

1 **Identification of growth regulators using cross-species network analysis in plants**

2 Pasquale Luca Curci^{1,2,3}, Jie Zhang^{1,2+}, Niklas Mähler⁴⁺, Carolin Seyfferth^{1,2,4}, Chanaka
3 Mannapperuma⁴, Tim Diels^{1,2}, Tom Van Hautegeem^{1,2}, David Jonsen⁶, Nathaniel Street⁴, Torgeir
4 R. Hvidsten^{4,7}, Magnus Hertzberg⁶, Ove Nilsson⁵, Dirk Inze^{1,2}, Hilde Nelissen^{1,2}, Klaas
5 Vandepoele^{1,2,8*}

6 (1) Department of Plant Biotechnology and Bioinformatics, Ghent University, Technologiepark
7 71, 9052 Ghent, Belgium

8 (2) VIB Center for Plant Systems Biology, Technologiepark 71, 9052 Ghent, Belgium

9 (3) Institute of Biosciences and Bioresources, National Research Council (CNR), Via Amendola
10 165/A, 70126 Bari, Italy

11 (4) Umea Plant Science Centre (UPSC), Department of Plant Physiology, Umeå University, 90187
12 Umeå, Sweden.

13 (5) Umea Plant Science Centre (UPSC), Department of Forest Genetics and Plant Physiology,
14 Swedish University of Agricultural Sciences, 90183 Umeå, Sweden.

15 (6) SweTree Technologies AB, Skogsmarksgränd 7, SE-907 36 Umeå, Sweden

16 (7) Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life
17 Sciences, 1432 Ås, Norway

18 (8) Bioinformatics Institute Ghent, Ghent University, Technologiepark 71, 9052 Ghent, Belgium

19 + both authors contributed equally

20 * Corresponding author: Klaas.Vandepoele@psb.vib-ugent.be

21 **Short title:** Identification of growth regulators in plants

22 **Author contributions:** P.L.C., N.S., T.R.H., H.N., O.N., D.I. and K.V. designed the research;
23 P.L.C., J.Z., N.M., C.S., C.M., T.D., and T.V.H. performed research and data analysis; N.M.,
24 T.V.H., D.J., and M.H. contributed data and analytic tools, P.L.C., J.Z. and K.V. wrote the paper
25 with input from all co-authors.

26 **One-sentence summary:** Cross-species network analysis enables identification and validation of
27 growth regulators in Arabidopsis.

28 The author responsible for distribution of materials integral to the findings presented in this article
29 in accordance with the policy described in the Instructions for Authors
30 (<https://academic.oup.com/pphys/pages/General-Instructions>) is Klaas Vandepoele.

31 **Abstract**

32 With the need to increase plant productivity, one of the challenges plant scientists are facing is to
33 identify genes that play a role in beneficial plant traits. Moreover, even when such genes are found,
34 it is generally not trivial to transfer this knowledge about gene function across species to identify
35 functional orthologs. Here, we focused on the leaf to study plant growth. First, we built leaf growth
36 transcriptional networks in Arabidopsis (*Arabidopsis thaliana*), maize (*Zea mays*), and aspen
37 (*Populus tremula*). Next, known growth regulators, here defined as genes that when mutated or
38 ectopically expressed alter plant growth, together with cross-species conserved networks, were
39 used as guides to predict novel Arabidopsis growth regulators. Using an in-depth literature
40 screening, 34 out of 100 top predicted growth regulators were confirmed to affect leaf phenotype
41 when mutated or overexpressed and thus represent novel potential growth regulators. Globally,
42 these growth regulators were involved in cell cycle, plant defense responses, gibberellin, auxin,
43 and brassinosteroid signaling. Phenotypic characterization of loss-of-function lines confirmed two
44 predicted growth regulators to be involved in leaf growth (*NPF6.4* and *LATE MERISTEM*
45 *IDENTITY2*). In conclusion, the presented network approach offers an integrative cross-species
46 strategy to identify genes involved in plant growth and development.

47 **Introduction**

48 The need to increase plant productivity reveals that, despite the detailed information gained on
49 plant genomes, modelling plant growth and translating the molecular knowledge obtained in model
50 plant species to crops is not trivial (Nuccio et al., 2018; Simmons et al., 2021, Inze and Nelissen,
51 2022). Plant organ growth is one of the processes that is well-studied in model plants (Vercruyssen
52 et al., 2020a), playing a major role in affecting plant productivity (Sun et al., 2017). New plant
53 organs are formed and then grow continuously throughout development. Upon adverse conditions,

54 growth adjustments are among the first plant responses, rendering growth regulation an important
55 yield component (Gray and Brady, 2016; Nowicka, 2019). The growth of plants involves complex
56 mechanisms controlling processes from the cellular to the whole-organism level (Verbraeken et
57 al., 2021). However, which growth zones or cell types are most important in controlling organ
58 growth is not always clear.

59 Numerous genes, which we refer to as growth regulators, have been identified that when mutated
60 or ectopically expressed alter organ size, such as leaf size, in plants. Detailed transcriptome and
61 functional analyses have revealed that many of these genes are part of functional modules
62 conserved across plant species (Vercruysse et al., 2020b). Previous research has shown that largely
63 similar cellular and molecular pathways govern the fundamental growth processes in dicots and
64 monocots (Anastasiou et al., 2007; Nelissen et al., 2016). This observation is based on the presence
65 of functionally conserved orthologous growth regulators which promote organ growth in both
66 dicots and monocots. Notable examples are genes encoding CYTOCHROME P450, FAMILY 78,
67 SUBFAMILY A, POLYPEPTIDE 8 (CYP78A), AUXIN-REGULATED GENE INVOLVED IN
68 ORGAN SIZE (ARGOS), rate limiting GA biosynthesis enzymes, BRASSINOSTEROID
69 INSENSITIVE 1 (BRI1), ANGUSTIFOLIA3 and GROWTH-REGULATING FACTORS
70 (Powell and Lenhard, 2012; Vercruysse et al., 2020a).

71 The complex and highly dynamic nature of the regulatory networks controlling complex traits
72 makes the identification of growth regulatory genes challenging (Baxter, 2020). Moreover,
73 duplication events across the plant kingdom have caused a general enlargement of gene families
74 and, with it, plant- and tissue-specific functional specialization (Jones and Vandepoele, 2020). It
75 became clear that, even when the gene space is well characterized and conserved, the translation
76 from model species to crops is not straightforward (Gong et al., 2022; Inze and Nelissen, 2022).
77 One of the bottlenecks lies in the complexity of crop genomes, such as polyploidy, and the
78 subsequent difficulty in identifying functional orthologs.

79 Gene orthology information is essential to transfer functional annotations from model plants with
80 high-quality annotations (e.g. *Arabidopsis thaliana*) to other species. Functional annotations
81 derived from experimental evidence can be used to identify relevant orthologs and drive gene
82 function discovery in crops (Lee et al., 2015, 2019). This approach is not straightforward, mainly
83 for two reasons: first, the orthology approach normally leads to the identification of complex (one-

84 to-one, one-to-many and many-to-many) orthology relationships (Movahedi et al., 2011; Van Bel
85 et al., 2012); second, for genes with multiple orthologs, it has been observed that the ortholog with
86 the highest protein sequence similarity is often not the ortholog with the most similar regulation,
87 indicating that identifying functionally conserved orthologs is challenging (Patel et al., 2012;
88 Netotea et al., 2014).

89 Biological networks offer the means to study the complex organization of gene interactions.
90 Densely connected network clusters form gene modules, defined as groups of linked genes with
91 similar expression profiles (i.e. co-expressed genes), which also tend to be co-regulated and
92 functionally related (Heyndrickx and Vandepoele, 2012; Klie et al., 2012). Although transferring
93 network links from better annotated species to crops is the most intuitive approach and has proven
94 to be helpful (Ficklin and Feltus, 2011; Obertello et al., 2015), it has been shown that only ~20-
95 40% of the co-expression links are conserved in pairwise comparison of *Arabidopsis* (*Arabidopsis*
96 *thaliana*), *Populus*, and rice (*Oryza sativa*) (Netotea et al., 2014). On the other hand, it has been
97 shown that using gene modules that are conserved across species can increase the amount of
98 biological knowledge transferred from one species to another (Mutwil et al., 2011; Heyndrickx
99 and Vandepoele, 2012; Cheng et al., 2021). Such conserved gene modules mirror biological
100 processes conserved across species, meaning that the orthologous genes present in these modules
101 are involved in the same process and potentially perform the same function (Stuart et al., 2003;
102 Ruprecht et al., 2011). Significantly conserved cross-species modules (with many shared
103 orthologs) can be used to transfer gene function annotations and analyze expression conservation
104 for paralogs involved in complex many-to-many orthology relationships. A guilt-by-association
105 approach can also then be used to infer functions of unknown genes from the functions of co-
106 expressed annotated genes (Wolfe et al., 2005; Lee et al., 2010; De Smet and Marchal, 2010; Klie
107 et al., 2012; Rhee and Mutwil, 2014).

108 Here, we aimed at developing an integrative approach to identify functionally conserved
109 regulators, leveraging high-resolution transcriptomes and the power of cross-species network
110 biology. In particular, we chose leaf as a system to study plant growth, as high-quality datasets
111 covering cell proliferation and expansion are available in three plant species: two dicotyledonous
112 plants, the annual plant *Arabidopsis* and the perennial plant aspen (*Populus tremula*), and one
113 monocotyledonous plant, maize (*Zea mays*). We leveraged these data to construct aggregated gene

114 networks for each species and identified, through gene neighborhood conservation analysis, genes
115 with cross-species network conservation. Subsequently, we used known plant growth regulators,
116 belonging to various functional modules and influencing growth of different plant organs, as guide
117 genes to predict putative growth regulators among these conserved genes. For the top 100 predicted
118 growth regulators, we screened the literature to investigate if predictions linked to leaf growth
119 were obtained. For a subset of highly ranked predictions with no reported information on plant
120 growth, we performed phenotypic analyses and succeeded in validating two novel *Arabidopsis*
121 growth regulators.

122 **Results**

123 **Network construction and gene neighborhood conservation analysis**

124 To perform network construction based on gene expression information, we used transcriptomic
125 data from leaves, which were selected as a representative system to study plant growth. This choice
126 was primarily motivated by the well-known similarities in leaf growth regulation across dicots and
127 monocots, which make cross-species comparison of gene networks straightforward and useful for
128 gene function discovery (Vercruyssen et al., 2020b). Secondly, our motivation relied on the
129 availability of large-scale expression profiling studies, which allow selecting similar samples and
130 constructing a congruent dataset for the different species. Expression compendia were built for
131 *Arabidopsis*, maize and aspen that contained a minimum of 24 leaf samples (Figure 1, step 1;
132 Supplemental Table S1; Supplemental Methods). These expression compendia all include
133 developmental stages with active cell proliferation and cell expansion. The *Arabidopsis* expression
134 compendium was composed of three main developmental phases: cell proliferation, cell expansion
135 and the transition between these two phases. For maize, the developmental expression
136 compendium included a newly generated high-resolution dataset and covered cell proliferation,
137 cell expansion and mature phases of development (Supplemental Methods). For aspen, samples
138 covered the developmental stages ranging from the very youngest leaf primordia to fully expanded
139 and mature leaves. In total, expression data covered 20,313 genes for *Arabidopsis*, 29,383 genes
140 for maize, and 35,309 genes for aspen (Supplemental Dataset S1).

141 The network construction was performed for each species with Seidr, a toolkit to perform multiple
142 gene network inferences and combine their results into a unified meta-network (Schiffthaler et al.,

143 2018). For each network inference algorithm included, a fully connected weighted gene network
144 was constructed. These were in turn aggregated into a weighted meta-network (simply “network”
145 hereinafter, Figure 1, step 2). When applying a weight threshold, the network density was defined
146 as the ratio between the number of links with a weight higher than this threshold and the number
147 of links in the weighted network. To dissect the network structure, several thresholds were used to
148 subset the networks into more stringent density subnetworks (DSs). For each species network, five
149 DSs were obtained ranging from DS1 (top 0.1% links) with an average of 358,455 links, to DS5
150 (top 10% links) with an average of 35,845,512 links (Figure 1, step 3), with higher densities
151 corresponding to a higher number of neighbors for each gene in the network (Supplemental Figure
152 S1). A gene’s neighborhood is defined as all genes connected with this gene for a given network.

153 Genes showing gene neighborhood conservation across species are part of conserved functional
154 modules controlling distinct biological processes. This implies that the conserved network
155 containing these genes confers a selective advantage and therefore that these genes are functionally
156 related (Stuart et al., 2003). However, which gene neighborhood size to select to identify conserved
157 growth-related functional modules is not straightforward, as being too stringent might lead to the
158 loss of valuable interactions while being too relaxed might include non-functional interactions
159 potentially representing noise (Movahedi et al. 2012). To identify genes showing network
160 conservation in different species, a gene neighborhood conservation analysis was performed using
161 each DS and the information on the orthology relationships between Arabidopsis, maize and aspen
162 genes (Figure 1, step 4a). The network neighborhood of a gene is represented by all genes
163 connected to it, at a given threshold. This concept was used to identify “triplets” (Supplemental
164 Dataset S2), each containing three orthologous genes across Arabidopsis, maize and aspen with
165 statistically significant overlaps between their gene network neighborhoods (see Methods). In an
166 example triplet (Figure 1, step 4a), a specific Arabidopsis gene *AI*, will have an ortholog *ZI* in
167 maize and another ortholog *PI* in aspen and these three genes will have a significant overlap of
168 their gene network neighborhoods. Due to the complex orthology relationships that exist in plants,
169 each gene can belong to one or multiple triplets as it can have one or more orthologs. For example,
170 an Arabidopsis gene with only one ortholog in maize and aspen, assuming they have significant
171 overlap of their gene network neighborhoods, will belong to one triplet. In contrast, another
172 Arabidopsis gene with two orthologs in maize and three in aspen, assuming they also all have
173 significant overlaps of their gene network neighborhoods, will belong to six triplets. We refer to

174 the set of unique genes that are part of triplets as “triplet genes”. Next, the conserved gene
175 neighborhoods were used to dissect the complex network structures of these plants and to
176 functionally harness the orthology relationships. The cross-species networks are available in an
177 interactive web application (<https://beta-complex.plantgenie.org>).

178 **Delineation of conserved growth regulators**

179 Since the output of cell proliferation and expansion are strongly contributing to leaf size, we
180 hypothesized that the generated triplets were an excellent source to extract orthologs potentially
181 altering plant growth, representing conserved GRs. Growth regulators typically act by stimulating
182 cell proliferation (yielding a higher cell number, as in the case of GRF (GROWTH-
183 REGULATING FACTOR) and GIF (GRF-INTERACTING FACTOR) proteins (Lee et al., 2009))
184 and/or cell expansion (as in the case of *ZHD5* (*ZINC-FINGER HOMEODOMAIN 5*) (Hong et al.,
185 2011)). We generated a list of known GRs (“primary-GRs”) covering 71 primary-GRs from
186 Arabidopsis, 71 from aspen and eight from maize. While the Arabidopsis and maize GRs mainly
187 have a role in controlling leaf size, the aspen GRs are affecting stem size. In both organs, cell
188 proliferation and expansion play an important role in controlling growth (Serrano-Mislata and
189 Sablowski, 2018). This list of genes was obtained by collecting scientific literature and by
190 phenotypic analysis of mutant and over-expression lines in Arabidopsis, maize, and aspen. We
191 then used the triplets to transfer GRs from maize and aspen to Arabidopsis (“translated-GRs”). In
192 other words, primary-GRs from maize and aspen, also identified as triplet genes, were used to
193 extract Arabidopsis orthologs with gene neighborhood conservation. The primary-GRs and
194 translated-GRs were finally merged and filtered for high expression variation in the Arabidopsis
195 expression compendium to retain only those active during either cell proliferation or cell
196 expansion. The resulting set, named “expression-supported GRs” (Supplemental Table S2,
197 Supplemental Figure S2), was composed of 82 GRs, including 24 Arabidopsis primary-GRs and
198 58 translated-GRs (*GRF2* and *GA20OX1* (*GIBBERELLIN 20-OXIDASE 1*) were shared between
199 primary-GR and translated-GR sets). According to their expression profiles in Arabidopsis, 35
200 expression-supported GRs showed maximal expression during cell proliferation, including several
201 proliferation marker genes like GROWTH-REGULATING FACTORS (e.g. *GRF1*, *GRF2*, *GRF3*),
202 AINTEGUMENTA (*ANT* (Mizukami and Fischer, 2000) and *KLUH* (Anastasiou et al., 2007)),
203 and 47 expression-supported GRs had increased expression during cell expansion, such as

204 *GA20Ox1* (Barboza et al., 2013) and *BR ENHANCED EXPRESSION 2* (*BEE2* (Friedrichsen et al.,
205 2002)).

206 The 82 expression-supported GRs (from here on simply referred to as “GRs”) represent our guide
207 genes, obtained by the integration of prior knowledge on plant growth and the cross-species gene
208 neighborhood conservation approach, to identify candidate GRs.

209 **Functional analysis of cross-species conserved networks underlying leaf cell proliferation** 210 **and expansion**

211 To explore cross-species conserved genes that function during cell proliferation and expansion,
212 we performed a Gene Ontology (GO (Ashburner et al., 2000)) functional enrichment analysis of
213 the Arabidopsis triplet genes from each DS across two sets: (1) all triplet genes (All) and (2) the
214 subset of triplet genes including the 82 GRs and their co-expressed triplet genes (Growth regulator-
215 related triplet genes) (Figure 2). The total number of triplets ranged from 1,739 (DS1) to 243,645
216 (DS5) (Figure 2A; Supplemental Dataset S2). To assess the significance of these numbers, a
217 permutation approach was employed where the orthology relationships were randomized 500
218 times and the number of triplets obtained from each permutation was recorded. The number of
219 triplets observed were highly significant with not a single permutation for any DS exceeding the
220 number of triplets observed in the non-permuted data (p-value<0.002). The number of unique
221 Arabidopsis triplet genes ranged from 211 (DS1) to 6,526 (DS5) indicating that less sparse
222 networks tend to have more genes and more conserved gene neighborhoods (Figure 2A).
223 Interestingly, GRs and their network neighbors on average made up 71% of the triplet genes across
224 the five DSs, suggesting that leaf growth-related gene networks are well conserved during leaf
225 development across plant species. For simplicity, from here on we will refer to triplet genes at a
226 specific DS as, for example at DS1, “genes conserved at DS1”. The functional enrichment (Figure
227 2B) showed that triplet genes from the most stringent subnetwork (DS1) were enriched for basal
228 biological processes during leaf development, including photosynthesis (e.g. glucose metabolic
229 process, response to light and carbon fixation) and translation (e.g. large and small ribosomal
230 subunits). Processes such as cell division and cell cycle regulation were significantly enriched for
231 genes conserved at DS2 and DS3, including genes coding for cyclins (type A, B, D and P), cyclin
232 dependent kinases (*CDK*) and their subunits (*CKS*), and other genes involved in the spindle
233 formation (i.e. *MICROTUBULE-ASSOCIATED PROTEINS (MAP)65-4 and -5*). Cell expansion-

234 related processes were identified among genes conserved at DS3 and included genes coding for
235 expansins (EXP) and xyloglucan endotransglucosylases/hydrolases (XTH). Genes conserved at
236 the two least stringent subnetworks (DS4 and DS5) were enriched for GO terms related to cell wall
237 organization (e.g. lignan biosynthesis, pectin degradation, lignin metabolism), defense response to
238 biotic and abiotic stresses (e.g. defense response to oomycetes, response to salt stress and heat
239 stress), and transmembrane transport and hormone signaling (e.g. response to auxin, ethylene and
240 brassinosteroid). The category “regulation of transcription” was enriched for genes conserved at
241 DS3, DS4, and DS5. GRs were significantly over-represented in subnetworks starting from DS2,
242 indicating that GRs have highly conserved gene network neighborhoods. Most of the GRs (87%)
243 were conserved in one or more DSs (Figure 2C).

244 Among the GRs conserved at DS2, 32% were transcription factors (TFs), including regulators of
245 cell cycle (e.g. *AINTEGUMENTA*) and cell elongation such as *BEE2* and its homolog *HBII*
246 (Supplemental Figure S3). These results suggest a conserved role of these TFs in leaf development
247 across the three plant species. Genes involved in hormone-mediated transcriptional regulation
248 (*INDOLEACETIC ACID-INDUCED PROTEIN (IAA)3*, *IAA14*, *IAA30*, and *AUXIN RESISTANT*
249 (*AUX1*) were also detected. Cell growth regulators, including the GRF family, were found
250 conserved and, among them, *GRF2* was conserved at DS2. Literature information on differentially
251 expressed gene (DEG) sets from perturbation experiments was also included in the functional
252 enrichment analyses for several primary-GRs. In particular, genes up- and down-regulated in
253 *SAMBA* loss-of-function mutants (Eloy et al., 2012) and *JAW (JAGGED AND WAVY)*
254 overexpression lines (Gonzalez et al., 2010) were significantly enriched in the GR-related set
255 (Figure 2B). Whereas *SAMBA* plays a key role in organ size control (seeds, leaves and roots),
256 transgenic overexpression lines of *JAW* showed enlarged leaves and an increased cell number,
257 indicative of prolonged cell proliferation (Gonzalez et al., 2010; Eloy et al., 2012). An additional
258 functional enrichment analysis was performed focusing on TF families to identify their cross-
259 species conservation level. In particular, genes conserved from DS2 to DS5 (Supplemental Figure
260 S4) were significantly enriched for the ETHYLENE RESPONSE FACTOR (ERF) family (q-value
261 < 0.01), which has a recognized role in plant growth (Dubois et al., 2018). At DS3, among others,
262 MYB and WRKY TF families, known to be involved in developmental processes, appeared
263 strongly conserved. At the least stringent DSs (DS4 and DS5) we could observe other conserved
264 TF families like DOF (regulating the transcriptional machinery in plant cells), MIKC-MADS

265 (involved in floral development) and NAC (with functions in plant growth, development and stress
266 responses) (Lehti-Shiu et al., 2017). For TFs conserved at DS2, a significant enrichment was
267 observed for the CONSTANS-like TF-family when considering GR-related triplet genes and
268 included *BBX3*, *BBX4*, *BBX14* and *BBX16*. A number of BBX proteins have been linked with
269 photomorphogenesis, neighborhood detection, and photoperiodic regulation of flowering (Vaishak
270 et al., 2019).

271 **Network-based prediction of novel growth regulators**

272 Apart from analyzing the conservation level of known GRs, we subsequently investigated if new
273 GRs could be identified. To obtain high-quality GR predictions, a combined strategy was adopted
274 to leverage the known GRs and the gene neighborhood conservation analysis through a guilt-by-
275 association (GBA) approach. The GBA principle states that genes with related function tend to be
276 protein interaction partners or share features such as expression patterns or close network
277 neighborhood (Oliver Stephen, 2000). First, gene function prediction through GBA was
278 performed, where the known GRs were used as guide genes for network-based gene function
279 discovery (Figure 1, step 4b). Gene functions were assigned through functional enrichment in the
280 Arabidopsis networks, at different DSs. As a result, genes that were part of network neighborhoods
281 significantly enriched for guide GRs were classified as predicted GRs, and a GBA score was
282 assigned to quantify the strength of the predicted GRs (see Materials and Methods). Secondly, the
283 predictions (Figure 1, step 4b) were filtered for those already identified as triplet genes (Figure 1,
284 step 4a). These filtered predictions (Figure 1, step 5), forming the predicted GR set, were labelled
285 with their species names if they were part of the guide GRs (primary or translated-GR) or with
286 “new” if they were novel (Supplemental Table S3). This approach led to 2206 GR predictions, of
287 which 66 were guide GRs. For the latter, 11 were uniquely from the Arabidopsis GR primary set,
288 53 uniquely from the aspen translated-GRs, and the remaining two were shared among species.
289 Note that the recovery of known GR genes would be zero in case the network would be random
290 and not capture growth-related transcriptional information. From DS1 to DS5, the subsets of GR
291 predictions covered 175, 496, 421, 891 and 223 genes, respectively (Supplemental Table S3).
292 Overall, the biological processes observed for the conserved predictions agreed with those
293 observed for all triplet genes (Figure 2).

294 To evaluate the reliability of the predicted GR set and its potential use for discovering genes with
295 a significant effect on plant growth, the public phenotype database RARGE II (Akiyama et al.,
296 2014), covering 17,808 genes and 35,594 lines, was screened obtaining a list of 391 Arabidopsis
297 genes that, if mutated, caused a phenotype change in Arabidopsis leaf length, width and/or size
298 (RARGE II leaf trait genes, Supplemental Table S4). When investigating the gene recovery for the
299 RARGE II leaf trait genes (Figure 3), a clear trend was observed in phenotype recovery ranging
300 from DS1, with higher recovery (~3 and ~4.3 fold enrichment compared to what is expected by
301 chance for proliferation and expansion, respectively), to DS5, with almost no recovery. This result
302 indicates that, among all DSs, DS5 is the least suitable one to identify genes with a potential effect
303 on leaf phenotype.

304 **Validation of GR predictions using literature and leaf phenotyping**

305 To validate the assumption that the GR predictions top ranked by GBA are more likely to show a
306 plant growth-related phenotype, an in-depth literature analysis was performed to summarize the
307 connection with different growth-related pathways (Supplemental Table S5) and to score known
308 growth-related phenotypes for the top 100 GR predictions (Supplemental Table S6). For 61 of
309 these 100 predicted genes, mutant lines and/or lines with ectopic expression were reported. For 34
310 out of the 61 genes (55.7%), obvious alterations to leaf size and shape as well as petiole length
311 were reported when mutated or overexpressed (Supplemental Table S6).

312 Functional analysis of the 34 genes with described leaf phenotypes revealed their involvement in
313 several biological processes and pathways such as cell cycle regulation, hormone response,
314 photosynthesis, carbon utilization and cell wall modification (Figure 4). Importantly, we could
315 find conserved relationships between five specific genes active in the expansion phase:
316 CATIONIC AMINO ACID TRANSPORTER (*CAT*)2, *THIOREDOXIN X (THX)*, BETA
317 CARBONIC ANHYDRASE (*BCA*)4, *CA2*, and *PMDH2*. Among them, *CAT2* and *BCA4* were
318 also high ranked by GBA score. For the proliferation cluster, we could observe strong relationships
319 between *ANT*, *OBFBINDING PROTEIN 1 (OBP1)*, *GRF2*, *CYCD3;3*, *GLABRA 1 (GLI)*, *HTA8*
320 (*HISTONE H2A 8*), and *AN3*. Among them, we identified TFs mainly involved in cell cycle
321 process (*ANT*, *OBP1*, *GRF2*), cell wall (*GLI*), and hormone signaling pathways such as jasmonate
322 (*GLI*), abscisic acid (*ANT*), and gibberellin (*GLI*). Twenty-seven of the 61 predictions with knock-
323 down mutations and/or ectopic expression lines did not show a association with leaf growth, which

324 may be partially due to the redundancy of large gene families or that the leaf phenotype was not
325 explored in those studies. Additionally, three of these 27 genes have been reported to influence
326 root or hypocotyl development, which may also contribute to overall plant growth and organ size.

327 To further validate the role of these candidate GRs in the leaf development, the system that we
328 chose to study plant growth, we collected the mutants of nine genes among the 27 predicted GRs
329 which have not been reported with a leaf phenotype (Supplemental Table S7). Molecular
330 identification of these mutants was conducted and a detailed analysis of leaf growth in controlled
331 long-day soil-grown conditions was made (Supplemental Figure S5). By following the projected
332 rosette area (PRA), compactness and stockiness of each mutant line over time, this phenotypic
333 characterization revealed that the mutants of two GR candidate genes showed altered rosette
334 growth. The mutant lines of a putative nitrate transporter gene *NPF6.4/NRT1.3*, *sper3-1* and *sper3-*
335 *3*, both displayed decreased PRA compared with the wild-type plants (Figure 5A). The *sper3-1*
336 harbored a mutation at a conserved glutamate of *NRT1.3*, while the T-DNA line *sper3-3* was a
337 knockout allele (Tong et al., 2016). The reduction in size of *sper3-3* was smaller and occurred later
338 in development compared with *sper3-1*. Before bolting (26 DAS), *sper3-1* and *sper3-3* were 37.3%
339 and 13.2% smaller, respectively, compared with the wild-type (Supplemental Table S7). Both
340 *sper3-1* and *sper3-3* showed significantly reduced leaf number compared to wild type (Figure 5,
341 Supplemental Figure S6). Besides *NPF6.4*, the mutants of *LATE MERISTEM IDENTITY2 (LMI2)*
342 which has been reported to be required for correct timing of the meristem identity transition
343 (Pastore et al., 2011), also showed altered rosette growth. In standard long-day conditions in soil,
344 a significant reduction of PRA was detected in *lmi2-1*, which displayed elevated *LMI2* expression
345 in seedlings. By contrast, the *lmi2-2* mutants in which the T-DNA insertion gave rise to a truncated
346 non-functional *LMI2* protein, exhibited significantly increased PRA and were 13.5% larger than
347 the wild-type plants at 26 DAS (Figure 5B and Supplemental Table S7). Among *LMI2* mutants,
348 *lmi2-2* showed significantly increased leaf number (Figure 5, Supplemental Figure S6). Both
349 *NPF6.4* and *LMI2* were highly ranked by GBA (rank 18 and 20, respectively), which further
350 implies that the predictions with a low GBA score are more likely to show a leaf phenotype.

351 Although the leaf was the model system chosen and analyzed in this study, we do not exclude that
352 the predicted candidate GRs, including the validated *NPF6.4* and *LMI2*, might also alter the growth
353 of other organs. Taken together, these experimentally validated genes lend additional support to
354 the potential of our predictions for plant growth regulation.

355 **Discussion**

356 In this study, we developed an integrative approach to identify candidate genes responsible for
357 altering plant growth. To accomplish this, we used cross-species gene network analysis focusing
358 on the leaf, given its similarities between dicots and monocots (Nelissen et al., 2016). To identify
359 relevant context-specific gene interactions, it is highly recommended to focus the gene network
360 analysis on a specific condition or context, rather than integrating multiple conditions (e.g.
361 different stresses, growth conditions, development stages) (Pavlidis and Gillis, 2012; Liseron-
362 Monfils and Ware, 2015; Serin et al., 2016). For this reason, expression datasets were generated
363 and compiled capturing two main features of leaf growth: cell proliferation and cell expansion.
364 These two processes are governed by similar cellular and molecular pathways across monocots
365 and dicots (Nelissen et al., 2016), which inspired the selection of transcriptional datasets from two
366 dicots (*Arabidopsis* and aspen) and one monocot (maize). The network construction was carried
367 out integrating multiple inference methods to leverage the power and complementarity of different
368 network inference algorithms (Marbach et al., 2012; Schiffthaler et al., 2018). To evaluate the
369 strength of different biological signals in our network, the gene interactions, obtained after
370 applying different network density cutoffs (DS1-5), were studied. Given that thousands of genes
371 are expressed during leaf development, prioritizing candidate growth regulators starting from
372 different developmental expression datasets is a major challenge. To do so, we relied on two main
373 approaches: the guilt-by-association principle, which is frequently used for gene discovery, and
374 network neighborhood conservation analysis, which detects significantly overlapping network
375 neighborhoods across species to identify reliable functional orthologs (Movahedi et al., 2011;
376 Netotea et al., 2014).

377 From the gene neighborhood conservation analysis on five different density subnetworks, we
378 observed that, with an increased network density, the number of genes with conserved network
379 neighborhood also grew. This is expected and is probably due to a greater statistical power when
380 comparing larger neighborhoods (Netotea et al., 2014). Overall, as previously observed
381 (Vercruysse et al., 2020b), the integration of different sequence-based orthology detection methods
382 was important because of their complementarity, highlighting complex orthology relationships and
383 evaluating the strength of the orthology support. Overall, 36% of the *Arabidopsis* genes (7,320 out
384 of 20,313 genes present in the network) had conserved neighborhoods across *Arabidopsis*, aspen,

385 and maize, in any of the five density subnetworks. This result is similar to what has been found
386 across Arabidopsis, poplar and rice, although a different network construction pipeline was used
387 there (Netotea et al., 2014).

388 From a plant breeding perspective, we were interested in cross-species functionally conserved
389 predictions with experimental evidence in more than one species. *GA20-oxidase1* represents a
390 well-known example of a GR that is functionally conserved across monocots and dicots. This gene
391 was confirmed in our analyses to be conserved at the network neighborhood level. *GA20-oxidase1*
392 is in fact a rate limiting enzyme for gibberellin growth hormone biosynthesis in Arabidopsis,
393 aspen, maize and rice (Gonzalez et al., 2010; Nelissen et al., 2012; Qin et al., 2013; Eriksson et
394 al., 2000). To validate the functional relevance of the predicted GRs, we screened the top 100 GR
395 predictions and observed that, among the 34 Arabidopsis predicted genes with a known leaf
396 phenotype in Arabidopsis, six were also already known to affect plant growth in aspen (here stem
397 size). This result is not unexpected as overlapping regulatory mechanisms and genes are shared
398 between primary and secondary meristems, which are responsible for the formation of plant tissues
399 and organs (Baucher et al., 2007). The six translated-GRs were *AUX1*, *IAA3/SHY2*, *AUXIN*
400 *RESISTANT 5 (AXR5)*, *ATBS1 INTERACTING FACTOR 3 (AIF3)*, *AIF4*, and *HOMOLOG OF*
401 *BEE2 INTERACTING WITH IBH 1 (HB11)* and their expression in Arabidopsis was peaking at the
402 cell expansion phase. The first three genes are auxin-related genes. Auxin is important for
403 regulating root meristem growth and is crucial for root initiation and lateral root number. *AUX1*
404 was translated from aspen Potra002054g16021 while *IAA3/SHY2* and *AXR5* were translated from
405 aspen Potra000605g04596. For both these aspen genes, generated aspen RNAi lines exhibited an
406 increase in stem size, an important indicator for tree biomass yield, connecting back to the
407 underlying regulatory processes in the meristematic tissues (Supplemental Table S2). *AUX1* is an
408 auxin transport protein which regulates auxin distribution across source (young leaf) and sink
409 organs (young roots) (Marchant et al., 2002). *IAA3/SHY2* is crucial for root meristem development
410 in Arabidopsis, being the converging point of cytokinin and auxin regulatory circuit (Li et al.,
411 2020). Arabidopsis mutants for *AUX1* and *IAA3/SHY2* showed alterations in number and size of
412 lateral roots (Tian and Reed, 1999; Marchant et al., 2002) while *AXR5* is an auxin response factor
413 and mutant plants for this gene are tolerant to auxin and show alterations of root and shoot tropisms
414 (Yang et al., 2004). Our network results and phenotypes in aspen and Arabidopsis indicate that
415 these genes also play an important role in meristem growth in other organs apart from root. *HB11*,

416 *AIF3*, and *AIF4*, encode a tier of interacting bHLH transcription factors downstream of BR and
417 regulate the cell elongation in leaf blade and petiole (Bai et al., 2013; Ikeda et al., 2013). *AIF3* and
418 *AIF4* were translated from Potra004144g24626 while *HBII* was translated from
419 Potra186144g28414. These two aspen genes have been tested with an overexpression approach in
420 aspen trees showing even a bigger increase in stem size as compared with the auxin-related aspen
421 genes Potra000605g04596 and Potra002054g16021 (Supplemental Table S2). Arabidopsis
422 mutants for these genes (*HBII*, *AIF3*, and *AIF4*) have been linked with alteration of petiole length
423 (Supplemental Table S6).

424 *LMI2* was a highly ranked GR prediction. Importantly, *LMI2* (a MYB TF) is not a paralog of *LATE*
425 *MERISTEM IDENTITY 1* (*LMII*, a homeobox TF), also predicted here. Although *LMII* and *LMI2*
426 belong to different TF families, they both function downstream of LEAFY to regulate meristem
427 transition (Pastore et al., 2011). *LMII* was reported to regulate leaf growth in Arabidopsis and
428 other species (Vlad et al., 2014; Andres et al., 2017; Li et al., 2021). Arabidopsis *LMII* loss-of-
429 function mutant showed decreased leaf serration and promoted tissue growth in stipules (Vuolo et
430 al., 2018). The observed phenotype of mutated *LMI2* was related to an increase of the number of
431 cauline leaves and secondary inflorescences (Pastore et al., 2011). Here, *LMI2* transgenic lines
432 were subjected to phenotypic analysis, which demonstrated that a *LMI2* loss-of-function mutant
433 showed increased leaf number and rosette area. We do not exclude that other organs and/or traits
434 might also be affected by the loss of functionality of this gene. The neighborhood conservation of
435 both *LMII* and *LMI2* suggests that it would be worthwhile to further explore their roles in leaf
436 shape control across monocots and dicots.

437 Other known examples of functionally conserved predictions across monocots and dicots were
438 GRFs (e.g. the highly ranked *GRF2*), which have a recognized role in leaf size regulation, and
439 AN3/GIF1, a transcriptional co-activator protein (Nelissen et al., 2016). This was also testified by
440 their network conservation in stringent density subnetworks (DS2). A second gene, *GLI*, had its
441 network neighborhood conserved with GRMZM2G022686 from maize. This maize gene encodes
442 for the MYB-related protein *Myb4*. This protein plays important roles in plant improved tolerance
443 to cold and freezing in Arabidopsis and barley (Soltész et al., 2012), but no connections with
444 growth have been observed for this gene. Arabidopsis *SUC2* showed conservation with
445 GRMZM2G307561, a sucrose/H⁺ symporter which remobilize sucrose out of the vacuole to the

446 growing tissues. Mutants for this gene showed reduced growth and the accumulation of large
447 quantities of sugar and starch in vegetative tissues in Arabidopsis (Srivastava et al., 2008), while
448 in maize mutants, slower growth, smaller tassels and ears, and fewer kernels were observed (Leach
449 et al., 2017). This gene is thus also important for growth, development, and yield across monocots
450 and dicots.

451 The application of a cross-species approach is an important feature of our methodology. To
452 perform GR predictions, translated-GRs from aspen and maize were also used as guide genes,
453 together with triplets to focus on the conserved parts of the inferred leaf networks. As a result,
454 among the cross-species conserved predictions with experimental evidence in more than one
455 species described above, *AUX1*, *IAA3/SHY2*, *AXR5*, *AIF3*, *AIF4*, *HBI1*, *AN3/GIF1*, *GL1*, and
456 *SUC2* couldn't have been predicted using solely primary-GRs from Arabidopsis. This observation
457 indicates that the integration of information of different plant species enhances the detection of
458 GRs.

459 A total of 11 primary-GRs from Arabidopsis showed no network neighborhood conservation. Lack
460 of conservation might be the result of (1) missing orthologs in a target species or (2) different
461 network gene neighbors across species, which in turn might be caused by different transcriptional
462 control. One clear example of no conservation due to a lack of orthologs is PEAPOD 2 (*PPD2*),
463 which is a TIFY transcriptional regulator part of the PEAPOD (PPD) pathway. This pathway plays
464 an important role in cell proliferation and, with its PPD/KIX/SAP module, is involved in leaf,
465 flower, fruit, and seed development. This pathway is present in most vascular plant lineages, but
466 was lost in monocot grasses (Schneider et al., 2021). The reason for this absence might be found
467 back in intrinsic differences between eudicots and grasses, being mainly lack of meristemoids and
468 functional redundancy for the regulation of cell proliferation. Surprisingly, several non-grass
469 monocot species such as banana (*Musa acuminata*) and oil palm (*Elaeis guineensis*), the
470 angiosperm *Amborella trichopoda* and lycophytes, carry PPD/KIX/SAP orthologs, although
471 information about their functionality is missing (Schneider et al., 2021). Another gene with
472 orthologs but lacking network neighborhood conservation was *AHK3*, a cytokinin receptor that
473 controls cytokinin-mediated leaf longevity. This might be explained by knock-out experiments on
474 *AHK* receptors showing contrasting effects on flowering time or floral development across
475 Arabidopsis and rice (Burr et al., 2020). Another non-conserved GR was *ZHD5* that regulates

476 floral architecture and leaf development and is regulated by *MIF1* (*MINI ZINC-FINGER 1*) (Hong
477 et al., 2011), which also lacked network conservation. *ZHD5* regulation might thus be different
478 across species. Similarly, *FBX92* (*F-BOX PROTEIN92*) was not conserved, which might be
479 explained by the opposite effects on leaf size shown by *ZmFBX92* and *AtFBX92* gain of function
480 in Arabidopsis due to the presence of an F-box-associated domain in *AtFBX92*, lacking in
481 *ZmFBX92*. *FBX92* orthologs might thus undergo different transcriptional regulation (Baute et al.,
482 2017). *EPF1* (*EPIDERMAL PATTERNING FACTOR 1*) was also a non-conserved GR. This gene
483 affects stomatal density and water use efficiency. Recent work suggested that, in monocots and
484 dicots, *EPF1* orthologs probably have different temporal dynamics of gene expression in the
485 stomatal lineage (Buckley et al., 2020), which might result in different network gene neighbors.

486
487 Based on the validation results of our GR prediction pipeline, a correlation between network size
488 and recovery of genes affecting leaf size was observed. In particular, with increasing network size,
489 the recovery rate decreased, indicating that DS5 is not a recommended network density to use to
490 find growth regulators. The network neighborhood conservation of genes in the most stringent
491 networks involved different basal biological processes, suggesting their functional similarity
492 across monocots and dicots. Not surprisingly, genes involved in cell cycle regulation and plant
493 hormonal response were found, as both processes have a key role in leaf development. Several cell
494 cycle regulators were predicted as GRs, like the cyclin gene *CYCD3;3*, the *CDK* inhibitor *KRP3*
495 (*KIP-RELATED PROTEIN*), and a *DOF* transcription factor gene *OBP1* (*OBF BINDING*
496 *PROTEIN 1*) that controls cell cycle progression (Dewitte et al., 2007; Skirycz et al., 2008; Jun et
497 al., 2013). The auxin-responsive transcription factor gene *MONOPTEROS* (*MP*) is crucial for leaf
498 vascular development (Hardtke and Berleth, 1998), while the Aux/IAA gene that represses auxin
499 signaling, *AXR2*, whose gain-of-function leads to strong inhibition of leaf growth (Mai et al.,
500 2011), was also predicted. Besides auxin, brassinosteroid (BR) and gibberellin (GA) coordinately
501 play key roles in regulating plant cell elongation. The other two predicted transcription factor
502 genes, *HB25* (*HOMEBOX PROTEIN 25*) and *MYR1*, which modulate bioactive GA biosynthesis,
503 were also shown to have an effect on the petiole growth (Bueso et al., 2014). It is noteworthy that
504 nearly half of all the 34 genes with leaf phenotype were transcription regulators, which highlights
505 the importance of TF-mediated gene expression regulation during leaf development. In addition to
506 hormone-related genes and TFs, genes related to photosynthesis are also important for leaf

507 development. A carotenoid biosynthesis gene *LCY* and a chloroplast redox-regulating gene
508 *THIOREDOXIN X* were predicted as GR and have been shown to affect leaf size (Li et al., 2009;
509 Pulido et al., 2010). Moreover, the cytoplasmic carbonic anhydrase genes *CA2* and *BCA4* were
510 identified, consistent with the view that carbon utilization in leaves is closely linked to leaf area
511 (DiMario et al., 2016). Cell wall modification is considered to be another important determinant
512 of leaf development. The predicted candidate genes *LACCASE11 (LAC11)* and *CUTICLE*
513 *DESTRUCTING FACTOR 1 (CDEF1)*, encoding for a laccase that associates with the lignin
514 deposition in cell wall and a cutinase essential for the degradation of cell wall components,
515 respectively, are also involved in regulating leaf growth and morphology (Takahashi et al., 2010;
516 Qin et al., 2013). Among Arabidopsis genes with a reported phenotype in the RARGE II loss-of-
517 function dataset, *ACO2 (ACC OXIDASE 2)* led to increased leaf size, and AT3G43270, a member
518 of Plant invertase/pectin methylesterase inhibitor superfamily, to smaller leaves. GRs translated
519 from aspen led, through our integrative network approach, to the prediction of *NITRATE*
520 *TRANSPORTER 1.3 (NPF6.4/NRT1.3)* as a potential GR. In Arabidopsis shoot, the expression of
521 *AtNPF6.4/NRT1.3* was induced by nitrate (Okamoto et al., 2003) while, in *Medicago truncatula*,
522 *MtNRT1.3* shares 70% identity with *AtNPF6.4/NRT1.3* and was reported to be a dual-affinity
523 nitrate transporter (Morre-Le Paven et al., 2011). It was also hypothesized that *NPF6.4/NRT1.3*
524 may play a role in supplying nitrate to photosynthesizing cells (Tong et al., 2016). In our
525 experiments, we showed that this gene, when mutated, is altering leaf growth. This cross-species
526 conserved gene would thus contribute to nitrogen assimilation, that, closely interacting with carbon
527 metabolism, sustains plant growth and development (Nunes-Nesi et al., 2010). Due to the
528 relevance and the strong interconnection of the processes where *NPF6.4/NRT1.3* and many of the
529 candidate GRs here predicted, are involved in, future experimental work will have to reveal the
530 role of these candidate GRs in other organs.

531 In conclusion, the approach developed in this study fully exploits the potential of integrative
532 biology to translate and expand yield-related functional annotations in different plant species, as
533 such accelerating crop breeding.

534 **Materials and Methods**

535 **Integration of developmental expression datasets and network construction**

536 Transcriptomic datasets were obtained from a list of studies in Arabidopsis, maize and aspen
537 covering samples from the main leaf developmental phases (Supplemental Table S1, Supplemental
538 Methods, Supplemental Dataset S1). Details about these datasets and the processing of these
539 samples were reported in Supplemental Methods. Maize data was mainly composed by a
540 developmental compendium generated in this work (Supplemental Methods). The network
541 inference was carried out with Seidr (Schiffthaler et al., 2018), which infers gene networks by
542 using multiple inference algorithms and then aggregating them into a meta-network. This approach
543 has been shown to strongly improve the accuracy of the results (Marbach et al., 2012). Each
544 network was subset into five density subnetworks (DSs) using five different network density
545 values. This procedure consisted in selecting the top 0.1, 0.5, 1, 5 and 10% top Seidr links in each
546 species-specific network and generating five DSs (from the most stringent DS1 to the least
547 stringent DS5).

548 **Orthology and network neighborhood conservation**

549 To compute cross-species gene network neighborhood conservation, orthology information
550 between genes from Arabidopsis, maize and aspen was computed using the PLAZA comparative
551 genomics platform (Van Bel et al., 2018). A custom version of this platform was built covering in
552 total 15 eukaryotic species including *Arabidopsis thaliana* (TAIR10), *Eucalyptus grandis* (v2.0),
553 *Populus trichocarpa* (v3.01), *Populus tremula* (v1.1), *Vitis vinifera* (12X March 2010 release),
554 *Zea mays* (AGPv3.0), *Oryza sativa* ssp. *Japonica* (MSU RGAP 7), *Triticum aestivum* (TGACv1),
555 *Amborella trichopoda* (Amborella v1.0), *Picea abies* (v1.0), *Pinus taeda* (v1.01), *Selaginella*
556 *moellendorffii* (v1.0), *Physcomitrium patens* (v3.3), *Chlamydomonas reinhardtii* (v5.5) and
557 *Micromonas commode* (v3.0). PLAZA allows identifying orthologs using different methods
558 (evidences), corresponding to orthologous gene families inferred through sequence-based
559 clustering with OrthoFinder (Emms and Kelly, 2015), phylogenetic trees, and multispecies Best-
560 Hits-and-Inparalogs families (Van Bel et al., 2012). The PLAZA orthology relationships were
561 extracted and filtered retaining all orthologs having a requirement of 2/3 orthology evidences and,
562 for those with 1/3 evidence and >25 orthologs, the ones corresponding to the best 25 blast hits

563 (sorted by e-value) were retained. The generated orthology output was used for the following
564 pipeline steps.

565 The generated DSs and the orthology information were used to compare the three species using a
566 network neighborhood conservation analysis (ComPIEx analysis, as in Netotea et al. 2014). In this
567 analysis, the network neighborhood of a gene (i.e. all genes with a link to it) was considered
568 conserved if it had a statistically significant ($q < 0.05$) overlap with the network neighborhood of
569 its ortholog in the other species (Netotea et al., 2014). Here, the comparison was performed for all
570 pairs of networks between the datasets of the three species, and the output of this analysis was
571 collated to create “triplets”. The triplets are sets of three orthologous genes—one per
572 network/species—that have a significantly conserved network neighborhood in all three pairs of
573 comparisons. Since the test is not commutative, the neighborhoods had to be significantly
574 conserved in both directions of the test. To estimate the false discovery rate (FDR) of the detection
575 of triplets, a permutation strategy was adopted. For 500 runs of ComPIEx, ortholog relationships
576 were shuffled, keeping the relative number of orthologs per gene and per species, and then
577 comparing the number of triplets computed from randomization with those resulting using the
578 original (unshuffled) orthologs.

579 **Functional analyses and prediction of growth regulators**

580 Gene Ontology (Ashburner et al., 2000) functional annotations for Arabidopsis, maize and aspen
581 were retrieved from TAIR (download 25/12/2018), Gramene (AGPv3.30,
582 <http://bioinfo.cau.edu.cn/agriGO/download.php>), and PlantGenIE
583 (ftp://ftp.plantgenie.org/Data/PopGenIE/Populus_tremula/v1.1/annotation/), respectively, and
584 filtered for the genes present in the corresponding species networks. We focused on biological
585 processes (BP) and excluded the general GO BP terms with ≥ 1500 genes as well as GO terms
586 with ≤ 10 genes to avoid biases towards very general and specific terms. For each gene, all GO
587 annotations were recursively propagated in order to include parental GO terms. Functional over-
588 representation analyses were performed using the hypergeometric distribution together with
589 Benjamini-Hochberg (BH) correction for multiple testing (Benjamini and Hochberg, 1995). To get
590 a complete view on all relevant processes related to plant growth, information from literature was
591 collected on growth regulators (GRs). Experimentally validated genes in Arabidopsis, maize and
592 aspen (primary-GRs) were retrieved from public databases (Gonzalez et al., 2010; Beltramino et

593 al., 2018). Experimentally validated aspen genes were obtained by access to SweTree
594 Technologies private database that contains data from the large-scale testing of >1,000 genes and
595 their growth-related properties (here only “stem size” was taken into consideration), an effort
596 where more than 1,500 recombinant DNA constructs were used to either introduce a gene product
597 or alter the level of an existing gene product by over-expression or RNA interference in aspen
598 trees, whose growth characteristics were then monitored in greenhouse and field experiments to
599 provide extensive gene-to-yield data. The Arabidopsis GR primary set was then enlarged with high
600 quality GR orthologs from maize and aspen using the triplets (“translated-GRs”) to obtain a
601 combined GR set. The combined set was finally filtered with genefilter package from
602 Bioconductor (Gentleman et al., 2021) to remove genes with small expression variance
603 ($\text{var.func}=\text{IQR}$, $\text{var.cutoff}=0.8$) and focus on genes active during proliferation or expansion phases
604 of leaf development (“expression-supported GRs”, Supplemental Table S2). Other information on
605 functional categories (Vercruyssen et al., 2020a) and differentially expressed genes from relevant
606 studies on plant development was also included in the functional enrichment analyses (Anastasiou
607 et al., 2007; Gonzalez et al., 2010; Eloy et al., 2012; Vercruyssen et al., 2014).

608 The expression-supported GRs were used as guide genes to perform network-guided gene function
609 prediction via a guilt-by-association (GBA) approach. This approach is based on the assumption
610 that genes close to the input GRs in the network are likely to have similar functions. The GBA
611 approach was applied to attribute functions based on GO enrichment in the modules of each DS
612 yielding five sets of gene predictions. By this procedure, gene neighborhoods significantly
613 enriched for guide GRs were functionally annotated (hypergeometric distribution). This allowed
614 to predict candidate GRs and estimate, for each of them, a corresponding FDR adjusted p-value
615 (or q-value), which was renamed “GBA-score”. The GBA score is a confidence score that ranks
616 genes high if they are connected with many GRs in the network (in fact high ranked genes have a
617 low GBA score as this is an indicator of a strong enrichment). For an example GR prediction (in
618 one of any of the five DSs), the GBA-score from the five DSs was summarized taking the mean
619 of the GBA-scores and setting the GBA-score to 0.05 for the DSs where the gene was not
620 predicted. This yielded a list of GR predictions that was then further filtered by only retaining
621 those predictions having conserved neighborhood in at least one DS. To perform a validation of
622 the gene function predictions, the RARGE II (Akiyama et al., 2014) database was interrogated to
623 retrieve a list of Arabidopsis genes that, when mutated, showed an increased or decreased length,

624 width and size for rosette leaf, vascular leaf and cauline leaf (leaf trait genes). This gene set was
625 used to analyze the recovery at each DS of leaf growth-related phenotypes. For the top 100
626 predictions ranked by GBA-score a manual literature search was performed to retrieve all genes
627 with a reported phenotype including information about the biological pathway the gene might be
628 active in, and other public functional annotations.

629 **Rosette growth phenotyping**

630 The *Arabidopsis thaliana* ecotype Columbia-0 (Col-0) was used as the wild-type in this study. The
631 T-DNA insertion lines for At4g26530 (Salk_080758/*fba5-1*), At3g21670 (Salk_001553/*sper3-3*),
632 At3g61250 (Salk_066767/*lmi2-1*, Salk_020792/*lmi2-2*), At4g25240 (Salk_113731), At1g63470
633 (Salk_123590/*ahl5*), At4g37980 (Salk_001773/*chr hpl*), At2g38530 (Salk_026257/*ltp2-1*),
634 At4g28950 (Salk_019272), and At1g12240 (Salk_016136) were confirmed using PCR with a T-
635 DNA primer and gene-specific primers (Supplemental Table S8) (Lu et al., 2012; Zhao et al., 2013;
636 Jacq et al., 2017; Tanaka et al., 2018; Pastore et al., 2011; Tong et al., 2016). All tested seeds were
637 stratified in the darkness at 4 °C for 3 days and then sown on soil in the 7 cm wide square pots
638 with a density of four seeds per pot. After 8 days in the growth room (with controlled temperature
639 at 22 °C and light intensity 110 $\mu\text{mol m}^{-2} \text{s}^{-1}$ in a 16 h/8 h cycle), the four seedlings were screened,
640 leaving one seedling per pot, which most closely resembled the genotype average. The plants were
641 imaged in a phenotyping platform (MIRGIS) with fixed cameras located directly above the plants,
642 which images plants at the same time every day. These images were then processed to extract the
643 rosette growth parameters of each plant. The mean PRA, compactness and stockiness values were
644 calculated over time for each genotype.

645 **Accession Numbers**

646 Sequence data from this article have been submitted to ENA (E-MTAB-11108). NPF6.4/NRT1.3
647 and LATE MERISTEM IDENTITY2 have locus identifier AT3G21670 and AT3G61250,
648 respectively.

649 **Supplemental data**

650 Supplemental Figure S1. Number of neighbors per gene at each density subnetwork in
651 Arabidopsis.

652 Supplemental Figure S2. Expression patterns for the expression-supported growth regulators in
653 Arabidopsis.

654 Supplemental Figure S3. Expression-supported growth regulators with neighborhood conservation
655 at each network density level.

656 Supplemental Figure S4. Functional enrichment of cross-species conserved transcription factors
657 (TF) grouped by TF family.

658 Supplemental Figure S5. Identification of T-DNA insertion lines.

659 Supplemental Figure S6. The rosette leaf numbers of the wild-type Col-0 and the mutants of
660 *NRT1.3* and *LMI2*.

661 Supplemental Table S1. Overview of the expression datasets used for the network computation.

662 Supplemental Table S2. List of expression-supported growth regulators.

663 Supplemental Table S3. Predicted growth regulators.

664 Supplemental Table S4. List of RARGE II leaf trait genes known to affect leaf phenotype if
665 mutated.

666 Supplemental Table S5. Top 100 predicted growth regulators annotated.

667 Supplemental Table S6. In depth literature analysis for the top 100 predicted growth regulators.

668 Supplemental Table S7. List of genes tested for leaf phenotype in this study.

669 Supplemental Table S8. Primers used for T-DNA identification and qPCR.

670 Supplemental Dataset S1. Expression datasets for Arabidopsis, maize, and aspen.

671 Supplemental Dataset S2. Triplets generated with ComPlex.

672 Supplemental Methods. Detailed methods for expression dataset retrieval, generation, and
673 processing.

674

675 **Funding information:** TRH, KV and NRS was partially funded by the Research Council of
676 Norway (project number 287465).

677 **Acknowledgements**

678 We thank Taku Takahashi (Okayama University, Japan) for kindly sending us the *sper3-*
679 *1* and *sper3-3* seeds, and Julie Pevernagie for her technical support.

680 **Figure legends**

681 **Figure 1. Outline of the cross-species network approach to identify candidate growth**
682 **regulators.** For Arabidopsis, maize and aspen, the expression data (step 1) is used as input to
683 construct a fully connected meta-network per species (step 2). Subsequently, each meta-network
684 is split into five density subnetworks (DSs) by applying specific density cutoffs (step 3). These
685 DSs are the input for two different analyses: they are used first as input to compute cross-species
686 gene neighborhood conservation (step 4a). Secondly, they are used to predict functions via guilt-
687 by-association (step 4b). This leads to gene function annotations of query genes (blue circles)
688 based on prior knowledge on growth regulators (purple circles). Edge thickness defines in which
689 subnetwork the interaction is conserved (line thickness represents the DS and ranges from 1, the
690 most stringent DS represented by the thickest line, to 5, the least stringent DS represented by the
691 thinnest line). Finally, the results of these two analyses (steps 4a and 4b) are integrated to obtain a
692 list of candidate growth regulators (step 5).

693 **Figure 2. Triplets and their functional enrichments in cross-species conserved leaf networks.**
694 (A) The number of triplet genes showing cross-species gene neighborhood conservation is plotted
695 for all density subnetworks (DSs). (B) The biological process functional over-representation at
696 each DS is summarized for two sets: (1) all triplet genes (All) and (2) growth regulators and their
697 network neighbor (Growth regulator-related) triplet genes, subset of all triplet genes. Functional
698 categories marked with asterisks (*) belong to leaf growth modules described in Vercruysse et al.
699 (2020) and to the differentially expressed gene sets from relevant studies on plant development
700 (Bezhani et al., 2007; Gonzalez et al., 2010; Eloy et al., 2012; Vercruyssen et al., 2014; Vanhaeren
701 et al., 2017). For clarity, long biological process names have been abbreviated (§). (C) Overview
702 of growth regulators with (and without) cross-species neighborhood conservation at different DSs.

703 **Figure 3. Recovery of RARGE II leaf trait genes for each density subnetwork split in**
704 **proliferation and expansion.** The grey dashed line indicates the leaf-related phenotype gene
705 recovery expected by chance (within the RARGE II dataset).

706 **Figure 4. Gene-function network of the 34 phenotype-related genes out of the top 100**
707 **predicted growth regulators.** Predictions are clustered by expression profile (proliferation on the
708 left and expansion on the right). Node label colours from dark green (weak) to yellow (strong)
709 represent the reliability of the gene prediction (GBA score). Node border colours indicate known
710 growth regulators from Arabidopsis (black), known growth regulators from aspen (red), and
711 Arabidopsis known growth regulator paralogs (violet). Diamonds represent transcription factors.
712 Links from dark orange thick (DS1) to light orange thin (DS5) represent the density subnetwork
713 where the genes were found connected. Genes are linked with their respective growth-related
714 pathways (centered if connecting to both proliferation and expansion related genes) by grey links.
715 Anti-correlation links (connecting proliferation with expansion genes) were removed for clarity.

716 **Figure 5. Mutants of predicted growth regulators *NRT1.3* and *LMI2* showed altered rosette**
717 **growth.** (A-B) Dynamic growth analysis of projected rosette area, compactness and stockiness
718 over time of wild-type Col-0 and the mutants of *NRT1.3* (A) and *LMI2* (B) in soil. Values are
719 means \pm SD. For phenotypic analysis of mutants of *LMI2*, sample sizes (n) were n=16 for Col-0,
720 n=16 for *lmi2-2*, and n=17 for *lmi2-1*. For phenotypic analysis of mutants of *NRT1.3*, n=14 for
721 Col-0, n=15 for *sper3-1*, and n=13 for *sper3-3*. The asterisks represent the time points at which
722 differences in the PRA become significant between the mutants and wild-type, as determined by
723 Student's t test (*, $P < 0.05$; **, $P < 0.01$). The experiments were repeated three times with similar
724 results, and one representative experiment is shown. (C-D) Phenotype of 26-day-old mutants of
725 *NRT1.3* (C) and *LMI2* (D). Scale bar = 1 cm.

726 References

- 727 Akiyama K, Kurotani A, Iida K, Kuromori T, Shinozaki K, Sakurai T (2014) RARGE II: An
728 integrated phenotype database of arabidopsis mutant traits using a controlled vocabulary.
729 Plant Cell Physiol 55: 1–10
- 730 Anastasiou E, Kenz S, Gerstung M, MacLean D, Timmer J, Fleck C, Lenhard M (2007) Control
731 of Plant Organ Size by KLUH/CYP78A5-Dependent Intercellular Signaling. Dev Cell 13:
732 843–856
- 733 Andres RJ, Coneva V, Frank MH, Tuttle JR, Samayoa LF, Han SW, Kaur B, Zhu L, Fang H,

734 Bowman DT, et al (2017) Modifications to a LATE MERISTEM IDENTITY gene are
735 responsible for the major leaf shapes of Upland cotton (*Gossypium hirsutum* L.). *Proc Natl*
736 *Acad Sci U S A* 114: E57–E66

737 Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight
738 SS, Eppig JT, et al (2000) Gene ontology: Tool for the unification of biology. *Nat Genet* 25:
739 25–29

740 Bai MY, Fan M, Oh E, Wang ZY (2013) A triple helix-loop-helix/basic helix-loop-helix cascade
741 controls cell elongation downstream of multiple hormonal and environmental signaling
742 pathways in *Arabidopsis*. *Plant Cell* 24: 4917–4929

743 Barboza L, Effgen S, Alonso-blanco C, Kooke R, Keurentjes JJB, Koornneef M (2013)
744 *Arabidopsis* semidwarfs evolved from independent mutations in GA20ox1 , ortholog to green
745 revolution dwarf alleles in rice and barley. 110: 15818–15823

746 Baucher M, Jaziri M El, Vandeputte O (2007) From primary to secondary growth : origin and
747 development of the vascular system. 58: 3485–3501

748 Baute J, Polyn S, De Block J, Blomme J, Van Lijsebettens M, Inzé D (2017) F-box protein FBX92
749 affects leaf size in *Arabidopsis thaliana*. *Plant Cell Physiol* 58: 962–975

750 Baxter I (2020) We aren't good at picking candidate genes, and it's slowing us down. *Curr Opin*
751 *Plant Biol* 54: 57–60

752 Beltramino M, Ercoli MF, Debernardi JM, Goldy C, Rojas AML, Nota F, Alvarez ME,
753 Vercruyssen L, Inzé D, Palatnik JF, et al (2018) Robust increase of leaf size by *Arabidopsis*
754 *thaliana* GRF3-like transcription factors under different growth conditions. *Sci Rep* 8: 1–13

755 Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful
756 Approach to Multiple Testing. *J R Stat Soc Ser B* 57: 289–300

757 Bezhani S, Winter C, Hershman S, Wagner JD, Kennedy JF, Chang SK, Pfluger J, Su Y, Wagner
758 D (2007) Unique, shared, and redundant roles for the Arabidopsis SWI/SNF chromatin
759 remodeling ATPases Brahma and Splayed. *Plant Cell* 19: 403–416

760 Buckley CR, Caine RS, Gray JE (2020) Pores for Thought: Can Genetic Manipulation of Stomatal
761 Density Protect Future Rice Yields? *Front Plant Sci* 10: 1783

762 Bueso E, Muñoz-Bertomeu J, Campos F, Brunaud V, Martínez L, Sayas E, Ballester P, Yenush L,
763 Serrano R (2014) ARABIDOPSIS THALIANA HOMEODOMAIN25 uncovers a role for
764 gibberellins in seed longevity. *Plant Physiol* 164: 999–1010

765 Burr CA, Sun J, Yamburenko M V., Willoughby A, Hodgens C, Boeshore SL, Elmore A, Atkinson
766 J, Nimchuk ZL, Bishopp A, et al (2020) The HK5 and HK6 cytokinin receptors mediate
767 diverse developmental pathways in rice. *Development*. 147:dev191734

768 Cheng CY, Li Y, Varala K, Bubert J, Huang J, Kim GJ, Halim J, Arp J, Shih HJS, Levinson G, et
769 al (2021) Evolutionarily informed machine learning enhances the power of predictive gene-
770 to-phenotype relationships. *Nat Commun* 12: 1–15

771 De Smet R, Marchal K (2010) Advantages and limitations of current network inference methods.
772 *Nat Rev Microbiol* 8: 717–729

773 Dewitte W, Scofield S, Alcasabas AA, Maughan SC, Menges M, Braun N, Collins C, Nieuwland
774 J, Prinsen E, Sundaresan V, et al (2007) Arabidopsis CYCD3 D-type cyclins link cell
775 proliferation and endocycles and are rate-limiting for cytokinin responses. *Proc Natl Acad
776 Sci U S A* 104: 14537–14542

777 DiMario RJ, Quebedeaux JC, Longstreth DJ, Dassanayake M, Hartman MM, Moroney J V. (2016)
778 The cytoplasmic carbonic anhydrases β CA2 and β CA4 are required for optimal plant growth
779 at low CO₂. *Plant Physiol* 171: 280–293

780 Dubois M, Van den Broeck L, Inzé D (2018) The Pivotal Role of Ethylene in Plant Growth. *Trends*

781 Plant Sci 23: 311–323

782 Eloy NB, Gonzalez N, Van Leene J, Maleux K, Vanhaeren H, De Milde L, Dhondt S, Vercruyssen
783 L, Witters E, Mercier R, et al (2012) SAMBA, a plant-specific anaphase-promoting
784 complex/cyclosome regulator is involved in early development and A-type cyclin
785 stabilization. *Proc Natl Acad Sci U S A* 109: 13853–13858

786 Emms DM, Kelly S (2015) OrthoFinder: solving fundamental biases in whole genome
787 comparisons dramatically improves orthogroup inference accuracy. *BMC Bioinformatics* 16: 157

788 Eriksson ME, Israelsson M, Olsson O, Moritz T (2000) Increased gibberellin biosynthesis in
789 transgenic trees promotes growth, biomass production and xylem fiber length. *Nat Biotechnol*
790 18: 784–788

791 Ficklin SP, Feltus FA (2011) Gene coexpression network alignment and conservation of gene
792 modules between two grass species: Maize and rice. *Plant Physiol* 156: 1244–1256

793 Friedrichsen DM, Nemhauser J, Muramitsu T, Maloof JN, Alonso J, Ecker JR, Furuya M, Chory
794 J (2002) Three redundant brassinosteroid early response genes encode putative bHLH
795 transcription factors required for normal growth. *Genetics* 162: 1445–1456

796 Gentleman R, Carey VJ, Huber W, Hahne F (2022). *genefilter: genefilter: methods for filtering*
797 *genes from high-throughput experiments*. R package version 1.78.0.

798 Gong P, Demuyneck K, De Block J, Aesaert S, Coussens G, Pauwels L, Inzé D, Nelissen H (2022)
799 Modulation of the DA1 pathway in maize shows that translatability of information from
800 Arabidopsis to crops is complex. *Plant Sci*. 321:111295

801 Gonzalez N, de Bodt S, Sulpice R, Jikumaru Y, Chae E, Dhondt S, van Daele T, de Milde L,
802 Weigel D, Kamiya Y, et al (2010) Increased leaf size: Different means to an end. *Plant Physiol*
803 153: 1261–1279

804 Gray SB, Brady SM (2016) Plant developmental responses to climate change. *Dev Biol* 419: 64–
805 77

806 Hardtke CS, Berleth T (1998) The Arabidopsis gene MONOPTEROS encodes a transcription
807 factor mediating embryo axis formation and vascular development. *EMBO J* 17: 1405–1411

808 Heyndrickx KS, Vandepoele K (2012) Systematic identification of functional plant modules
809 through the integration of complementary data sources. *Plant Physiol* 159: 884–901

810 Hong SY, Kim OK, Kim SG, Yang MS, Park CM (2011) Nuclear import and DNA binding of the
811 ZHD5 transcription factor is modulated by a competitive peptide inhibitor in Arabidopsis. *J*
812 *Biol Chem* 286: 1659–1668

813 Inze D, Nelissen H (2022) The translatability of genetic networks from model to crop species:
814 lessons from the past and perspectives for the future. *New Phytol*

815 Ikeda M, Mitsuda N, Ohme-Takagi M (2013) ATBS1 Interacting Factors negatively regulate
816 Arabidopsis cell elongation in the triantagonistic bHLH system. *Plant Signal Behav.*
817 8:e23448

818 Jacq A, Pernot C, Martinez Y, Domergue F, Payré B, Jamet E, Burlat V, Pacquit VB (2017) The
819 Arabidopsis lipid transfer protein 2 (AtLTP2) is involved in cuticle-cell wall interface
820 integrity and in etiolated hypocotyl permeability. *Front Plant Sci* 8: 1–17

821 Jones DM, Vandepoele K (2020) Identification and evolution of gene regulatory networks: insights
822 from comparative studies in plants. *Curr Opin Plant Biol* 54: 42–48

823 Jun SE, Okushima Y, Nam J, Umeda M, Kim GT (2013) Kip-related protein 3 is required for
824 control of endoreduplication in the shoot apical meristem and leaves of Arabidopsis. *Mol*
825 *Cells* 35: 47–53

826 Klie S, Mutwil M, Persson S, Nikoloski Z (2012) Inferring gene functions through dissection of

827 relevance networks: Interleaving the intra- and inter-species views. *Mol Biosyst* 8: 2233–
828 2241

829 Leach KA, Tran TM, Slewinski TL, Meeley RB, Braun DM (2017) Sucrose transporter2
830 contributes to maize growth, development, and crop yield. *J Integr Plant Biol* 59: 390–408

831 Lee BH, Ko JH, Lee S, Lee Y, Pak JH, Kim JH (2009) The Arabidopsis GRF-Interacting Factor
832 gene family performs an overlapping function in determining organ size as well as multiple.
833 *Plant Physiol* 151: 655–668

834 Lee I, Ambaru B, Thakkar P, Marcotte EM, Rhee SY (2010) Rational association of genes with
835 traits using a genome-scale gene network for Arabidopsis thaliana. *Nat Biotechnol* 28: 149–
836 156

837 Lee T, Lee S, Yang S, Lee I (2019) MaizeNet: a co-functional network for network-assisted
838 systems genetics in Zea mays. *Plant J* 99: 571–582

839 Lee T, Yang S, Kim E, Ko Y, Hwang S, Shin J, Shim JE, Shim H, Kim H, Kim C, et al (2015)
840 AraNet v2: An improved database of co-functional gene networks for the study of
841 Arabidopsis thaliana and 27 other nonmodel plant species. *Nucleic Acids Res* 43: D996–
842 D1002

843 Lehti-Shiu MD, Panchy N, Wang P, Uygun S, Shiu SH (2017) Diversity, expansion, and
844 evolutionary novelty of plant DNA-binding transcription factor families. *Biochim Biophys*
845 *Acta - Gene Regul Mech* 1860: 3–20

846 Li J, Yuan D, Wang P, Wang Q, Sun M, Liu Z, Si H, Xu Z, Ma Y, Zhang B, et al (2021) Cotton
847 pan-genome retrieves the lost sequences and genes during domestication and selection.
848 *Genome Biol* 22: 1–26

849 Li T, Kang X, Lei W, Yao X, Zou L, Zhang D, Lin H (2020) SHY2 as a node in the regulation of
850 root meristem development by auxin, brassinosteroids, and cytokinin. *J Integr Plant Biol* 62:

851 1500–1517

852 Li Z, Ahn TK, Avenson TJ, Ballottari M, Cruz JA, Kramer DM, Bassi R, Fleming GR, Keasling
853 JD, Niyogi KK (2009) Lutein accumulation in the absence of zeaxanthin restores
854 nonphotochemical quenching in the arabidopsis thaliana npq1 mutant. *Plant Cell* 21: 1798–
855 1812

856 Liseron-Monfils C, Ware D (2015) Revealing gene regulation and associations through biological
857 networks. *Curr Plant Biol* 3–4: 30–39

858 Lu W, Tang X, Huo Y, Xu R, Qi S, Huang J, Zheng C, Wu C ai (2012) Identification and
859 characterization of fructose 1,6-bisphosphate aldolase genes in Arabidopsis reveal a gene
860 family with diverse responses to abiotic stresses. *Gene* 503: 65–74

861 Mai YX, Wang L, Yang HQ (2011) A Gain-of-Function Mutation in IAA7/AXR2 Confers Late
862 Flowering under Short-day Light in Arabidopsis. *J Integr Plant Biol* 53: 480–492

863 Marbach D, Costello JC, Küffner R, Vega NM, Prill RJ, Camacho DM, Allison KR, Kellis M,
864 Collins JJ, Aderhold A, et al (2012) Wisdom of crowds for robust gene network inference.
865 *Nat Methods* 9: 796–804

866 Marchant A, Bhalerao R, Casimiro I, Eklöf J, Casero PJ, Bennett M, Sandberg G (2002) AUX1
867 promotes lateral root formation by facilitating indole-3-acetic acid distribution between sink
868 and source tissues in the Arabidopsis seedling. *Plant Cell* 14: 589–597

869 Mizukami Y, Fischer RL (2000) Plant organ size control: AINTEGUMENTA regulates growth
870 and cell numbers during organogenesis. *Proc Natl Acad Sci U S A* 97: 942–947

871 Morre-Le Paven MC, Viau L, Hamon A, Vandecasteele C, Pellizzaro A, Bourdin C, Laffont C,
872 Lapied B, Lepetit M, Frugier F, et al (2011) Characterization of a dual-affinity nitrate
873 transporter MtNRT1.3 in the model legume *Medicago truncatula*. *J Exp Bot* 62: 5595–5605

874 Movahedi S, van De Peer Y, Vandepoele K (2011) Comparative network analysis reveals that
875 tissue specificity and gene function are important factors influencing the mode of expression
876 evolution in arabidopsis and rice. *Plant Physiol* 156: 1316–1330

877 Movahedi S, Van Bel M, Heyndrickx KS, Vandepoele K (2012) Comparative co-expression
878 analysis in plant biology. *Plant, Cell Environ* 35: 1787–1798

879 Mutwil M, Klie S, Tohge T, Giorgi FMM, Wilkins O, Campbell MMM, Fernie ARR, Usadel B,
880 Nikoloski Z, Persson S (2011) PlaNet: Combined Sequence and Expression Comparisons
881 across Plant Networks Derived from Seven Species. *Plant Cell* 23: 895–910

882 Nelissen H, Gonzalez N, Inzé D (2016) Leaf growth in dicots and monocots: So different yet so
883 alike. *Curr Opin Plant Biol* 33: 72–76

884 Nelissen H, Rymen B, Jikumaru Y, Demuynck K, Van Lijsebettens M, Kamiya Y, Inzé D,
885 Beemster GTS (2012) A local maximum in gibberellin levels regulates maize leaf growth by
886 spatial control of cell division. *Curr Biol* 22: 1183–1187

887 Netotea S, Sundell D, Street NR, Hvidsten TR (2014) ComPIEx: Conservation and divergence of
888 co-expression networks in *A. thaliana*, *Populus* and *O. sativa*. *BMC Genomics* 15: 106

889 Nowicka B (2019) Target genes for plant productivity improvement. *J Biotechnol* 298: 21–34

890 Nuccio ML, Paul M, Bate NJ, Cohn J, Cutler SR (2018) Where are the drought tolerant crops? An
891 assessment of more than two decades of plant biotechnology effort in crop improvement.
892 *Plant Sci* 273: 110–119

893 Nunes-Nesi A, Fernie AR, Stitt M (2010) Metabolic and signaling aspects underpinning the
894 regulation of plant carbon nitrogen interactions. *Mol Plant* 3: 973–996

895 Obertello M, Shrivastava S, Katari MS, Coruzzi GM (2015) Cross-species network analysis
896 uncovers conserved nitrogen-regulated network modules in rice. *Plant Physiol* 168: 1830–

897 1843

- 898 Oliver Stephen (2000) Guilt-by-association goes global. *Nature* 403: 601–603
- 899 Okamoto M, Vidmar JJ, Glass ADM (2003) Regulation of NRT1 and NRT2 gene families of
900 *Arabidopsis thaliana*: Responses to nitrate provision. *Plant Cell Physiol* 44: 304–317
- 901 Pastore JJ, Limpuangthip A, Yamaguchi N, Wu M-F, Sang Y, Han S-K, Malaspina L, Chavdaroff
902 N, Yamaguchi A, Wagner D (2011) LATE MERISTEM IDENTITY2 acts together with
903 LEAFY to activate APETALA1. *Development* 138: 3189–3198
- 904 Patel R V., Nahal HK, Breit R, Provart NJ (2012) BAR expressolog identification: Expression
905 profile similarity ranking of homologous genes in plant species. *Plant J* 71: 1038–1050
- 906 Pavlidis P, Gillis J (2012) Progress and challenges in the computational prediction of gene function
907 using networks. *F1000Research* 1: 2012–2013
- 908 Powell AE, Lenhard M (2012) Control of organ size in plants. *Curr Biol* 22: R360–R367
- 909 Pulido P, Spínola MC, Kirchsteiger K, Guinea M, Pascual MB, Sahrawy M, Sandalio LM, Dietz
910 KJ, González M, Cejudo FJ (2010) Functional analysis of the pathways for 2-Cys
911 peroxiredoxin reduction in *Arabidopsis thaliana* chloroplasts. *J Exp Bot* 61: 4043–4054
- 912 Qin X, Liu JH, Zhao WS, Chen XJ, Guo ZJ, Peng YL (2013) Gibberellin 20-Oxidase Gene
913 OsGA20ox3 regulates plant stature and disease development in rice. *Mol Plant-Microbe*
914 *Interact* 26: 227–239
- 915 Rhee SY, Mutwil M (2014) Towards revealing the functions of all genes in plants. *Trends Plant*
916 *Sci* 19: 212–221
- 917 Ruprecht C, Mutwil M, Saxe F, Eder M, Nikoloski Z, Persson S (2011) Large-scale co-expression
918 approach to dissect secondary cell wall formation across plant species. *Front Plant Sci* 2: 1–

919 13

920 Schiffthaler B, Serrano A, Street N, Delhomme N (2018) Seidr: A gene meta-network calculation
921 toolkit. bioRxiv 250696

922 Schneider M, Gonzalez N, Pauwels L, Inzé D, Baekelandt A (2021) The PEAPOD Pathway and
923 Its Potential To Improve Crop Yield. *Trends Plant Sci* 26: 220–236

924 Serin EAR, Nijveen H, Hilhorst HWM, Ligterink W (2016) Learning from co-expression
925 networks: Possibilities and challenges. *Front Plant Sci* 7: 444

926 Serrano-Mislata A, Sablowski R (2018) The pillars of land plants: new insights into stem
927 development. *Curr Opin Plant Biol* 45: 11–17

928 Simmons CR, Lafitte HR, Reimann KS, Brugière N, Roesler K, Albertsen MC, Greene TW,
929 Habben JE (2021) Successes and insights of an industry biotech program to enhance maize
930 agronomic traits. *Plant Sci*. doi: 10.1016/j.plantsci.2021.110899

931 Skiryicz A, Radziejwoski A, Busch W, Hannah MA, Czeszejko J, Kwaśniewski M, Zanol MI,
932 Lohmann JU, De Veylder L, Witt I, et al (2008) The DOF transcription factor OBP1 is
933 involved in cell cycle regulation in *Arabidopsis thaliana*. *Plant J* 56: 779–792

934 Soltész A, Vágújfalvi A, Rizza F, Kerepesi I, Galiba G, Cattivelli L, Coraggio I, Crosatti C (2012)
935 The rice *Osmyb4* gene enhances tolerance to frost and improves germination under
936 unfavourable conditions in transgenic barley plants. *J Appl Genet* 53: 133–143

937 Srivastava AC, Ganesan S, Ismail IO, Ayre BG (2008) Functional characterization of the
938 *Arabidopsis AtSUC2* sucrose/H⁺ symporter by tissue-specific complementation reveals an
939 essential role in phloem loading but not in long-distance transport. *Plant Physiol* 148: 200–
940 211

941 Stuart JM, Segal E, Koller D, Kim SK (2003) A gene-coexpression network for global discovery

942 of conserved genetic modules. *Science* 302: 249–55.

943 Sun X, Cahill J, Van Hautegeem T, Feys K, Whipple C, Novák O, Delbare S, Versteede C,
944 Demuyneck K, De Block J, et al (2017) Altered expression of maize PLASTOCHRON1
945 enhances biomass and seed yield by extending cell division duration. *Nat Commun* 8: 14752

946 Takahashi K, Shimada T, Kondo M, Tamai A, Mori M, Nishimura M, Hara-Nishimura I (2010)
947 Ectopic expression of an esterase, which is a candidate for the unidentified plant cutinase,
948 causes cuticular defects in *Arabidopsis thaliana*. *Plant Cell Physiol* 51: 123–131

949 Tanaka T, Ikeda A, Shiojiri K, Ozawa R, Shiki K, Nagai-Kunihiro N, Fujita K, Sugimoto K,
950 Yamato KT, Dohra H, et al (2018) Identification of a hexenal reductase that modulates the
951 composition of green leaf volatiles. *Plant Physiol* 178: 552–564

952 Tian Q, Reed JW (1999) Control of auxin-regulated root development by the *Arabidopsis thaliana*
953 SHY2/AA3 gene. *Development* 126: 711–721

954 Tong W, Imai A, Tabata R, Shigenobu S, Yamaguchi K, Yamada M, Hasebe M, Sawa S, Motose
955 H, Takahashi T (2016) Polyamine resistance is increased by mutations in a nitrate transporter
956 gene NRT1.3 (AtNPF6.4) in *Arabidopsis thaliana*. *Front Plant Sci* 7: 1–10

957 Vaishak KP, Yadukrishnan P, Bakshi S, Kushwaha AK, Ramachandran H, Job N, Babu D, Datta
958 S (2019) The B-box bridge between light and hormones in plants. *J Photochem Photobiol B*
959 *Biol* 191: 164–174

960 Van Bel M, Diels T, Vancaester E, Kreft L, Botzki A, Van De Peer Y, Coppens F, Vandepoele K
961 (2018) PLAZA 4.0: An integrative resource for functional, evolutionary and comparative
962 plant genomics. *Nucleic Acids Res* 46: D1190–D1196

963 Van Bel M, Proost S, Wischnitzki E, Movahedi S, Scheerlinck C, Van de Peer Y, Vandepoele K
964 (2012) Dissecting plant genomes with the PLAZA comparative genomics platform. *Plant*
965 *Physiol* 158: 590–600

966 Vanhaeren H, Nam YJ, De Milde L, Chae E, Storme V, Weigel D, Gonzalez N, Inzé D (2017)
967 Forever young: The role of ubiquitin receptor DA1 and E3 ligase big brother in controlling
968 leaf growth and development. *Plant Physiol* 173: 1269–1282

969 Verbraeken L, Wuyts N, Mertens S, Cannoot B, Maleux K, Demuyneck K, de Block J, Merchie J,
970 Dhondt S, Bonaventure G, et al (2021) Drought affects the rate and duration of organ growth
971 but not inter-organ growth coordination. *Plant Physiol* 186: 1336–1353

972 Vercruyssen J, Baekelandt A, Gonzalez N, Inzé D (2020a) Molecular networks regulating cell
973 division during Arabidopsis leaf growth. *J Exp Bot* 71: 2365–2378

974 Vercruyssen J, Van Bel M, Osuna-Cruz CM, Kulkarni SR, Storme V, Nelissen H, Gonzalez N, Inzé
975 D, Vandepoele K (2020b) Comparative transcriptomics enables the identification of
976 functional orthologous genes involved in early leaf growth. *Plant Biotechnol J* 18: 553–567

977 Vercruyssen L, Verkest A, Gonzalez N, Heyndrickx KS, Eeckhout D, Han S-K, Jegu T, Archacki
978 R, Van Leene J, Andriankaja M, et al (2014) ANGUSTIFOLIA3 Binds to SWI/SNF
979 Chromatin Remodeling Complexes to Regulate Transcription during Arabidopsis Leaf
980 Development. *Plant Cell* 26: 210–229

981 Vlad D, Kierzkowski D, Rast MI, Vuolo F, Dello Ioio R, Galinha C, Gan X, Hajheidari M, Hay
982 A, Smith RS, et al (2014) Leaf shape evolution through duplication, regulatory diversification
983 and loss of a homeobox gene. *Science* 343: 780–783

984 Vuolo F, Kierzkowski D, Runions A, Hajheidari M, Mentink RA, Gupta M Das, Zhang Z, Vlad
985 D, Wang Y, Pecinka A, et al (2018) LMI1 homeodomain protein regulates organ proportions
986 by spatial modulation of endoreduplication. *Genes Dev* 32: 1361–1366

987 Wolfe CJ, Kohane IS, Butte AJ (2005) Systematic survey reveals general applicability of “guilt-
988 by-association” within gene coexpression networks. *BMC Bioinformatics* 6: 227

989 Yang X, Lee S, So JH, Dharmasiri S, Dharmasiri N, Ge L, Jensen C, Hangarter R, Hobbie L,

990 Estelle M (2004) The IAA1 protein is encoded by AXR5 and is a substrate of SCF TIR1.
991 Plant J 40: 772–782

992 Zhao J, Favero DS, Peng H, Neff MM (2013) Arabidopsis thaliana AHL family modulates
993 hypocotyl growth redundantly by interacting with each other via the PPC/DUF296 domain.
994 Proc Natl Acad Sci U S A 110: 1–10

995