestor of eudicots (Figure 2). The larger subgroup includes all 12 sampled eudicot taxa, many of which
264 have additional paralogs due to lineage-specific duplications. Examples include Brassicaceae (Arabidopsis
265 AtSULTR3.1 and AtSULTR3.2), Salicoid (poplar PtSULTR3.1a and PtSULTR3.1b), and legume-specific
266 duplications (Figure 2). In contrast, the smaller eudicot subgroup is only retained in blue columbine and
267 six other sampled taxa in both rosid and asterids, all of which are perennial species (Figure 2). Distinct
268 orthologs were also identified in other perennial eudicot genomes available on Phytozome v13, such as
269 birch (*Betula platyphylla*), papaya (*Carica papaya*), chestnut (*Castanea dentata*), orange (*Citrus sinensis*),
270 cassava (*Manihot esculenta*), peach (*Prunus persica*), and castor bean (*Ricinus communis*) of rosid, and
271 coffee (*Coffea arabica*) and hydrangea (*Hydrangea quercifolia*) of asterids. These findings suggest that
272 this SULTR3.1 subgroup may play an important role in the functionality of perennial species.
273
274 The duplication history of the SULTR3.1 branch has led to inconsistent gene and subfamily
275 nomenclatures in the literature. Specifically, Arabidopsis AtSULTR3.1 and AtSULTR3.2 are paralogs within
276 the same SULTR3.1 (large) subgroup, unlike AtSULTR3.3, AtSULTR3.4, and AtSULTR3.5 which represent
277 phylogenetically distinct subfamilies. In poplar, PtSULTR3.1 and PtSULTR3.2 are descendants of the
278 ancestral eudicot duplication in the larger and smaller SULTR3.1 subgroups, respectively, with their
279 Salicoid paralogs designated with “a” and “b” suffixes (*e.g.*, PtSULTR3.1a and PtSULTR3.1b) in the
280 literature (Dürr et al. 2010; Malcheska et al. 2017). In these examples, AtSULTR3.1/3.2 are orthologous to
281 PtSULTR3.1a/b, while the perennial-specific PtSULTR3.2a/b do not have direct orthologs in Arabidopsis
282 (Figure 2). We preserve the use of SULTR3.2 as a gene name in keeping with convention and refer to this
283 subfamily as SULTR3.1 for now. Phylogenomics-informed gene (re)naming is clearly needed but is beyond

284 the scope of this study (see HUGO commentary in this issue). Our analysis strongly supports four distinct
285 SULTR3 (AC-IIB) subfamilies SULTR3.1, SULTR3.5 (node AC-IIB-1), SULTR3.3, and SULTR3.4 (node AC-IIB-2),
286 alongside the two, well studied AC-IIA subfamilies SULTR1 and SULTR2. Together with the more divergent
287 AC-I SULTR4 group, angiosperm taxa contain seven evolutionarily conserved SULTR subfamilies (Figure 1),
288 rather than the historically proposed four (Takahashi et al. 2000).
289
290 **Exon-intron structure varies with the evolution of *SULTR subfamilies***
291 Analysis of gene structure revealed variations that generally align with the SULTR subfamily evolution
292 inferred from our phylogenetic trees. *SULTR4* genes of the small AC-I clade have 17 exons with
293 evolutionarily conserved junctions in all species we examined (Figure 3). However, multiple exon-fusion
294 events have occurred in the expanded AC-II subfamilies. Specifically, three fusions involving exons 8 and
295 9, exons 10 to 12, and exons 13 and 14 are shared among all AC-II subfamilies and across species,
296 including the basal spikemoss members (Figure 3). Among them, *Sm184750* basal to AC-IIB is likely the
297 founding member with just these three exon fusions. This gene structure is retained in *SULTR1* (AC-IIA),
298 *SULTR3.3* and *SULTR3.4* (AC-IIB-2) members.
299
300 *SULTR2* members (AC-IIA) have an additional fusion between exons 6 and 7. This fusion is absent in the
301 basal Amborella *SULTR2* which retains the same gene structure as the presumed AC-II founding member
302 *Sm184750* and the sister *SULTR1* branch (Figure 3). Interestingly, this fusion is present in the more
303 distantly related spikemoss *Sm169988* basal to AC-IIB (Figure 3). Similarly, the *SULTR3.1* and *SULTR3.5*
304 subfamilies share a conserved fusion of exons 1 and 2 across basal, monocot, and dicot lineages of the
305 AC-IIB-1 subclade, as well as the less related *Sm113596* basal to AC-IIA (Figure 3). The independent
306 occurrence of the same exon fusion event in unrelated lineages may suggest convergent evolution.
307 Alternatively, the minor inconsistencies noted above may reflect limitations in current phylogenetic
308 reconstructions in correctly positioning distant orthologs from early-divergent lineages, such as moss
309 and spikemoss. This limitation is evident for the SULTR4 clade (Figure 1) as previously reported
310 (Takahashi et al. 2012).
311
312 Finally, gene structure analysis also revealed multiple unusual or lineage-specific fusion events. Examples
313 include grass-specific fusions, including secondary fusions between above-mentioned exon fusion
314 events, in *SULTR3.1* and *SULTR3.4* members (Figure 3, Phytozome v13). Furthermore, half of the
315 Arabidopsis *SULTR* family members have unique exon fusion events conserved in Brassicaceae, indicative

316 of their unusual evolution. This suggests that using Arabidopsis genes as sole references for comparative
317 studies may introduce confounding factors for evolutionary inference.
318
319 **Protein modeling revealed binding pocket substitutions in the expanded SULTR3 subfamilies**
320 The recently resolved protein structure of Arabidopsis AtSULTR4.1 (Wang et al. 2021) was used as a
321 template to identify key substrate binding residues in other Arabidopsis and poplar SULTR proteins. The
322 key residue Glu347 (based on AtSULTR4.1 numbering), essential for co-substrate proton sensing and
323 transport, is conserved in all SULTRs (Figure 4). Key residues lining the sulfate binding pocket, including
324 Gln112, Tyr116, Ser392, and Arg(Lys)393, are also conserved across SULTR proteins (Figure 4). Mutations
325 to Ala at any of these positions resulted in a near-total loss of sulfate transport activity (Wang et al.
326 2021). However, two other residues in the predicted sulfate binding pocket showed subfamily-specific
327 substitutions. Ala153 is replaced by Ser in the SULTR3.3 and SULTR3.4 subfamilies (node AC-IIB-2) (Figure
328 4). Depending on whether the hydroxyl group of Ser interferes with productive hydrogen bonding in the
329 pocket, these substitutions may be catalytically important. Furthermore, Ser390 is replaced by Pro in the
330 SULTR3.1 and SULTR3.5 subfamilies (node AC-IIB-1) or by Ala in SULTR3.3 (Figure 4). Ser390 is thought to
331 interact with Arg393, which likely provides positive electrostatic potential to stabilize substrate binding
332 (Wang et al. 2021). Whether these SULTR3-specific substitutions contribute to the reported difficulty of
333 demonstrating its *in vitro* transport activity (Takahashi et al. 2000; Kataoka et al. 2004a) is unknown.
334 Interestingly, however, an Ala residue at position 113 of AC-I, adjacent to the highly conserved Gln112, is
335 replaced by Asp in the high-affinity SULTR1 subfamily, by Ser in the low-affinity SULTR2 subfamily, and by
336 Gly in the enigmatic SULTR3 (AC-IIB) subfamilies (Figure 4, shaded in yellow). In addition, a basic Lys/Arg
337 residue at position 118 is conserved in nearly all SULTRs except for the low-affinity SULTR2 where it is
338 replaced by Thr or Asn (Figure 4, shaded in yellow). Most of the other substitutions at or near the
339 binding pocket lining are not predicted to affect its catalytic properties.
340
341 **Expression divergence and tissue partitioning in expanded *SULTR3* subfamilies**
342 We analyzed the expression of poplar *SULTR* genes using published RNA-Seq datasets, including the JGI
343 tissue atlas from *P. tremula* × *P. alba* INRA 717-1B4 (Zhou et al. 2023). The AC-I *PtaSULTR4* transcripts
344 were detected in all tissues, with *PtaSULTR4.1* being higher than *PtaSULTR4.2* in mature leaves and
345 young elongating roots (Figure 5). Among the AC-IIA members, the two *PtaSULTR1.1* paralogs were
346 expressed in all tissues but showed distinct tissue preferences. *PtaSULTR1.1a* transcript levels were
347 highest in the bark and xylem of the primary stem (internodes 3-6), while *PtaSULTR1.1b* transcripts were

348 low overall but predominated in roots (Figure 5). The *PtaSULTR2.1* paralogs were generally poorly
349 expressed, except in primary xylem where *PtaSULTR2.1b* was strongly expressed. *PtaSULTR2.2*, the
350 ancient inverted tandem duplicate of *PtaSULTR1.1b*, was well expressed throughout the plant with a
351 clear bias for mature leaves (Figure 5).
352
353 Members of the expanded AC-IIB subfamilies exhibited substantial expression divergence or partitioning,
354 with many showing restricted tissue distribution. For instance, the singleton *PtaSULTR3.5* was only
355 detected at low levels in buds and roots (Figure 5). No bark expression was detected for the *PtaSULTR3.1*
356 and *PtaSULTR3.5* (AC-IIB-1) subfamilies. *PtaSULTR3.1a* was poorly expressed overall, but the paralogous
357 *PtaSULTR3.1b* was well expressed in both leaves and roots (Figure 5). The *PtaSULTR3.2* pair showed
358 clearly xylem-biased and developmentally conditioned expression in stems. Both genes were expressed
359 in primary xylem, but *PtaSULTR3.2a* expression increased substantially in secondary xylem as
360 *PtaSULTR3.2b* expression sharply decreased (Figure 5). Primary xylem-biased expression was also
361 detected for the *PtaSULTR3.4* paralogs, with *PtaSULTR3.4a* being higher than *PtaSULTR3.4b. PtaSULTR3.3*
362 is the only AC-IIB subfamily that showed broad tissue expression, but this expression was partitioned
363 between the paralogs. *PtaSULTR3.3a* was mainly expressed in stem and root tissues, whereas
364 *PtaSULTR3.3b* was detected in leaves (Figure 5).
365
366 **Differential responses of xylem-biased *SULTR3.2a* and *SULTR3.4a* to drought and lignification**
367 Some *PtaSULTR* genes have previously been shown to be sensitive to seasonal changes in wood or to
368 drought stress in leaves and roots based on qRT-PCR studies (Dürr et al. 2010; Malcheska et al. 2013;
369 Malcheska et al. 2017). Here, we explored data from stress and transgenic perturbation RNA-Seq
370 experiments (Xue et al. 2016; Tsai et al. 2020; Swamy et al. 2015) to analyze *SULTR* expression responses,
371 focusing particularly on vascular tissues and the expanded *SULTR3* subfamilies (excluding *PtaSULTR3.5*
372 which was below detection). We observed a partitioned drought response both between subfamilies and
373 between paralogs, consistent with the partitioned tissue preference discussed above (Figures 5-6). The
374 dominant *PtaSULTR3.1* subfamily members, *PtaSULTR3.1b* and *PtaSULTR3.2a*, were significantly
375 upregulated by drought in roots and secondary xylem, respectively, while their poorly expressed
376 paralogs were unresponsive (Figure 6). Similar patterns were observed for the *PtaSULTR3.3* and
377 *PtaSULTR3.4* subfamilies, with the dominant *PtaSULTR3.3a* and *PtaSULTR3.4a* significantly upregulated in
378 drought-stressed bark and root, respectively. Although both *PtaSULTR3.1b* and *PtaSULTR3.4a* were
379 drought-responsive in root tissue, the expression of *PtaSULTR3.1b* was an order of magnitude higher

than that of *PtaSULTR3.4a*. In summary, the data showed that each of the three vascular tissues responded to drought stress by upregulating a distinct AC-IIB (*SULTR3*) subfamily member: *PtaSULTR3.1b* in roots, *PtaSULTR3.2a* in xylem, and *PtaSULTR3.3a* in bark.

We next examined *SULTR3* expression in the secondary xylem (internodes 60-80) of erect or inclined poplar trees during normal wood (NW) or tension wood (TW) development (Swamy et al. 2015). The secondary xylem-predominant *PtaSULTR3.2a* was down-regulated by more than 70% in TW (Figure 7A). TW in angiosperms is characterized by cellulose-enriched gelatinous fibers with overall reduced lignin accrual (Groover 2016). To further investigate the potential connection between *SULTR3* expression and cell wall composition, we utilized RNA-Seq data from young developing xylem (15 to 30 cm from the shoot tip) of *4CL1* (*4-coumarate:CoA ligase*)-knockout poplar with reduced lignin accumulation (Zhou et al. 2015; Tsai et al. 2020). We observed a 40-50% downregulation of both *PtaSULTR3.2a* and *PtaSULTR3.4a* in the low-lignin *4cl1* mutants (Figure 7B). Interestingly, *PtaSULTR3.4a* transcript abundance was higher than that of *PtaSULTR3.2a*, presumably reflecting the younger stems used in this study compared to the TW experiment (Swamy et al. 2015; Tsai et al. 2020). Together, these results suggest that *PtaSULTR3.2a* expression correlated positively with lignification during stem development, while *PtaSULTR3.4a*'s response was conditional and limited to primary xylem, where it was most highly expressed (Figures 5 and 7).

Gene coexpression of *PtaSULTR3.2a* with lignification and *PtaSULTR3.4a* with phosphorus starvation response

To further explore the functional association of the xylem-biased *PtaSULTR3.2a* and *PtaSULTR3.4a*, we performed gene-centered coexpression analysis using 179 published RNA-seq datasets from xylem, bark, and stem tissues (Supplemental Table 2). We identified a small group of 70 genes co-expressed with *PtaSULTR3.2a* (Pearson correlation coefficient or PCC ≥ 0.7), which showed significant enrichment in Gene Ontology (GO) terms associated with "regulation of secondary cell wall biogenesis" and "defense response" (Figure 8A). Specifically, *PtaSULTR3.2* was coexpressed with genes encoding multiple lignin biosynthetic pathway enzymes, including phenylalanine ammonia lyase (PAL), 4-coumarate:CoA ligase (4CL), hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase, ferulic acid 5-hydroxylase, C1 metabolism (serine hydroxymethyltransferase and S-adenosylmethionine synthase), and transcriptional factors regulating secondary wall biogenesis, such as SND (secondary wall-associated NAC domain) and MYB proteins (Zhong et al. 2011) (Supplemental Table 3). We then plotted PCC

values from 46 genes involved in lignin biosynthesis, C1 metabolism, and their upstream transcriptional regulators (Supplemental Table 3). We found a median PCC of 0.63 with *PtaSULTR3.2a* (median = 0.63) but no correlation was observed with *PtaSULTR3.4a* (Figure 8B). For *PAL* and *4CL* gene families involved in both lignin and flavonoid biosynthesis, *PtaSULTR3.2a* coexpression was observed specifically for lignin-associated *PtaPAL2* (Kao et al. 2002), *Pta4CL1* and *Pta4CL5* isoforms (Harding et al. 2002; Tsai et al. 2020). These results further substantiate findings from Figure 7, indicating a previously undescribed link between *SULTR3.2a* and lignification.

In comparison, more than five times as many genes (360) were co-expressed with *PtaSULTR3.4a*, with significant GO enrichment in “cellular response to phosphate starvation”, “sulfolipid biosynthetic process”, “response to nitrate”, and “transporter activity”, including “sulfate transport” and “phosphate ion transport” (Figure 8A). Indeed, *PtaSULTR3.4a* was strongly co-expressed with *PtaSULTR3.4b* (PCC = 0.91) and *PtaSULTR2.1b* (0.82), as well as genes encoding multiple phosphate transporters (Supplemental Table 3), consistent with previous reports associating the SULTR3.4 subfamily with phosphorus transport (Yamaji et al. 2017; Ding et al. 2020). *PtaSULTR3.4a* was also highly co-expressed with genes known to be sensitive to phosphorus starvation, including orthologs of SPX domain-containing proteins and purple acid phosphatases (Supplemental Table 3) (Duan et al. 2008; Wang et al. 2011; Gan et al. 2015; Chen et al. 2022). Our analysis suggests a close coordination of *PtaSULTR3.4a* with a network of transporters and phosphorus responses genes.

DISCUSSION

The evolutionary expansion and functional diversification of SULTR transporters from bacteria to higher plants has been extensively discussed (Takahashi et al. 2012). However, important questions remain, especially regarding the large and vascular plant-specific SULTR3 group, whose members exhibit little to no sulfate transport activity on their own (Takahashi et al. 2000; Kataoka et al. 2004a). SULTR3 members exhibit variability in their membrane localization and have reported functions ranging from facilitating chloroplast sulfate assimilation and enhancing sulfate transport by other SULTR partners to phosphorus distribution via the vasculature (Kataoka et al. 2004a; Cao et al. 2013; Zhao et al. 2016; Yamaji et al.

2017; Chen et al. 2019; Ding et al. 2020; Gu et al. 2022). By including multiple early-diverging species at key evolutionary junctures for phylogenetic inference, we have updated the SULTR family classification into seven distinct subfamilies dating back to basal angiosperms, with four belonging to the enigmatic SULTR3 group. We further identified a perennial-specific subgroup of the SULTR3.1 subfamily, with the poplar member *PtaSULTR3.2a* being highly coexpressed with lignin biosynthetic genes and responsive to lignin alterations.

Early evolution of seven divergent SULTR subfamilies in angiosperms

Our phylogenetic analysis supports several features of the plant SULTR family evolution previously suggested by Takahashi et al. (2012), based on the placement of moss and spikemoss members. These include the split of AC-I (SULTR4) and AC-II (SULTR1-SULTR3) clades early during land plant evolution, and the subsequent split of AC-IIA (SULTR1-SULTR2) and AC-IIB (SULTR3) subclades, with AC-IIB arising specifically in vascular plants (Figure 1). While AC-IIA is further split into the SULTR1 and SULTR2 subfamilies, the much large AC-IIB was previously considered as one (SULTR3) subfamily with four subgroups (Takahashi et al. 2012). However, our phylogenetic reconstruction with basal angiosperm lineages uncovered four strongly supported SULTR3 subfamilies that predate the split of Amborella and tulip tree (Figure 1). Examination of the recently sequenced camphor tree (*Cinnamomum kanehirae* v3) (Chaw et al. 2019) and water lily (*Nymphaea colorata* v1.2)(Zhang et al. 2020) genomes available on Phytozome v13 confirmed the presence of four AC-IIB subfamilies in early diverging angiosperm taxa, suggesting their origin in the last common ancestor of all flowering plants.

Gene structure analysis provided further support for distinct evolution of the AC-IIB subfamilies. It is important to note that gene structure changes, such as intron losses, can impact expression and contribute to functional diversification as effectively, if not more so, than mutations within exons (Morello and Breviario 2008). This was evidenced by preferential yet partitioned expression, both among *SULTR3* subfamilies and between gene duplicates, in vascular tissues of poplar (Figure 5). Drought stress stimulated expression of *PtaSULTR3.1b*, *PtaSULTR3.2a*, and *PtaSULR3.3a* in a root-, xylem-, and bark-biased manner, respectively (Figure 6). These examples serve to illustrate the diverse roles of the four SULTR3 subfamilies beyond the historically established vacuolar (SULTR4), high-affinity (SULTR1), and low-affinity (SULTR2) sulfate transport activities. The four SULTR3 subfamilies likely differentiated as mechanisms for internal distribution and utilization of sulfate evolved in angiosperms (Takahashi 2019).

Phylogenetic insights into substrate binding pocket sequences

Despite differences in tissue expression and transport kinetics, the proton and sulfate binding pocket residues predicted to be essential for activity (Wang et al. 2021) are conserved in AC-I SULTR4 and AC-IIA SULTR1/SULTR2 subfamilies (Figure 4). However, catalytic divergence among the expanded AC-IIB (SULTR3) coincides with several subfamily-specific substitutions within or near the binding pocket. AC-IIB-1 SULTR3.1 and SULTR3.5 are the only SULTR3 subfamilies with demonstrated sulfate uptake activity in yeast (Krusell et al. 2005; Xun et al. 2021). Mutations to AtSULTR3.1 were more detrimental to sulfate uptake by isolated chloroplasts than mutations to any of the other AC-IIB members (Cao et al. 2013; Chen et al. 2019). Another SULTR3.1 subfamily member, AtSULTR3.2, has been convincingly linked with sulfate translocation in developing seeds of Arabidopsis (Zuber et al. 2010). SULTR3.1 and SULTR3.5 differ from the canonical AC-I and AC-IIA at position 390 (Pro vs. Ser). Interestingly, the basic residue conserved at position 118 of the high-affinity SULTR1 is also conserved in most SULTR3.1/SULTR3.5 members but is replaced by a non-basic residue in the perennial-specific subgroup (e.g., PtaSULTR3.2s) as in the low affinity SULTR2. SULTR3.5 is unique in having a Cys at position 155 (Figure 4). Cys can form intra- and inter-protein disulfide linkages although its steric ability to do so depends on the tertiary structure. Also in SULTR3.5, Lys replaces Arg at conserved residue 393 important for sulfate binding pocket hydrogen-bonding, and thus stability and function (Wang et al. 2021). Although a basic residue Arg(Lys)393 is preserved across eukaryotic homologs (Wang et al. 2021), Lys can be methylated with additional consequences for protein function, most classically in the cases of histones and calmodulins (Harding et al. 1997; Martin and Zhang 2005). Non-histone Lys methylation as a regulator of hydrogen-bonding and other protein properties affecting diverse cellular processes has been recently reviewed (Wu et al. 2017; Cornett et al. 2019).

AC-IIB-2 subfamilies SULTR3.3 and SULTR3.4 differ from the other SULTRs at position 153, replacing Ala with Ser in the sulfate-binding pocket (Figure 4). Interestingly, residue Ser390 is conserved in SULTR3.4 but is replaced by Ala in SULTR3.3. In multiple studies, SULTR3.4 subfamily members have been shown to transport phosphate instead of sulfate (Yamaji et al. 2017; Ding et al. 2020; Gu et al. 2022). For the SULTR3.3 subfamily with dual substitutions at positions 153 and 390, no *in vitro* transport activity for phosphate or sulfate was detected with rice OsSULTR3.3 but its mutations reduced phytate and total phosphorus concentrations in the rice grains (Zhao et al. 2016). The functional divergence of AC-IIB-2 members as phosphorus distribution transporters may hint at the importance of Ala153 for substrate selectivity, although other sequence and protein conformational differences could also contribute.

508

509 ***PtaSULTR3.4* coexpression with phosphorus transport and starvation response genes**

510 Poplar *PtaSULTR3.4a* and its less expressed genome duplicate *PtaSULTR3.4b* were preferentially
511 expressed in primary xylem. Of the seven Salicoid genome duplicate pairs in the SULTR family,
512 *PtaSULTR3.4s* are the only pair that retains expression conservation in stem vascular tissues. Consistent
513 with a role of the SULTR3.4 subfamily in phosphorus distribution and metabolism (Yamaji et al. 2017;
514 Ding et al. 2020), *PtaSULTR3.4a*-coexpressed genes included several phosphate transporters and
515 phosphate starvation response markers (Gan et al. 2015; Chen et al. 2022). This group also included
516 genes encoding nitrate-responsive regulators, transporters for nitrate, amino acids, auxin, and various
517 metal ions, as well as plasma membrane H^+-ATPases necessary to energize transporters for
518 transmembrane nutrient uptake (Palmgren 2001). The coordinated acquisition of nitrogen and
519 phosphorus, and the crosstalk between nitrate- and phosphorus-responsive signaling pathways have
520 been extensively reported, including in poplar (Gan et al. 2015; Hu and Chu 2020; Ueda et al. 2020). This
521 coordination also extends to other mineral nutrients (Raddatz et al. 2020; Wei et al. 2023). Similarly,
522 sulfur assimilation is sensitive to availability of not only sulfate but also nitrate (Kopriva et al. 2004). Our
523 gene coexpression analysis implicated SULTR3.4 as an important player in these coordinated responses.

524

525 **Secondary xylem-biased *PtaSULTR3.2a* expression and lignification**

526 Extensive prior work has illustrated temporal associations between *SULTR* gene expression and sulfate
527 circulation in poplar stems throughout the year (Rennenberg et al. 2007; Dürr et al. 2010; Malcheska et
528 al. 2013). The sulfate assimilation pathway interfaces with C1 metabolism for the production of S-
529 adenosyl Met (SAM), which is a methyl donor for the biosynthesis of lignin and other cell wall
530 polysaccharides (Rajinikanth et al. 2007; Srivastava et al. 2015; Zhang et al. 2019). However, there is little
531 information linking SULTRs with wood formation. A connection was recently suggested via a xylem-
532 expressed microRNA, *miR395c*, which regulated sulfate assimilation genes and low-affinity SULTR2.1b to
533 affect ABA synthesis and secondary cell wall biosynthesis in poplar (Liu et al. 2022).

534

535 In poplar, AC-IIB-1 member *PtaSULTR3.2a* is the dominant *SULTR* transcript in secondary xylem
536 undergoing secondary cell wall thickening. A functional association between PtaSULTR3.2a and
537 lignification received support from reduced *PtaSULTR3.2a* expression in low lignin xylem, either via
538 genetic perturbation in *4cl1* mutants or from TW of trees subjected to stem bending (Figure 7). In
539 addition, xylem *PtaSULTR3.2a* expression increased during drought stress which generally results in

increased lignification (Moura et al. 2010). We show strong coexpression of *PtaSULTR3.2a* with lignin biosynthesis, C1 metabolism, and secondary cell wall biogenesis regulatory genes (Rajinikanth et al. 2007; Zhong et al. 2011; Zhang et al. 2019; Donev et al. 2023), accounting for 19% of the 69 *PtaSULTR3.2a*-coexpressed genes. While strictly correlative, the conclusion that PtaSULTR3.2a was more connected than other PtaSULTRs with lignification in a woody perennial was supported by data from a very wide range of experimental approaches (Supplemental Table 2). In addition, sulfur-containing glutathione plays important role in detoxification and scavenging of reactive oxygen species (Takahashi et al. 2011), with documented involvement in lignin biosynthesis in response to oxidative stress perturbations (Van de Mortel et al. 2008; Tsai et al. 2020; Noor et al. 2024; Yuan et al. 2024). Given that PtaSULT3.2a belongs to a small subgroup of the SULTR3.1 subfamily that is retained only in perennial dicot species, the preferential expression of *PtaSULTR3.2a* in secondary xylem and its strong coexpression with lignification genes are all consistent with perennial-specific functional diversification.

In summary, our analysis suggests that the enigmatic SULTR3 contains four distinct subfamilies with an ancient origin predating angiosperm evolution. Variations in exon-intron structure, substrate binding pocket residues, and tissue expression preferences likely contribute to the functional diversification of the four SULTR3 subfamilies. The previously reported neofunctionalization of the SULTR3.4 subfamily associated with phosphorus distribution (Yamaji et al. 2017; Ding et al. 2020; Gu et al. 2022) gained credence in poplar based on coexpression analysis. In addition, we provided evidence supporting the neofunctionalization of PtaSULTR3.2a associated with lignification in poplar and likely other perennial species. These functional predictions await future experimental verification. Finally, this work highlights the need for phylogeny-informed gene and gene family nomenclatures as discussed elsewhere in this special issue.

ACKNOWLEDGEMENTS

The authors thank the advice of Ethan Baldwin and Jim Leebens-Mack on phylogenetic analysis and Ran Zhou on bioinformatics analysis.

AUTHORS' CONTRIBUTIONS

C.-J.T. conceived the idea, S.M.S. performed all analyses with assistance from C.H. for RNA-seq and gene coexpression, S.M.S. drafted the manuscript, S.A.H. and C.-J.T. revised the manuscript.

SUPPLEMENTARY DATA

Table S1. List of SULTR sequences used in the phylogenetic analysis

Table S2. List of RNA-Seq datasets from *Populus tremula* × *alba* stem tissues used in coexpression analysis

Table S3. List of genes co-expressed with *PtaSULTR3.2a* and *PtaSULTR3.4a*

FUNDING

The work was supported by the Center for Bioenergy Innovation (CBI), U.S. Department of Energy, Office of Science, Biological and Environmental Research Program under Award Number ERKP886. S.M.S was supported by the Foundation for Food and Agriculture Research fellowship program.

CONFLICT OF INTEREST: None declared.

DATA AVAILABILITY STATEMENT

Gene models and genome versions for protein sequences used for phylogenetic analysis are provided in Supplementary Table S1. Accession numbers for RNA-Seq datasets used in gene expression analysis are SRR6255832 to SRR6255842 for tissue expression, RR1292611, SRR1292615, and SRR1295980-SRR1295983 for stem bending response, SRP041959 for drought response, and PRJNA589632 for lignin perturbation response. Accession numbers for the 179 RNA-Seq datasets used for gene coexpression analysis are provided in Supplementary Table S2.

REFERENCES

Batool, S., V.V. Uslu, H. Rajab, N. Ahmad, R. Waadt, D. Geiger, M. Malagoli, C.B. Xiang, R. Hedrich, H. Rennenberg, C. Herschbach, R. Hell and M. Wirtz. 2018. Sulfate is incorporated into cysteine to trigger ABA production and stomatal closure. Plant Cell. 30:2973-2987.

Bauer, H., P. Ache, S. Lautner, J. Fromm, W. Hartung, Khaled A.S. Al-Rasheid, S. Sonnewald, U. Sonnewald, S. Kneitz, N. Lachmann, Ralf R. Mendel, F. Bittner, Alistair M. Hetherington and R. Hedrich. 2013. The stomatal response to reduced relative humidity requires guard cell-autonomous ABA synthesis. Current Biology. 23:53-57.

Bianchini, G. and P. Sánchez-Baracaldo. 2024. TreeViewer: Flexible, modular software to visualise and manipulate phylogenetic trees. Ecology and Evolution. 14:e10873.

Bray, N.L., H. Pimentel, P. Melsted and L. Pachter. 2016. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 34:525-7.

Cao, M.-J., Z. Wang, M. Wirtz, R. Hell, D.J. Oliver and C.-B. Xiang. 2013. SULTR3;1 is a chloroplast-localized sulfate transporter in Arabidopsis thaliana. The Plant Journal. 73:607-616.

Cao, M.-J., Z. Wang, Q. Zhao, J.-L. Mao, A. Speiser, M. Wirtz, R. Hell, J.-K. Zhu and C.-B. Xiang. 2014. Sulfate availability affects ABA levels and germination response to ABA and salt stress in rabidopsis thaliana. Plant Journal. 77:604-615.

Chaw, S.-M., Y.-C. Liu, Y.-W. Wu, H.-Y. Wang, C.-Y.I. Lin, C.-S. Wu, H.-M. Ke, L.-Y. Chang, C.-Y. Hsu, H.-T. Yang, E. Sudianto, M.-H. Hsu, K.-P. Wu, L.-N. Wang, J.H. Leebens-Mack and I.J. Tsai. 2019. Stout camphor tree genome fills gaps in understanding of flowering plant genome evolution. Nature Plants. 5:63-73.

Chen, N., S. Tong, J. Yang, J. Qin, W. Wang, K. Chen, W. Shi, J. Li, J. Liu and Y. Jiang. 2022. PtoWRKY40 interacts with PtoPHR1-LIKE3 while regulating the phosphate starvation response in poplar. Plant Physiology. 190:2688-2705.

Chen, Y., X. Xiao, R. Yang, Z. Sun, S. Yang, H. Zhang, B. Xing, Y. Li, Q. Liu, Q. Lu, Y. Shi, Y. Yuan, C. Miao and P. Li. 2024. Genome-wide identification and expression-pattern analysis of sulfate transporter (SULTR) gene family in cotton under multiple abiotic stresses and fiber development. Functional & Integrative Genomics. 24:108.

Chen, Z., P.X. Zhao, Z.Q. Miao, G.F. Qi, Z. Wang, Y. Yuan, N. Ahmad, M.J. Cao, R. Hell, M. Wirtz and C.B. Xiang. 2019. SULTR3s function in chloroplast sulfate uptake and affect ABA biosynthesis and the stress response. Plant Physiol. 180:593-604.

Cornett, E.M., L. Ferry, P.A. Defossez and S.B. Rothbart. 2019. Lysine methylation regulators moonlighting outside the epigenome. Mol Cell. 75:1092-1101.

Ding, G., G.J. Lei, N. Yamaji, K. Yokosho, N. Mitani-Ueno, S. Huang and J.F. Ma. 2020. Vascular cambium-localized AtSPDT mediates xylem-to-phloem transfer of phosphorus for its preferential distribution in Arabidopsis. Mol Plant. 13:99-111.

Ding, Y., X. Zhou, L. Zuo, H. Wang and D. Yu. 2016. Identification and functional characterization of the sulfate transporter gene GmSULTR1;2b in soybean. BMC Genomics. 17:373.

Donev, E.N., M. Derba-Maceluch, Z. Yassin, M.L. Gandla, S. Pramod, E. Heinonen, V. Kumar, G. Scheepers, F. Vilaplana, U. Johansson, M. Hertzberg, B. Sundberg, S. Winestrand, A. Hörnberg, B. Alriksson, L.J. Jönsson and E.J. Mellerowicz. 2023. Field testing of transgenic aspen from large greenhouse screening identifies unexpected winners. Plant Biotechnology Journal. 21:1005-1021.

Duan, K., K. Yi, L. Dang, H. Huang, W. Wu and P. Wu. 2008. Characterization of a sub-family of Arabidopsis genes with the SPX domain reveals their diverse functions in plant tolerance to phosphorus starvation. The Plant Journal. 54:965-975.

Dürr, J., H. Bücking, S. Mult, H. Wildhagen, K. Palme, H. Rennenberg, F. Ditengou and C. Herschbach. 2010. Seasonal and cell type specific expression of sulfate transporters in the phloem of Populus reveals tree specific characteristics for SO(4)(2-) storage and mobilization. Plant Mol Biol. 72:499-517.

Ernst, L., J.Q. Goodger, S. Alvarez, E.L. Marsh, B. Berla, E. Lockhart, J. Jung, P. Li, H.J. Bohnert and D.P. Schachtman. 2010. Sulphate as a xylem-borne chemical signal precedes the expression of ABA biosynthetic genes in maize roots. J Exp Bot. 61:3395-405.

Ewels, P.A., A. Peltzer, S. Fillinger, H. Patel, J. Alneberg, A. Wilm, M.U. Garcia, P. Di Tommaso and S. Nahnsen. 2020. The nf-core framework for community-curated bioinformatics pipelines. Nature Biotechnology. 38:276-278.

Gallardo, K., P.-E. Courty, C. Le Signor, D. Wipf and V. Vernoud. 2014. Sulfate transporters in the plant's response to drought and salinity: regulation and possible functions. Frontiers in Plant Science. 5

Gan, H., Y. Jiao, J. Jia, X. Wang, H. Li, W. Shi, C. Peng, A. Polle and Z.-B. Luo. 2015. Phosphorus and nitrogen physiology of two contrasting poplar genotypes when exposed to phosphorus and/or nitrogen starvation. Tree Physiology. 36:22-38.

Ge, S.X., D. Jung and R. Yao. 2020. ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics. 36:2628-2629.

Goodstein, D.M., S. Shu, R. Howson, R. Neupane, R.D. Hayes, J. Fazo, T. Mitros, W. Dirks, U. Hellsten, N. Putnam and D.S. Rokhsar. 2012. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Research. 40:D1178-D1186.

Gu, M., H. Huang, H. Hisano, G. Ding, S. Huang, N. Mitani-Ueno, K. Yokosho, K. Sato, N. Yamaji and J.F. Ma. 2022. A crucial role for a node-localized transporter, HvSPDT, in□loading phosphorus into barley grains. New Phytologist. 234:1249-1261.

Harding, S.A., J. Leshkevich, V.L. Chiang and C.J. Tsai. 2002. Differential substrate inhibition couples kinetically distinct 4-coumarate:coenzyme A ligases with spatially distinct metabolic roles in quaking aspen. Plant Physiology. 128:428-438.

Harding, S.A., S.H. Oh and D.M. Roberts. 1997. Transgenic tobacco expressing a foreign calmodulin gene shows an enhanced production of active oxygen species. The EMBO Journal. 16:1137-1144.

Heidari, P., S. Hasanzadeh, S. Faraji, S. Ercisli and F. Mora-Poblete. 2023. Genome-wide characterization of the sulfate transporter gene family in oilseed crops: *Camelina sativa* and *Brassica napus*. Plants. 12:628.

Hu, B. and C. Chu. 2020. Nitrogen–phosphorus interplay: old story with molecular tale. New Phytologist. 225:1455-1460.

Huang, Q., M. Wang and Z. Xia. 2018. The SULTR gene family in maize (*Zea mays* L.): Gene cloning and expression analyses under sulfate starvation and abiotic stress. Journal of Plant Physiology. 220:24-33.

Jones, D.T., W.R. Taylor and J.M. Thornton. 1992. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci. 8:275-82.

Kalyaanamoorthy, S., B.Q. Minh, T.K.F. Wong, A. von Haeseler and L.S. Jermiin. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nature Methods. 14:587-589.

Kao, Y.Y., S.A. Harding and C.J. Tsai. 2002. Differential expression of two distinct phenylalanine ammonia-lyase genes in condensed tannin-accumulating and lignifying cells of quaking aspen. Plant Physiology. 130:796-807.

Kataoka, T., N. Hayashi, T. Yamaya and H. Takahashi. 2004a. Root-to-shoot transport of sulfate in Arabidopsis. Evidence for the role of SULTR3;5 as a component of low-affinity sulfate transport system in the root vasculature. Plant Physiol. 136:4198-204.

Kataoka, T., A. Watanabe-Takahashi, N. Hayashi, M. Ohnishi, T. Mimura, P. Buchner, M.J. Hawkesford, T. Yamaya and H. Takahashi. 2004b. Vacuolar sulfate transporters are essential determinants controlling internal distribution of sulfate in Arabidopsis. Plant Cell. 16:2693-704.

Ko, J.-H., C. Prassinos, D. Keathley, K.-H. Han and C. Li. 2011. Novel aspects of transcriptional regulation in the winter survival and maintenance mechanism of poplar. Tree Physiology. 31:208-225.

Kopriva, S., T. Hartmann, G. Massaro, P. Hönicke and H. Rennenberg. 2004. Regulation of sulfate assimilation by nitrogen and sulfur nutrition in poplar trees. Trees. 18:320-326.

Kopriva, S., M. Malagoli and H. Takahashi. 2019. Sulfur nutrition: impacts on plant development, metabolism, and stress responses. Journal of Experimental Botany. 70:4069-4073.

Krusell, L., K. Krause, T. Ott, G. Desbrosses, U. Krämer, S. Sato, Y. Nakamura, S. Tabata, E.K. James, N. Sandal, J. Stougaard, M. Kawaguchi, A. Miyamoto, N. Suganuma and M.K. Udvardi. 2005. The sulfate transporter SST1 is crucial for symbiotic nitrogen fixation in Lotus japonicus root nodules. Plant Cell. 17:1625-36.

Kumar, S., G. Stecher, M. Li, C. Knyaz and K. Tamura. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Molecular Biology and Evolution. 35:1547-1549.

Liu, C., D. Ma, Z. Wang, N. Chen, X. Ma and X.-Q. He. 2022. MiR395c regulates secondary xylem development through sulfate metabolism in poplar. Frontiers in Plant Science. 13

Madeira, F., N. Madhusoodanan, J. Lee, A. Eusebi, A. Niewielska, A.R.N. Tivey, S. Meacham, R. Lopez and S. Butcher. 2024. Using EMBL-EBI Services via Web Interface and Programmatically via Web Services. Current Protocols. 4:e1065.

Malcheska, F., A. Ahmad, S. Batool, H.M. Müller, J. Ludwig-Müller, J. Kreuzwieser, D. Randewig, R. Hänsch, R.R. Mendel, R. Hell, M. Wirtz, D. Geiger, P. Ache, R. Hedrich, C. Herschbach and H. Rennenberg. 2017. Drought-enhanced xylem sap sulfate closes stomata by affecting ALMT12 and guard cell ABA synthesis. Plant Physiol. 174:798-814.

Malcheska, F., A. Honsel, H. Wildhagen, J. Dürr, C. Larisch, H. Rennenberg and C. Herschbach. 2013. Differential expression of specific sulphate transporters underlies seasonal and spatial patterns of sulphate allocation in trees. Plant, Cell & Environment. 36:1285-1295.

Martin, C. and Y. Zhang. 2005. The diverse functions of histone lysine methylation. Nature Reviews Molecular Cell Biology. 6:838-849.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011. 17:3.

Maruyama-Nakashita, A., A. Watanabe-Takahashi, E. Inoue, T. Yamaya, K. Saito and H. Takahashi. 2015. Sulfur-responsive elements in the 3′-nontranscribed intergenic region are essential for the induction of SULFATE TRANSPORTER 2;1 gene expression in arabidopsis roots under sulfur deficiency. Plant Cell. 27:1279-1296.

Morello, L. and D. Breviario. 2008. Plant spliceosomal introns: not only cut and paste. Curr Genomics. 9:227-38.

Moura, J.C.M.S., C.A.V. Bonine, J.D.O.F. Viana, M.C. Dornelas and P. Mazzafera. 2010. Abiotic and biotic stresses and changes in the lignin content and composition in plants. Journal of Integrative Plant Biology. 52:360-376.

Nguyen, L.T., H.A. Schmidt, A. von Haeseler and B.Q. Minh. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32:268-74.

Noor, I., H. Sohail, C. Wentao, K. Zhu, M. Hasanuzzaman, G. Li and J. Liu. 2024. Phosphorus-induced restructuring of the ascorbate–glutathione cycle and lignin biosynthesis alleviates manganese toxicity in peach roots. Tree Physiology. 44

Notredame, C., D.G. Higgins and J. Heringa. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol. 302:205-17.

Palmgren, M.G. 2001. PLANT PLASMA MEMBRANE H+-ATPases: Powerhouses for Nutrient Uptake. Annual Review of Plant Biology. 52:817-845.

Pfautsch, S., T. Hölttä and M. Mencuccini. 2015. Hydraulic functioning of tree stems—fusing ray anatomy, radial transfer and capacitance. Tree Physiology. 35:706-722.

Puresmaeli, F., P. Heidari and S. Lawson. 2023. Insights into the sulfate transporter gene family and its expression patterns in durum wheat seedlings under salinity. Genes. 14:333.

Raddatz, N., L. Morales de los Ríos, M. Lindahl, F.J. Quintero and J.M. Pardo. 2020. Coordinated Transport of Nitrate, Potassium, and Sodium. Frontiers in Plant Science. 11

Rajinikanth, M., S.A. Harding and C.-J. Tsai. 2007. The glycine decarboxylase complex multienzyme family in Populus. Journal of Experimental Botany. 58:1761-1770.

Rennenberg, H., C. Herschbach, K. Haberer and S. Kopriva. 2007. Sulfur metabolism in plants: Are trees different? Plant Biology. 9:620-637.

Rouached, H., M. Wirtz, R. Alary, R.d. Hell, A.B. Arpat, J.-C. Davidian, P. Fourcroy and P. Berthomieu. 2008. Differential regulation of the expression of two high-affinity sulfate transporters, SULTR1.1 and SULTR1.2, in *Arabidopsis* Plant Physiology. 147:897-911.

Shannon, P., A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker. 2003. Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Research. 13:2498-2504.

Srivastava, A.C., F. Chen, T. Ray, S. Pattathil, M.J. Peña, U. Avci, H. Li, D.V. Huhman, J. Backe, B. Urbanowicz, J.S. Miller, M. Bedair, C.E. Wyman, L.W. Sumner, W.S. York, M.G. Hahn, R.A. Dixon, E.B. Blancaflor and Y. Tang. 2015. Loss of function of folylpolyglutamate synthetase 1 reduces lignin content and improves cell wall digestibility in Arabidopsis. Biotechnology for Biofuels. 8:224.

Swamy, P.S., H. Hu, S. Pattathil, V.J. Maloney, H. Xiao, L.-J. Xue, J.-D. Chung, V.E. Johnson, Y. Zhu, G.F. Peter, M.G. Hahn, S.D. Mansfield, S.A. Harding and C.-J. Tsai. 2015. Tubulin perturbation leads to unexpected cell wall modifications and affects stomatal behaviour in Populus. Journal of Experimental Botany. 66:6507-6518.

Takahashi, H. 2019. Sulfate transport systems in plants: functional diversity and molecular mechanisms underlying regulatory coordination. Journal of Experimental Botany. 70:4075-4087.

Takahashi, H., P. Buchner, N. Yoshimoto, M.J. Hawkesford and S.-H. Shiu. 2012. Evolutionary relationships and functional diversity of plant sulfate transporters. Frontiers in Plant Science. 2

Takahashi, H., S. Kopriva, M. Giordano, K. Saito and R. Hell. 2011. Sulfur assimilation in photosynthetic organisms: Molecular functions and regulations of transporters and assimilatory enzymes. Annual Review of Plant Biology. 62:157-184.

Takahashi, H., A. Watanabe-Takahashi, F.W. Smith, M. Blake-Kalff, M.J. Hawkesford and K. Saito. 2000. The roles of three functional sulphate transporters involved in uptake and translocation of sulphate in Arabidopsis thaliana. The Plant Journal. 23:171-182.

Takahashi, H., M. Yamazaki, N. Sasakura, A. Watanabe, T. Leustek, J.A. Engler, G. Engler, M. Van Montagu and K. Saito. 1997. Regulation of sulfur assimilation in higher plants: a sulfate transporter induced in sulfate-starved roots plays a central role in Arabidopsis thaliana. Proc Natl Acad Sci U S A. 94:11102-7.

Tsai, C.-J., P. Xu, L.-J. Xue, H. Hu, B. Nyamdari, R. Naran, X. Zhou, G. Goeminne, R. Gao, E. Gjersing, J. Dahlen, S. Pattathil, M.G. Hahn, M. Davis, J. Ralph, W.A. Boerjan and S.A. Harding. 2020. Compensatory guaiacyl lignin biosynthesis at the expense of syringyl lignin in *4CL1*-knockout poplar. Plant Physiology. 183:123-136.

Ueda, Y., T. Kiba and S. Yanagisawa. 2020. Nitrate-inducible NIGT1 proteins modulate phosphate uptake and starvation signalling via transcriptional regulation of SPX genes. Plant Journal. 102:448-466.

Van de Mortel, J.E., H. Schat, P.D. Moerland, E.V.L. Van Themaat, S. Van Der Ent, H. Blankestijn, A. Ghandilyan, S. Tsiatsiani and M.G.M. Aarts. 2008. Expression differences for genes involved in lignin, glutathione and sulphate metabolism in response to cadmium in Arabidopsis thaliana and the related Zn/Cd-hyperaccumulator Thlaspi caerulescens. Plant, Cell & Environment. 31:301-324.

Vidmar, J.J., J.K. Schjoerring, B. Touraine and A.D.M. Glass. 1999. Regulation of the hvst1 gene encoding a high-affinity sulfate transporter from Hordeum vulgare. Plant Molecular Biology. 40:883-892.

Wang, L., K. Chen and M. Zhou. 2021. Structure and function of an Arabidopsis thaliana sulfate transporter. Nature Communications. 12:4455.

Wang, L., Z. Li, W. Qian, W. Guo, X. Gao, L. Huang, H. Wang, H. Zhu, J.W. Wu, D. Wang and D. Liu. 2011. The Arabidopsis purple acid phosphatase AtPAP10 is predominantly associated with the root surface and plays an important role in plant tolerance to phosphate limitation. Plant Physiol. 157:1283-99.

Wei, M., M. Zhang, J. Sun, Y. Zhao, S. Pak, M. Ma, Y. Chen, H. Lu, J. Yang, H. Wei, Y. Li and C. Li. 2023. PuHox52 promotes coordinated uptake of nitrate, phosphate, and iron under nitrogen deficiency in Populus ussuriensis. J Integr Plant Biol. 65:791-809.

Wu, Z., J. Connolly and K.K. Biggar. 2017. Beyond histones – the expanding roles of protein lysine methylation. The FEBS Journal. 284:2732-2744.

Xue, L.-J., C.J. Frost, C.-J. Tsai and S.A. Harding. 2016. Drought response transcriptomics are altered in poplar with reduced tonoplast sucrose transporter expression. Scientific Reports. 6:33655.

Xun, M., J. Song, J. Shi, J. Li, Y. Shi, J. Yan, W. Zhang and H. Yang. 2021. Genome-wide identification of sultr genes in *Malus domestica* and low sulfur-induced MhSultr3;1a to increase cysteine-improving growth. Frontiers in Plant Science. 12

Yamaji, N., Y. Takemoto, T. Miyaji, N. Mitani-Ueno, K.T. Yoshida and J.F. Ma. 2017. Reducing phosphorus accumulation in rice grains with an impaired transporter in the node. Nature. 541:92-95.

Yoshimoto, N., H. Takahashi, F.W. Smith, T. Yamaya and K. Saito. 2002. Two distinct high-affinity sulfate transporters with different inducibilities mediate uptake of sulfate in Arabidopsis roots. Plant Journal. 29:465-473.

Yuan, L., J. Dang, J. Zhang, L. Wang, H. Zheng, G. Li, J. Li, F. Zhou, A. Khan, Z. Zhang and X. Hu. 2024. A glutathione S-transferase regulates lignin biosynthesis and enhances salt tolerance in tomato. Plant Physiol. 196:2989-3006.

Zhang, J., M. Li, A.C. Bryan, C.G. Yoo, W. Rottmann, K.A. Winkeler, Cassandra M. Collins, V. Singan, E.A. Lindquist, S.S. Jawdy, L.E. Gunter, N.L. Engle, X. Yang, K. Barry, T.J. Tschaplinski, J. Schmutz, Y. Pu, A.J. Ragauskas, G.A. Tuskan, W. Muchero and J.-G. Chen. 2019. Overexpression of a serine hydroxymethyltransferase increases biomass production and reduces recalcitrance in the bioenergy crop Populus. Sustainable Energy & Fuels. 3:195-207.

Zhang, L., F. Chen, X. Zhang, Z. Li, Y. Zhao, R. Lohaus, X. Chang, W. Dong, S.Y.W. Ho, X. Liu, A. Song, J. Chen, W. Guo, Z. Wang, Y. Zhuang, H. Wang, X. Chen, J. Hu, Y. Liu, Y. Qin, K. Wang, S. Dong, Y. Liu, S. Zhang, X. Yu, Q. Wu, L. Wang, X. Yan, Y. Jiao, H. Kong, X. Zhou, C. Yu, Y. Chen, F. Li, J. Wang, W. Chen, X. Chen, Q. Jia, C. Zhang, Y. Jiang, W. Zhang, G. Liu, J. Fu, F. Chen, H. Ma, Y. Van de Peer and H. Tang. 2020. The water lily genome and the early evolution of flowering plants. Nature. 577:79-84.

Zhao, H., T. Frank, Y. Tan, C. Zhou, M. Jabnoune, A.B. Arpat, H. Cui, J. Huang, Z. He, Y. Poirier, K.-H. Engel and Q. Shu. 2016. Disruption of Os3;3 reduces phytate and phosphorus concentrations and alters the metabolite profile in rice grains. New Phytologist. 211:926-939.

Zhao, Q., J. Geng, Y. Du, S. Li, X. Yuan, J. Zhu, Z. Zhou, Q. Wang and J. Du. 2022. The common bean (*Phaseolus vulgaris*) SULTR gene family: genome-wide identification, phylogeny, evolutionary expansion and expression patterns. Biotechnology & Biotechnological Equipment. 36:724-736.

Zhong, R., R.L. McCarthy, C. Lee and Z.-H. Ye. 2011. Dissection of the transcriptional program regulating secondary wall biosynthesis during wood formation in poplar. Plant Physiology. 157:1452-1468.

Zhou, J., Y. Dong, Y. Liu, Y. Huang, W. Jiang, X. Zheng, H. Zhang, N. Gong and X. Bai. 2024. Identification and expression analysis of sulfate transporter genes family and function analysis of GmSULTR3;1a from soybean. Int J Mol Sci. 25

Zhou, R., J.W. Jenkins, Y. Zeng, S. Shu, H. Jang, S.A. Harding, M. Williams, C. Plott, K.W. Barry, M. Koriabine, M. Amirebrahimi, J. Talag, S. Rajasekar, J. Grimwood, R.J. Schmitz, R.K. Dawe, J. Schmutz and C.-J. Tsai. 2023. Haplotype-resolved genome assembly of *Populus tremula* × *P. alba* reveals aspen-specific megabase satellite DNA. Plant Journal. 16:1003-1017.

Zhou, X., T.B. Jacobs, L.-J. Xue, S.A. Harding and C.-J. Tsai. 2015. Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate:CoA ligase specificity and redundancy. New Phytologist. 208:298-301.

Zuber, H., J.C. Davidian, G. Aubert, D. Aimé, M. Belghazi, R. Lugan, D. Heintz, M. Wirtz, R. Hell, R. Thompson and K. Gallardo. 2010. The seed composition of Arabidopsis mutants for the group 3 sulfate transporters indicates a role in sulfate translocation within developing seeds. Plant Physiol. 154:913-26.

FIGURE LEGENDS

Figure 1. Maximum likelihood phylogenetic inference of 262 SULTR sequences from 22 plant genomes. Angiosperm SULTRs form two ancient clades, AC-I and AC-II. The AC-I clade represents the SULTR4 subfamily. The AC-II clade consists of two subclades, with AC-IIA represented by the SULTR1 and SULTR2 subfamilies, and AC-IIB by SULTR3. AC-IIB is present exclusively in vascular plants with four well-supported groups organized into two branches in all angiosperm taxa. The AC-IIB-1 branch contains SULTR3.1 and SULTR3.5 subfamilies, whereas AC-IIB-2 includes SULTR3.3 and SULTR3.4 subfamilies. Bootstrap values for major nodes are shown.

Figure 2. Maximum likelihood phylogenetic inference of AC-IIB members from 20 angiosperm genomes. The 125 angiosperm SULTR3 members form two branches each with two nodes, rooted on the two *Selaginella moellendorffi* sequences. The four color-coded nodes represent SULTR3.1 and SULTR3.5 of AC-IIB-1, and SULTR3.3 and SULTR3.4 of AC-IIB-2. The perennial-specific subgroup within the SULTR3.1 subfamily is highlighted in darker shade. Bootstrap values for major nodes are shown. Taxon symbols are the same as in Figure 1.

Figure 3. Exon-intron structures of *SULTR* genes of nine representative species. Genes are arranged by phylogenetic classification, with basal lineages shown on top of each group. Conserved exon fusion events are colored and gray denotes lineage-specific fusion events. Introns are not drawn to scale and first exons are trimmed for alignments. Boldface indicates perennial-specific SULTR3.1 subfamily members.

Figure 4. Sequence alignments of *Arabidopsis thaliana* and *Populus trichocarpa* SULTRs near the predicted substrate binding pocket. Amino acid residue numbering is based on the structure-resolved AT5G13550 (AtSULTR4.1*) shown on top, and other sequences are arranged by phylogenetic grouping. Green- and purple-shaded positions denote residues critical to proton and sulfate binding or interaction, respectively. Other residues of potential relevance are shared in yellow. Boldface indicates perennial-specific SULTR3.1 subfamily members.

Figure 5. Heatmap illustration of *SULTR* gene expression across *Populus* tissues. Genes are arranged by phylogenetic grouping and genome duplicates are denoted as a and b. Boldface indicates perennial-specific SULTR3.1 subfamily members. Data were from JGI tissue atlas of Zhou et al. 2023.

Figure 6. Drought stress response of *Populus SULTR3* members. The four *SULTR3* subfamilies showed partitioned tissue response to drought stress, with divergent expression strengths between genome duplicates. Bars are means ± SE of n = 3 biological replicates. Significance testing was conducted using the two-sample *t*-test against well-watered samples (**, $P \leq 0.01$; * $P \leq 0.05$). Data were from Xue et al. 2016.

Figure 7. Xylem expression responses of *Populus SULTR3* members to genetic or environmental perturbations. **A**, Transcript abundance in normal (NW) versus tension wood (TW) from stem internodes 60-80 (Swamy et al. 2015). **B,** Transcript abundance in control versus lignin-reduced *4cl1* xylem of young stem internodes 15-30 cm from the shoot tip (Tsai et al. 2020). Bars are means ± SE of n = 3 biological replicates. Significance testing was conducted using the two-sample *t*-test (***, $P \leq 0.001$; **, $P \leq 0.01$; * $P \leq 0.05$).

Figure 8. Gene-centered coexpression analysis of xylem-biased *PtaSULT3.2a* and *PtaSULTR3.4a*. **A,** GO enrichment of 69 and 360 genes co-expressed (Pearson correlation coefficient [PCC] $\geq$ 0.7) with *PtaSULT3.2a* and *PtaSULTR3.4a*, respectively, across 179 published RNA-seq datasets from stem vascular tissues. **B,** Boxplot distribution of PCC values against 46 genes annotated and color-coded for their involvement in lignin biosynthesis, C1 metabolism, and upstream transcription factors (TFs). The gene list is provided in Supplemental Table S3.

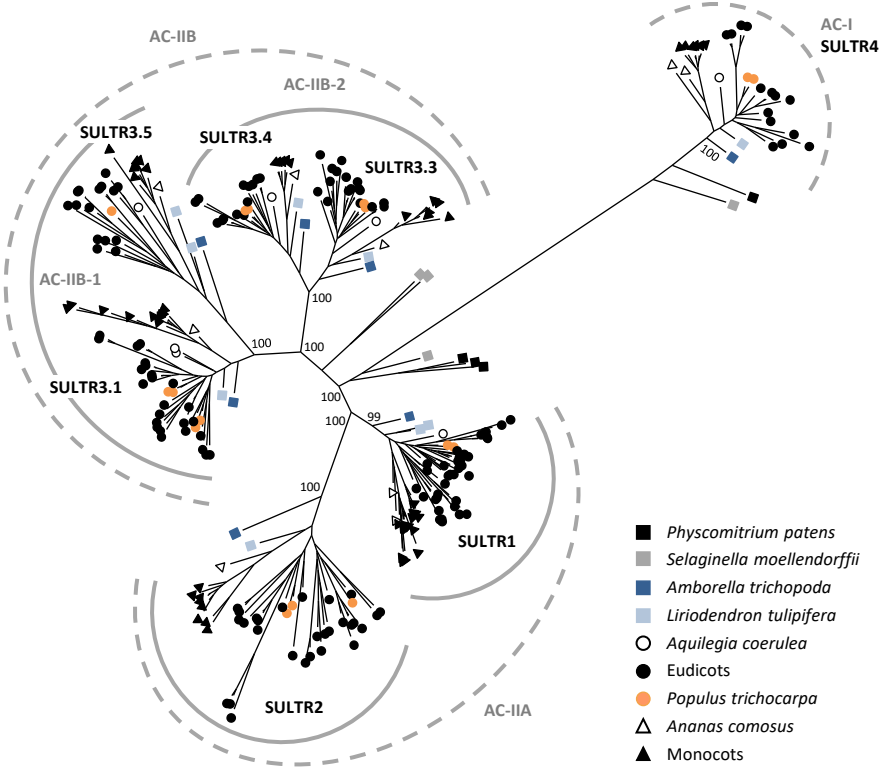
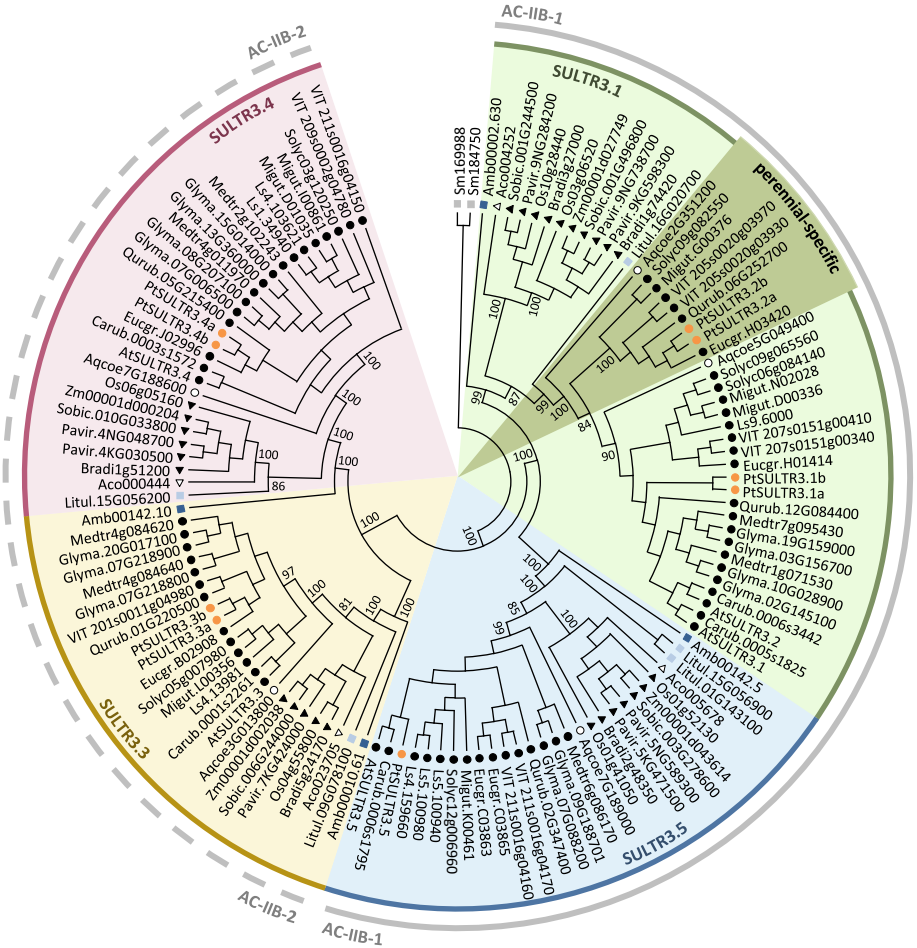


Figure 1. Maximum likelihood phylogenetic inference of 262 SULTR sequences from 22 plant genomes. Angiosperm SULTRs form two ancient clades, AC-I and AC-II. The AC-I clade represents the SULTR4 subfamily. The AC-II clade consists of two subclades, with AC-IIA represented by the SULTR1 and SULTR2 subfamilies, and AC-IIB by SULTR3. AC-IIB is present exclusively in vascular plants with four well-supported groups organized into two branches in all angiosperm taxa. The AC-IIB-1 branch contains SULTR3.1 and SULTR3.5 subfamilies, whereas AC-IIB-2 includes SULTR3.3 and SULTR3.4 subfamilies. Bootstrap values for major nodes are shown.


AC-IIB-1
AC-IIB-2
SULTR3.1
SULTR3.2
SULTR3.3
SULTR3.4
SULTR3.5
perennial-specific
Sm169988
Sm184750
Amb00002.630
Aco004252
Sobic.001G244500
Pavir.9NG284200
Os10g28440
Bradi3g27000
Os03g06520
Zm00001d027749
Sobic.001G496800
Pavir.9NG738700
Bradi1g74420
Litul.15G598300
Aqcoe2G351200
Solyc09g082550
Migut.G00376
VIT 205s0020g03970
Qurub.06G252700
PtSULTR3.2b
PtSULTR3.2a
Eucgr.H03420
Aqcoe5G049400
Solyc09g065560
Solyc06g084140
Migut.N02028
Migut.D00336
Ls9.6000
VIT 207s0151g00410
VIT 207s0151g00340
Eucgr.H01414
PtSULTR3.1b
PtSULTR3.1a
Qurub.12G084400
Medtr7g095430
Glyma.19G159000
Glyma.03G156700
Medtr1g071530
Glyma.10G028900
Glyma.02G145100
Carub.0006s3442
AtSULTR3.2
Carub.0005s1825
AtSULTR3.1
Amb00142.5
Litul.14G056900
Litul.01G143100
Aco005678
Os03g52130
Zm00001d043614
Sobic.003G278600
Pavir.5NG589300
Pavir.5NG471500
Bradi2g48350
Os03g41050
Aqcoe7G189000
Medtr6g086170
Glyma.09G188701
Glyma.07G088200
Qurub.02G347400
VIT 211s0016g04170
VIT 211s0016g04160
Eucgr.C03865
Eucgr.C03863
Migut.K00461
Solyc12g006960
Ls5.100940
Ls5.100980
Ls4.159660
AtSULTR3.5
Carub.0006s1795
PtSULTR3.5
Amb00010.161
Amb00010.5
Litul.09G078100
Aco023705
Litul.02G221170
Bradi5g24000
Os04g55800
Pavir.7NG424000
Sobic.006G244000
Zm00001d002088
Aqcoe3G013800
AtSULTR3.3
Carub.0001s2261
Ls4.139811
Migut.L00356
Eucgr.B029008
PtSULTR3.3a
PtSULTR3.3b
Solyc05g007980
VIT 201s0011g04980
Qurub.01G220500
Glyma.07G218800
Medtr4g084640
Glyma.20G017100
Glyma.07G218900
Medtr4g084620
Amb00142.10
Litul.15G056200
Aco000444
Os06g05160
Zm00001d000204
Sobic.010G033800
Pavir.4NG048700
Pavir.4KG030500
Bradi1g51200
AtSULTR3.4
Aqcoe7G188600
Eucgr.J02996
Carub.0003s1572
PtSULTR3.4b
PtSULTR3.4a
Qurub.07G096500
Glyma.08G207100
Glyma.15G014000
Glyma.13G360000
Medtr4g011970
Medtr2g102243
Ls4.103621
Migut.D01035
Migut.I00861
Solyc03g120230
VIT 209s0002g04780
VIT 211s0016g04150
100
99
87
100
100
100
84
90
100
100
100

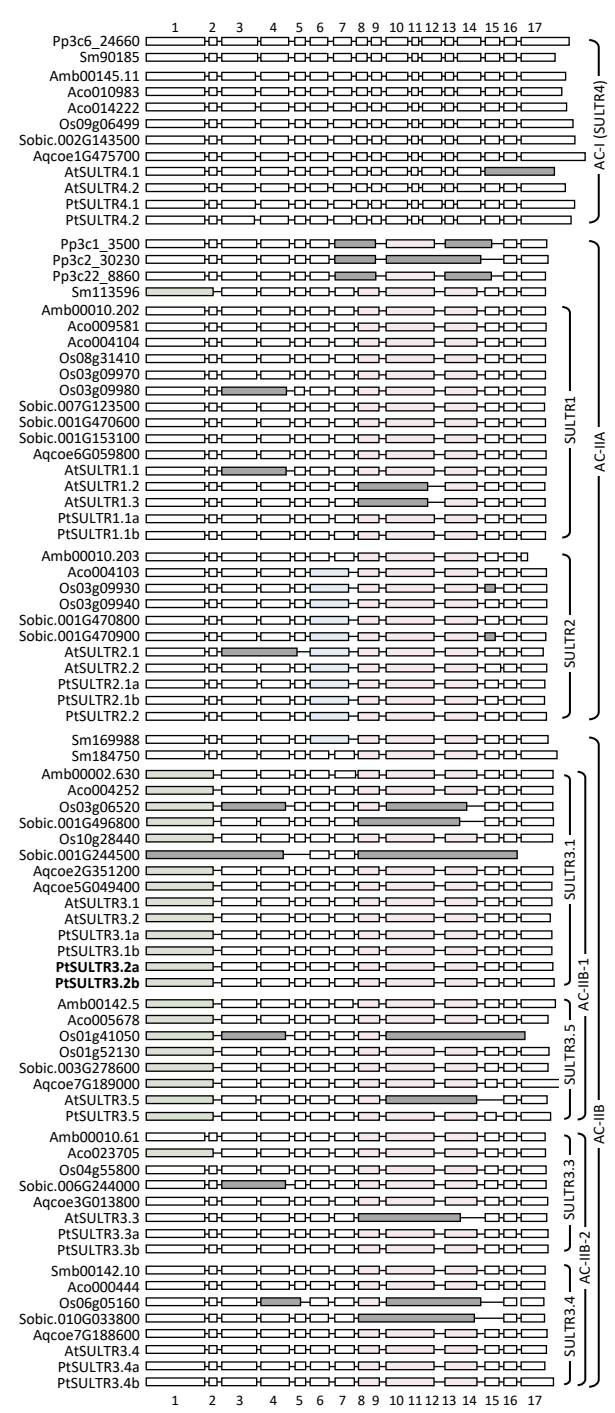


Figure 3. Exon-intron structures of *SULTR* genes of nine representative species. Genes are arranged by phylogenetic classification, with basal lineages shown on top of each group. Conserved exon fusion events are colored and gray denotes lineage-specific fusion events. Introns are not drawn to scale and first exons are trimmed for alignments. Boldface indicates perennial-specific SULTR3.1 subfamily members.

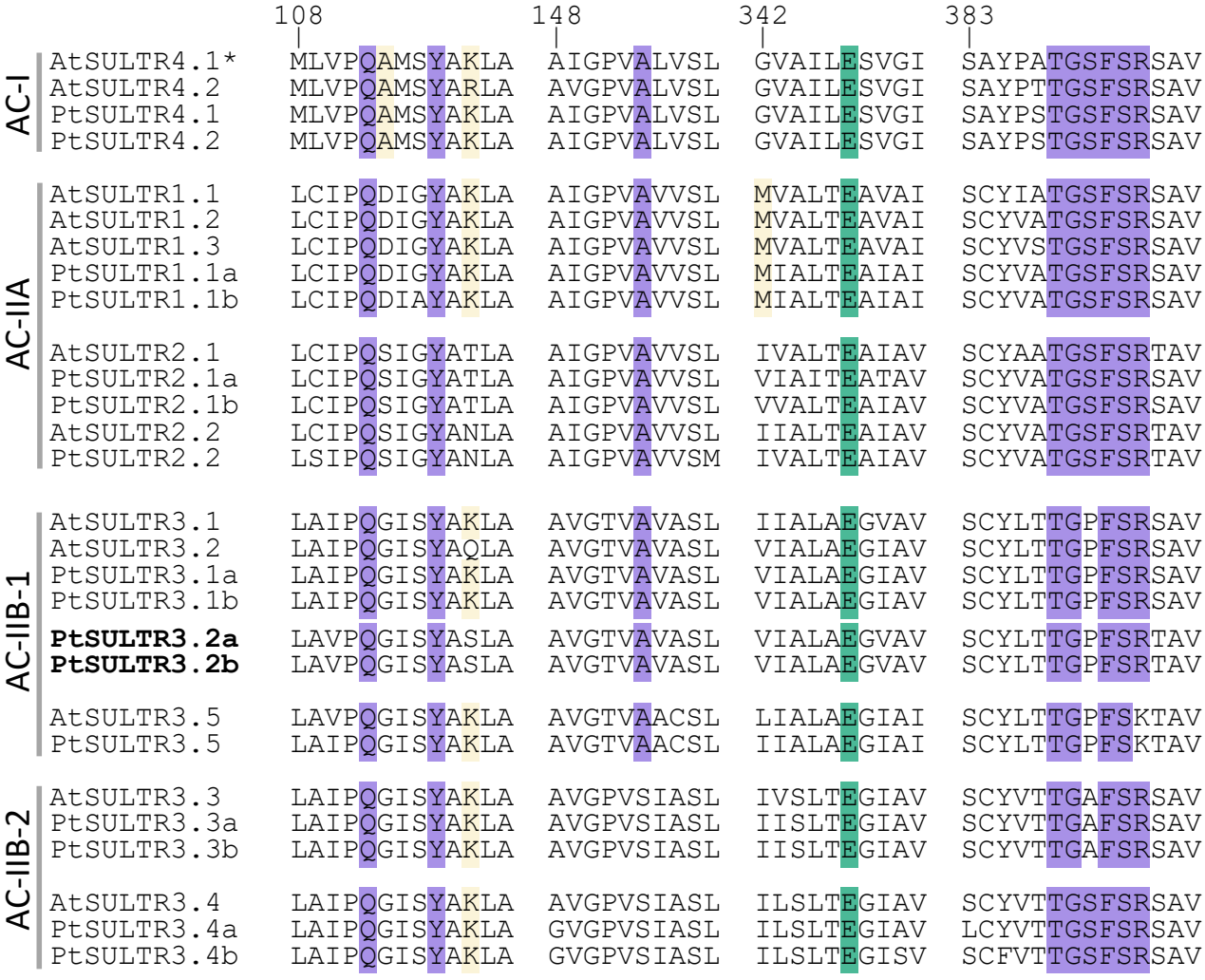


Figure 4. Sequence alignments of *Arabidopsis thaliana* and *Populus trichocarpa* SULTRs near the predicted substrate binding pocket. Amino acid residue numbering is based on the structure-resolved AT5G13550 (AtSULTR4.1*) shown on top, and other sequences are arranged by phylogenetic grouping. Green- and purple-shaded positions denote residues critical to proton and sulfate binding or interaction, respectively. Other residues of potential relevance are shared in yellow. Boldface indicates perennial-specific SULTR3.1 subfamily members.

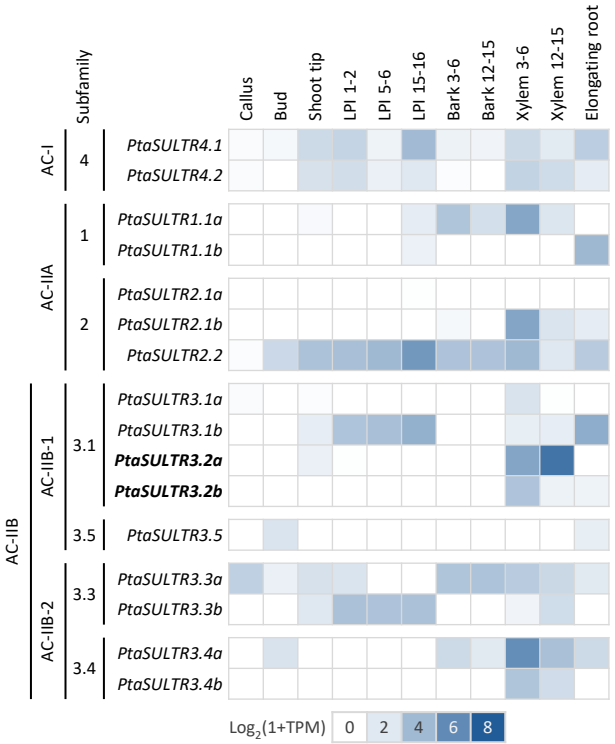


Figure 5. Heatmap illustration of *SULTR* gene expression across *Populus* tissues. Genes are arranged by phylogenetic grouping and genome duplicates are denoted as a and b. Boldface indicates perennial-specific SULTR3.1 subfamily members. Data were from JGI tissue atlas of Zhou et al. 2023.

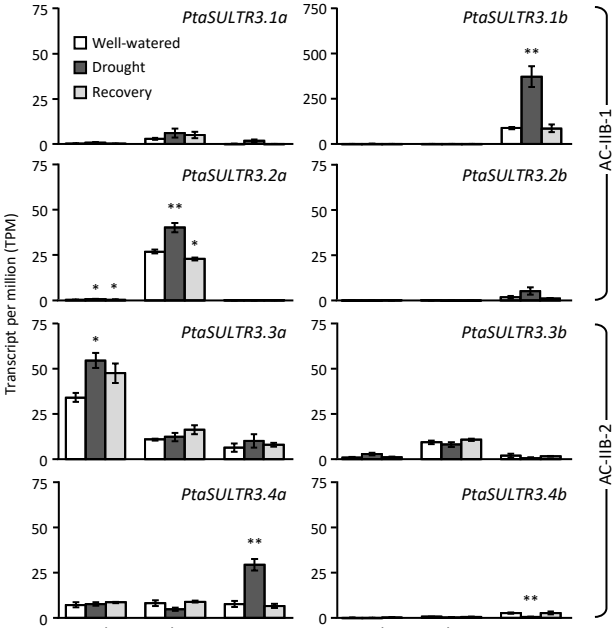


Figure 6. Drought stress response of *Populus SULTR3* members. The four *SULTR3* subfamilies showed partitioned tissue response to drought stress, with divergent expression strengths between genome duplicates. Bars are means ± SE of n = 3 biological replicates. Significance testing was conducted using the two-sample *t*-test against well-watered samples (**, $P \leq 0.01$; * $P \leq 0.05$). Data were from Xue et al. 2016.

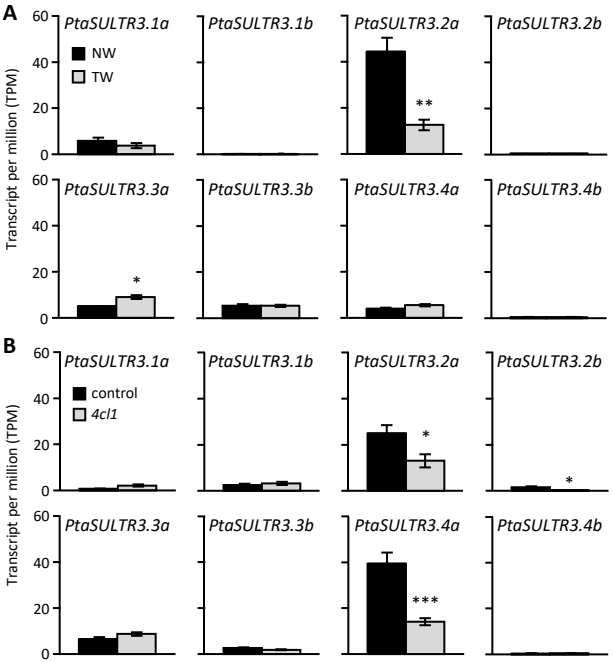


Figure 7. Xylem expression responses of *Populus SULTR3* members to genetic or environmental perturbations. **A**, Transcript abundance in normal (NW) versus tension wood (TW) from stem internodes 60-80 (Swamy et al. 2015). **B,** Transcript abundance in control versus lignin-reduced *4cl1* xylem of young stem internodes 15-30 cm from the shoot tip (Tsai et al. 2020). Bars are means ± SE of n = 3 biological replicates. Significance testing was conducted using the two-sample *t*-test (***, $P \leq 0.001$; **, $P \leq 0.01$; * $P \leq 0.05$).

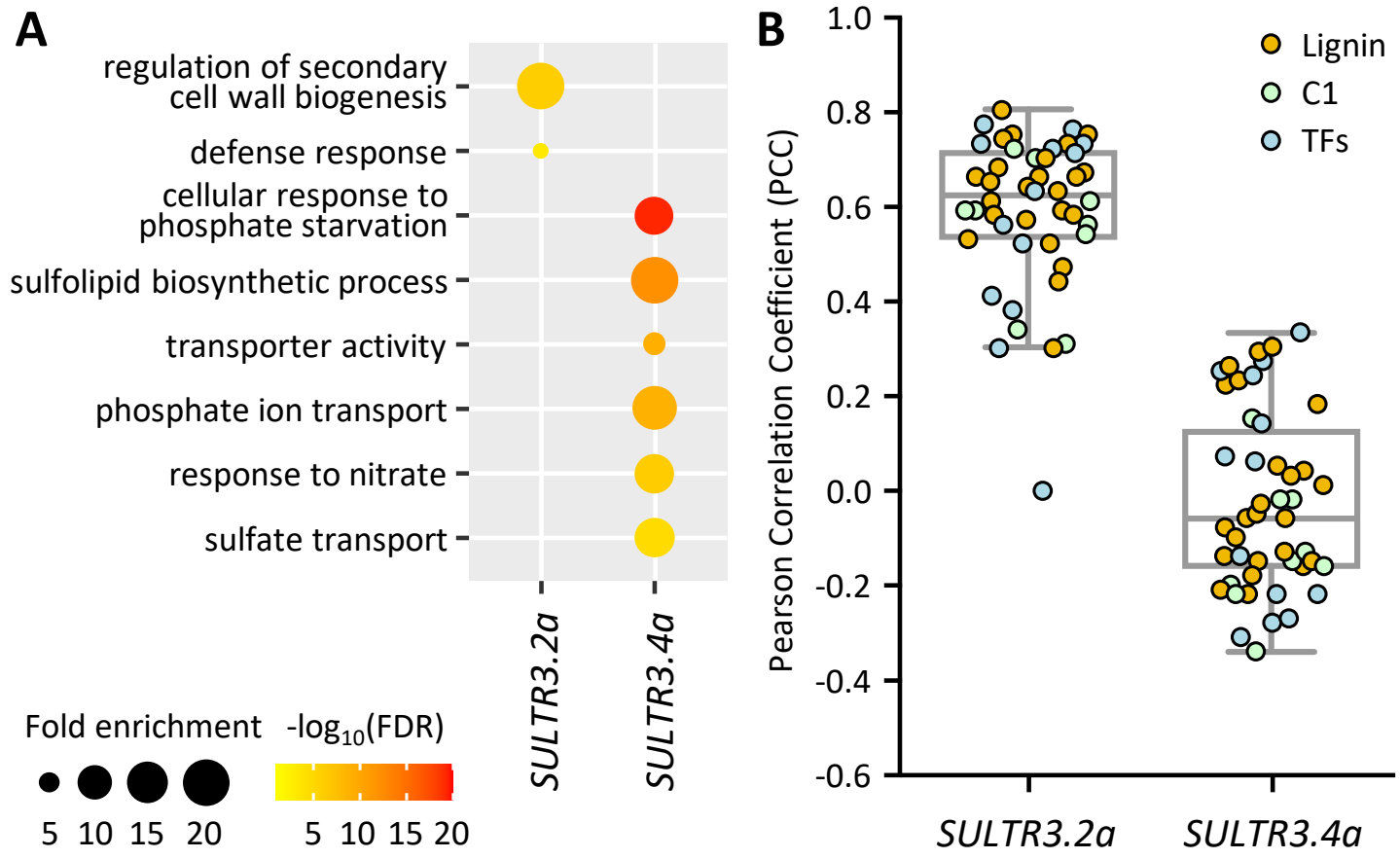


Figure 8. Gene-centered coexpression analysis of xylem-biased *PtaSULT3.2a* and *PtaSULTR3.4a*. **A,** GO enrichment of 69 and 360 genes co-expressed (Pearson correlation coefficient [PCC] ≥ 0.7) with *PtaSULT3.2a* and *PtaSULTR3.4a*, respectively, across 179 published RNA-seq datasets from stem vascular tissues. **B,** Boxplot distribution of PCC values against 46 genes annotated and color-coded for their involvement in lignin biosynthesis, C1 metabolism, and upstream transcription factors (TFs). The gene list is provided in Supplemental Table S3.

Table 1. SULTR family members in 22 plant species.

Species	AC-I	AC-IIA		AC-IIB				Total
				AC-IIB-1		AC-IIB-2		
	SULTR4	SULTR1	SULTR2	SULTR3.1	SULTR3.5	SULTR3.3	SULTR3.4	
Bryophyte								
*Physcomitrium patens**	1	3						4
Lycophyte								
*Selaginella moellendorffii**	1	1		2				4
Angiosperms								
*Amborella trichopoda**	1	1	1	1	1	1	1	7
*Liriodendron tulipifera**	1	2	1	1	2	1	1	9
Eudicots								
*Aquilegia coerulea**	1	1	0	2	1	1	1	7
Rosid								
Arabidopsis thaliana	2	3	2	2	1	1	1	12
Capsella rubella	2	3	2	2	1	1	1	12
Eucalyptus grandis	1	3	1	2	2	1	1	11
Glycine max	2	4	8	4	2	3	4	27
Medicago truncatula	1	3	4	2	1	2	2	15
Populus trichocarpa	2	2	3	4	1	2	2	17
Quercus rubra	1	4	2	2	1	1	1	12
Vitis vinifera	1	2	2	4	2	1	2	14
Asterids								
Lactuca sativa	2	4	2	1	3	1	2	15
Mimulus guttatus	1	6	2	3	1	1	2	16
Solanum lycopersicum	1	3	3	3	1	1	1	13
Monocots								
*Ananas comosus**	2	2	1	1	1	1	1	9
Brachypodium distachyon	1	3	1	2	1	1	1	10
Oryza sativa	1	3	2	2	2	1	1	12
Pancicum virgatum	2	5	2	3	2	1	2	17
Sorghum bicolor	1	3	2	2	1	1	1	11
Zea mays	1	3	1	1	1	1	1	9
Total	29	64	42	46	28	24	29	262