This is a post-peer-review, pre-copyedit version of an article published in *Plant Biotechnology Journal*. The final authenticated version is available online at: [https://doi.org/10.1111/pbi.13223](https://doi.org/10.1111/pbi.13223)

**TITLE**

Comparative transcriptomics enables the identification of functional orthologous genes involved in early leaf growth

**AUTHORS**

Jasmien Vercruysse¹,², Michiel Van Bel¹,², Cristina M. Osuna-Cruz¹,², Shubhada R. Kulkarni¹,², Véronique Storme¹,², Hilde Nelissen¹,², Nathalie Gonzalez¹,²,³,⁴, Dirk Inzé¹,²,⁴,*, Klaas Vandepoele¹,²,⁴

**AFFILIATIONS**

¹Department of Plant Biotechnology and Bioinformatics, Ghent University, Technologiepark 927, 9052 Gent, Belgium

²Center for Plant Systems Biology, VIB, Technologiepark 71, 9052 Gent, Belgium

³INRA, UMR1332 Biologie du fruit et Pathologie, INRA Bordeaux Aquitaine, CS20032, F-33882, Villenave d’Ornon cedex, France

⁴These authors contributed equally

**CONTACT INFORMATION**

*Correspondence: Dirk Inzé ([dirk.inze@psb.vib-ugent.be](mailto:dirk.inze@psb.vib-ugent.be))

**KEYWORDS**

*Arabidopsis thaliana, Zea mays*, orthology, data integration, networks, comparative transcriptomics, leaf development
ABSTRACT

Leaf growth is a complex trait for which many similarities exist in different plant species, suggesting functional conservation of the underlying pathways. However, a global view of orthologous genes involved in leaf growth showing conserved expression in dicots and monocots is currently missing. Here, we present a genome-wide comparative transcriptome analysis between Arabidopsis and maize, identifying conserved biological processes and gene functions active during leaf growth. Despite the orthology complexity between these distantly related plants, 926 orthologous gene groups including 2,829 Arabidopsis and 2,974 maize genes with similar expression during leaf growth were found, indicating conservation of the underlying molecular networks. We found 65% of these genes to be involved in one-to-one orthology, whereas only 28.7% of the groups with divergent expression had one-to-one orthology. Within the pool of genes with conserved expression, 19 transcription factor families were identified, demonstrating expression conservation of regulators active during leaf growth. Additionally, 25 Arabidopsis and 25 maize putative targets of the TCP transcription factors with conserved expression were determined based on the presence of enriched transcription factor binding sites. Based on large-scale phenotypic data, we observed that genes with conserved expression have a higher probability to be involved in leaf growth and that leaf-related phenotypes are more frequently present for genes having orthologs between dicots and monocots than clade-specific genes. This study shows the power of integrating transcriptomic with orthology data to identify or select candidates for functional studies during leaf development in flowering plants.

INTRODUCTION

Zea mays and Arabidopsis thaliana are two important model organisms for monocots and dicots, respectively, which diverged 140-200 million years ago (Chaw et al., 2004). Although they present numerous morphological and physiological differences, many developmental processes, such as leaf growth, show remarkable similarities (Li et al., 2010; Yu et al., 2015; Nelissen et al., 2016). In both organisms, leaf growth is characterized by two major consecutive phases: a cell proliferation phase, during which cells divide, and a cell expansion phase, during which cells increase their volume. The interplay between cell division and cell expansion determines final leaf size (Hepworth and Lenhard, 2014). However, the direction of growth is different in both species. In Arabidopsis, the cellular growth pattern is dispersed, resulting in round mature leaves with reticulate veins, while in maize, growth is directed in a longitudinal
fashion, generating narrow, elongated leaves with parallel veins (Nelson and Dengler, 1997; Nelissen et al., 2016). As soon as the leaf emerges from the shoot apical meristem, all cells from the primordium start dividing. After a few days, cells at the tip of the leaf cease division and start to elongate, marking the beginning of cell expansion. The leaf is then composed of both dividing and expanding cells during a so-called transition period, characterized by a cell-cycle arrest front moving from tip to base until most cells are no longer dividing and only expanding (Pyke et al., 1991; Donnelly et al., 1999; Andriankaja et al., 2012; Karidas et al., 2015; Nelissen et al., 2016).

Conservation of the cell division and cell expansion phases between both species suggests that the underlying molecular mechanisms are preserved. These conserved mechanisms involve orthologous genes that evolved from a gene in a common ancestor through a speciation event. Nevertheless, different events including gene duplication and functional divergence result in complex many-to-many orthology in plants, where potentially only a subset of orthologs have conserved functions whereas the other orthologs acquired different functions or are differently expressed. Sorting out the true functional orthologs which do share similar biological functions can be accomplished by combining transcriptomic data with orthology information (Movahedi et al., 2012; Patel et al., 2012). The resulting integrated network provides a powerful means to explore transcriptome-wide patterns of conservation and divergence between species (Movahedi et al., 2011). Several studies have already described orthologous genes with a conserved expression pattern and/or function in two or more species (Rensink et al., 2005; Libauilt et al., 2010; Mustroph et al., 2010; Narsai et al., 2011; Hefer et al., 2015). For example, members of the GROWTH-REGULATING FACTOR (GRF) transcription factor (TF) family plays key roles in leaf growth regulation and development and have been shown to interact with the co-activator GRF-INTERACTING FACTOR (GIF) both in Arabidopsis and maize (Kim and Kende, 2004; Zhang et al., 2008; Lee et al., 2009; Vercruyssen et al., 2014; Nelissen et al., 2015). EXPANSINs function in both Arabidopsis and maize as cell wall loosening proteins facilitating cell expansion and stimulating cell enlargement when overexpressed or applied exogenously (Cho and Cosgrove, 2000; Geilfus et al., 2015). In these examples, the similar function of orthologous genes hints at a conservation of the molecular mechanisms that regulate leaf growth.

To identify the molecular processes active during leaf development, several genome-wide transcriptome studies addressing changes in gene expression during early leaf growth, have been realized in Arabidopsis and maize (Li et al., 2010; Skirycz et al., 2011; Baerenfaller et al.,
2012; Avramova et al., 2015; Baute et al., 2016; Dubois et al., 2017; Nelissen et al., 2018). These experiments provide a birds-eye view on the expression of all genes possibly involved in leaf growth within a species. However, comparing gene expression data across multiple species requires accurate orthology information between the species. A network including orthology allows studying expression dynamics of different organisms at the gene level. Nonetheless, the current picture of all components involved in leaf growth shared between different plants is far from complete.

Here, we aimed at identifying common and divergent leaf growth related regulatory processes in a monocot and a dicot plant. To do so, the expression conservation between Arabidopsis and maize was studied during early leaf development. We built an integrated network based on expression data linked with complex gene orthology information to study similarities between the two species, allowing to list potential leaf growth regulating candidates. The level of conservation among Arabidopsis and maize TFs was also explored, indicating substantial conservation of transcriptional regulation.

RESULTS

Similar expression patterns during leaf development in Arabidopsis and maize

To discover common or divergent expression patterns between Arabidopsis and maize during leaf development, we compared gene expression profiles in both species. Two publicly available gene expression data sets were used (Andriankaja et al., 2012; Nelissen et al., 2018). Both sets include time-course gene expression information from samples harvested during leaf development and correspond to the three main developmental phases: cell proliferation, cell expansion, and the transition between these two phases (Figure S1a). For Arabidopsis, the expression information originated from six datapoints, in which two successive datapoints corresponded to one developmental phase. For maize, the expression information from samples harvested at nine datapoints was used, with four samples corresponding to the proliferating zone, three to the transitioning zone, and two to the expanding zone. In total, expression levels were measured for 29,920 Arabidopsis and 39,323 maize genes of which respectively 4,217 and 6,495 differentially expressed (DE) genes (DEGs) were selected.

To identify major expression trends in both sets, genes were clustered based on their co-expression during leaf development. Hierarchical clustering was used to group the DEGs in a
minimal set of clusters reporting coherent expression profiles within each cluster and distinct profiles between clusters (Methods, Figure S1b, S1c). To facilitate statistical data analysis, only clusters with at least 50 genes were retained. For the Arabidopsis dataset, initial clustering of the DEGs yielded 11 clusters, of which six containing more than 50 genes with distinct expression trends were retained: clusters A1 to A6 that represent 97% of the DEGs. For the maize data set, the same approach resulted in the selection of seven out of eight clusters (clusters Z1 to Z7), representing 99% of the DEGs. In both sets of clusters, three trends were visible, corresponding to a peak of expression during proliferation, during transition, or during expansion (Figure S1d, S1e). Additionally, in the maize dataset, a fourth expression trend was visible corresponding to expression peaks during proliferation and expansion, forming an inverse profile of the transition-specific expression trend. To identify molecular processes active during leaf growth, we performed GO enrichment on all Arabidopsis and maize gene clusters (Figure S2). Clusters with decreasing expression (A1 and Z1) were enriched for terms such as cell division and DNA replication, processes that occur during early leaf development (Andriankaja et al., 2012). In A2 however, terms such as responses to stress (wounding, hypoxia), meristem development as well as leaf polarity determination were enriched, of which the latter process is known to occur in very early leaf development (Stahle et al., 2009). In clusters with a transition-specific profile (A3, A4, Z2, Z3, and Z4), different terms were found enriched in function of the cluster. Cluster A3 was enriched for terms related to cell division, whereas A4 was enriched for galactolipid -and chlorophyll biosynthesis correlating with photosynthesis onset. Clusters Z2, Z3, and Z4 included genes leading to few enriched terms that suggested divergent biological processes such as root morphogenesis, fatty acid metabolic process, and protein translation respectively. Terms including photosynthesis and chlorophyll biosynthesis were found to be enriched in clusters with increasing expression during leaf development (A5, Z5, and Z6). This is to be expected since photosynthesis begins when leaves start to expand (Andriankaja et al., 2012; Van Dingenen et al., 2016). In contrast, A6 included terms related to plant defense (glucosinolate biosynthesis, (Wittstock and Burow, 2010)), energy production (myo-inositol hexakisphosphate biosynthesis), auxin signaling and response to brassinosteroid. For Z7 no enriched terms were found.

To conclude, we mostly found similar expression trends for Arabidopsis and maize DEGs during leaf development with coinciding gene functions, suggesting that common pathways control leaf development in both species.
Integration of orthology and expression information identifies conserved orthologs expressed during leaf development.

To identify conserved gene functions during leaf development in both species, gene orthology was determined and analyzed. Orthologous genes were identified with the PLAZA integrative orthology method (Van Bel et al., 2012) that combines three orthology prediction methods: orthologous gene families inferred through sequence-based clustering with OrthoMCL (O); reconciled phylogenetic trees (T); and multispecies best hits and inparalogs (BHI) families (B) (Methods). Whereas the latter are related to a simple BLAST-based approach, OrthoMCL and phylogenetic trees are more advanced methods better capturing complex orthology relationships. For instance, using the B evidence only, three orthologs of Arabidopsis GRF3 could be identified, whereas including O- and T evidence, twelve additional maize orthologs were identified (Figure 1a). To avoid missing any potential functionally conserved genes, we considered genes as orthologs when at least one type of evidence was found using PLAZA. At the genome-wide level, we found that 83.8% (22,987 genes) and 76.0% (29,874 genes), of the Arabidopsis and maize protein-encoding genes respectively, have orthologs in the other species (Figure 1b, 1c). Different levels of complexity were observed in the constructed orthology network composed of a total of 7,336 separated sub-networks (or orthologous groups): one-to-one (31%), one-to-many (28%), and many-to-many orthology (41%) (Figure 1d).

To identify orthology groups showing similar expression patterns in Arabidopsis and maize, we used the integrative orthology information to compare the expression clusters from Arabidopsis and maize. To do so, the overlap between two clusters was estimated by determining the number of Arabidopsis genes of a given cluster having orthologs in a given maize cluster and by calculating the significance of this overlap (Figure 2, Methods). Of the 1,811 genes of the Arabidopsis A1 cluster, there were 1,709 genes with maize orthologs, and for the maize Z1 cluster that includes 2,649 genes, there were 2,407 Arabidopsis orthologs. The comparison of these two clusters showed that 862 Arabidopsis genes had 924 maize orthologs and that this 50.4% overlap (862/1,709) was significant ($p<0.001$). The same comparisons were done for all clusters, and we observed a significant overlap for eight cluster comparisons ($p<0.001$).

Overall, when including significant and non-significant overlaps, four distinct scenarios could be distinguished (Figure 2). First, for five cluster pairs (A1-Z1, A5-Z5, A5-Z6, A6-Z5, and A6-Z6), a significant overlap of orthologous genes showing similar expression patterns was observed. This significant overlap suggests that in both species, gene function during leaf development.
development is probably conserved among orthologs with similar expression. Among these five cluster pairs, 217 one-to-one orthology pairs showing conserved expression were identified, including 69 groups with 110 unknown genes, which suggest 69 new maize and 41 new Arabidopsis candidates that might regulate leaf development in both species in a similar way (Table S1). Furthermore, the 217 one-to-one pairs with conserved expression corresponded to 91.2% of all DE one-to-one orthology pairs.

Second, in three cases (A2-Z3, A3-Z6, A4-Z6), there was a significant overlap between clusters with a divergent expression profile. For instance, in the A2-Z3 overlap (25 Arabidopsis and 22 maize genes), Arabidopsis genes showed high expression during proliferation, whereas for maize, the peak of expression occurred in the transition zone. The difference in expression was even more clear in the A3-Z6 and A4-Z6 significant overlaps. Clusters A3 and A4 showed high expression during transition, whereas in Z6 high expression occurred during cell expansion. This divergence could indicate that either the orthologous genes are involved in different processes or that they are part of a similar processes, however regulated in a different manner during leaf development. One-to-one orthology was only found in the A3-Z6 and A4-Z6 overlaps, corresponding to respectively 0.8% (two groups) and 1.3% (three groups) of all DE one-to-one orthology pairs.

Third, seven of all cluster comparisons did not show any significant overlap although the respective clusters showed a conserved expression profile (A2-Z1, A3-Z2, A3-Z3, A3-Z4, A4-Z2, A4-Z3, A4-Z4). Interestingly, six of these comparisons were between clusters that shared a transition-specific expression profile. The absence of significant overlap agrees with the poor GO enrichment overlap found for the five clusters (Figure S2). Finally, the other cluster comparisons between clusters with divergent profiles did not show any significant overlap.

**Detection of conserved processes involved in leaf development**

To identify the biological processes in which orthologs with conserved expression during leaf development take part, a GO enrichment analysis was done using the Arabidopsis shared orthologs from the five significant overlaps A1-Z1, A5-Z5, A5-Z6, A6-Z5, and A6-Z6, that exhibited conserved expression (Figure 3).

A significant overlap of orthologs was observed for the A1-Z1 comparison. The enriched terms were related to different steps of the cell cycle including the M-phase (cell plate formation and mitotic cytokinetic process) and the S-phase (regulation of DNA duplication). For example, we observed 16 members of the MINI-CHROMOSOME MAINTENANCE (MCM) complex and
19 members of the histone family. Interestingly, RNA methylation suggesting modification of the RNA metabolism and histone H3-K9 methylation resulting in gene silencing (Hu et al., 2019) were also found enriched in the orthologs showing a high expression during the proliferation phase.

The enriched terms in the A5-Z5 and A5-Z6 overlaps were mainly related to photosynthesis and energetic/carbohydrate metabolism. Within the genes part of the A5-Z5 overlap, 12 genes encoding glycoside hydrolases (including AtBGLU8 and AtBLG9), which are involved in cell expansion (Cosgrove, 2005; Minic and Jouanin, 2006), were present. Furthermore, 11 members of the glutathione-S-transferase (GST) gene family (including AtGSTU17 and AtGSTU29) were also found in the GO enrichment analysis. In leaves, GST transcripts have been reported to accumulate in older rather than in younger tissues (Gong et al., 2005), validating their allocation to the A5-Z5 overlap. In the shared Arabidopsis and maize orthologs from the A5-Z6 (389 Arabidopsis and 367 maize orthologs) overlap, aquaporins (11 genes all of which are either members of PLASMA MEMBRANE INTRINSIC PROTEIN or TONOPLAST MEMBRANE INTRINSIC PROTEIN families) were present. The aquaporin families have been shown to have an important role in expanding cells, causing turgor pressure by water uptake in Arabidopsis and maize (Ludevid et al., 1992; Chaumont et al., 1998).

In the A6-Z5 and A6-Z6 overlaps, instead of genes related to metabolism execution, we found genes related to regulatory processes such as hormone signaling, response to auxin and to brassinosteroid, and secondary metabolite biosynthesis (phenylpropanoid metabolic process). For example, the AUXIN-INDOLE-3-ACETIC-ACID (AUX/IAA) protein family, among which SHORT HYPOCOTYL 2 (SHY2) and SOLITARY ROOT (SLR) were found in these clusters. Additionally, genes encoding eight members of the SMALL AUXIN UP RNA (SAUR) protein family, of which some are known to regulate plasma membrane H^+-ATPase activity resulting in cell elongation, were found in the A6-Z6 overlap (Spartz et al., 2014). Several genes related to cell wall biosynthesis were also found, such as LACCASE17 (LAC17) and CHITINASE PROTEIN-LIKE2 (CTL2), involved in lignin biosynthesis, and XYLOGLUCAN ENDOTRANSGLUCOSYLASE/HYDROLASE7 (XTH7) (Berthet et al., 2011; Sánchez-Rodríguez et al., 2012).

In Arabidopsis, although the overlaps including A5 or A6 genes showed a general increased expression during leaf growth, it was obvious that the enriched categories are different in both overlaps with A5 containing genes related to metabolism while A6 contains genes that are more
related to regulatory processes (hormone signaling) and secondary metabolism. Interestingly, in maize, these two types of categories seem to be distributed both in clusters Z5 and Z6.

With this analysis, we were able to identify 1,967 Arabidopsis genes with 2,050 maize orthologs showing conserved expression and function, being part of a significant overlap. Conserved biological processes related, but not limited, to the cell cycle were found for 862 Arabidopsis and 924 maize orthologous genes with proliferation-specific expression in the A1-Z1 overlap. Interestingly, RNA -and histone methylation were also enriched. A large part of the 1,105 Arabidopsis and 1,126 maize orthologous genes with expansion-specific expression (A5-Z5, A5-Z6, A6-Z5, and A6-Z6) were clearly related to photosynthesis and carbohydrate metabolism, but also to hormonal signaling. To conclude, we were able to show large functional conservation among orthologs in dicots and monocots with conserved expression.

**Phenotypic analysis of Arabidopsis genes showing conserved and divergent expression profiles**

To further explore the impact of conserved expression on gene function during leaf development, publicly available phenotypic data were mapped onto the orthology and expression information we combined. TAIR (Berardini et al., 2015) and RARGE II (Akiyama et al., 2013) were consulted to generate a list of 738 Arabidopsis genes that if mutated, lead to a leaf-related phenotype such as an increased leaf rosette size (Methods). The genes in A1, A5, and A6 that had a significant overlap, were used in this analysis. For these clusters, three groups were defined: (i) a conserved group containing Arabidopsis genes with at least one ortholog showing conserved expression, (ii) a divergent group in which Arabidopsis genes only had orthologs with divergent expression, and (iii) a group of Arabidopsis genes lacking maize orthologs, thus representing species-specific genes. For example, cluster A5 included 119 genes with no maize orthologs and 756 which did have a DE maize ortholog. Of these 756 genes, 673 belonged to the conserved group and 83 to the divergent groups (Table S2). Phenotypic data was mapped onto these three groups for the three clusters. Globally, 4.0% (738/18,369) of the Arabidopsis genes lead to a leaf phenotype when mutated, while for A5 conserved and divergent groups, percentages of respectively 6.24% (42 genes; p-value 2.92E-02) and 7.23% (six genes; p-value 0.53) were found. For the A5 species-specific group, 3.36% of leaf-related phenotypes were found (four genes; p-value 0.87). These results suggest that DE A5 genes with DE maize orthologs during leaf development have a higher possibility to cause a leaf-related phenotype when mutated. For A1 and A6, the percentage of leaf-related phenotypes in the conserved group (3.58% and 2.86% respectively) and the divergent group (4.84% and 5.88%
respectively) was greater as compared to the species-specific group (2.94% and 2.47% respectively).

Over all groups, there were only nine genes that result in an increased leaf size when mutated, of which 3 could be confirmed (Storozhenko et al., 2002; Cheng et al., 2005; Gong et al., 2005; Brauner et al., 2014). In contrast, 49 genes were found that decreased leaf size when mutated, of which 13 could be confirmed (Chiang et al., 1995; Hobbie and Estelle, 1995; Kim et al., 2006; Peng et al., 2006; Lintala et al., 2007; Ohkama-Ohtsu et al., 2007; Mattioli et al., 2008; Xia et al., 2010; Šimková et al., 2012; Uematsu et al., 2012; Xiao et al., 2012; Yoshizawa et al., 2014; Liu et al., 2015). For both phenotypes, a larger number of genes was found in the conserved expression groups compared to the divergent expression groups, suggesting that expression conservation is an important property to identify genes involved in leaf development (Table S3).

In the A5 conserved group, five genes causing an increased leaf size through overexpression, corresponded to AT1G04680, GLUTAMATE:GLYOXYLATE AMINOTRANSFERASE1 (GGT1), 1-AMINOCYCLOPROPANE-1-CARBOXYLIC OXIDASE (ATACO2), RESPONSIVE TO DESICCATION22 (RD22), and STARCH-FREE1 (STF1). Apart from GGT1 and STF1, which are known to decrease leaf size when knocked out (Gong et al., 2005; Brauner et al., 2014), these genes were only described in RARGE II, and are thus interesting candidates for further characterization. Of these genes, we know that at least one maize ortholog showed conserved expression during leaf development and could thus also be involved in leaf size regulation. Additionally, in the A6 conserved group, TF DNA-BINDING WITH ONE ZINC FINGER2 (DOF2) was found to increase leaf size when overexpressed and is interestingly known to be expressed at the tip of the leaf in very early developmental stages (Gardiner et al., 2010). With this analysis, we conclude that for orthologs with conserved expression, a significant fraction is associated with known leaf phenotypes.

Apart from conservation analysis, also cases suggesting expression divergence of orthologs were investigated. For example, in the A5 divergent group, we identified GIBBERELLIN 3-OXIDASE 1 (GA3ox1), considered the last and rate-limiting enzyme in the gibberellin biosynthetic pathway leading to the production of bioactive gibberellins, that causes a semi-dwarfed phenotype in Arabidopsis when knocked out (Chiang et al., 1995; Fleet et al., 2003). Together with the expression of GA3ox1, which increases as leaf growth advances with a peak during cell expansion, the ga3ox1-3 mutant phenotype demonstrates a function for this gene in the control of leaf expansion (Davies, 2004; Mitchum et al., 2006). In contrast, the maize
GA3ox1 ortholog exhibited the inverse expression pattern with a peak in expression during cell proliferation and a decrease in expression as leaf growth advances coinciding with the accumulation of gibberellin at the transition zone, controlling the shift towards cell expansion at this boundary (Nelissen et al., 2012). Although both orthologs may metabolically carry out the same biochemical function, the regulation of this function is different in Arabidopsis and maize due the spatial organization of the leaf. In total, we could identify eight Arabidopsis genes of which four, three, and one belonged to the A1, A5, and A6 divergent groups respectively that had divergent expression patterns compared to their maize orthologs, and that exhibited an increase (one gene) or decrease (seven genes) in leaf size when mutated (Table S3). There were only four genes in the A5 (two genes) and A6 (two genes) groups with no orthologs that displayed an increase (one gene in A6NO) or decrease (two genes in A5NO and one gene in A6NO) in leaf size when mutated, suggesting an Arabidopsis-specific leaf-related function. These results reveal the potential to identify possible candidates that are either Arabidopsis-specific or gene candidates that have an altered expression profile, suggesting a change in regulation or function.

**Visualization of the integrated orthologous expression network**

To study the complexity of the genome-wide orthologous relations between clustered Arabidopsis and maize DEGs, a visual representation was generated using Cytoscape (Shannon et al., 2003). The complexity of the network was limited by only featuring the orthologous relations between DEGs of clusters A1-A6 and Z1-Z7 (Supplemental network file).

The filtered network included 2,371 Arabidopsis and 2,675 maize DEGs, visualized as nodes, that were connected through 4,829 edges, representing orthology. Genes were colored according to their expression profile trend. Additionally, edges were colored when the orthologs were part of a significant overlap. This approach allowed studying transcriptional conservation between both species at both the gene-and the genome-wide level.

The network was divided into 1,487 orthologous groups that can be seen as (parts of) gene families in which node color visualizes conserved/divergent expression. There were 421 orthologous groups with conserved expansion-specific expression, of which 67.2% had a one-to-one orthology relation. Conserved proliferation-specific expression was observed for 501 orthologous groups, of which 64.7% had one-to-one orthology relation. Only four groups containing ten genes were found with conserved transition-specific expression, of which three displayed one-to-one orthology. These data therefore revealed that 62.7% of the integrated orthology network showed expression conservation among orthologs. In contrast, 37.3% (561
groups) of the groups contained orthologs with divergent expression. Within these groups, only 161 corresponded to one-to-one orthology (28.7%), showing that orthologs with divergent expression have more complex orthology compared to completely conserved expression orthology groups. It seems thus more likely to observe orthologs with divergent expression in complex orthology groups with many-to-many orthology, suggesting neofunctionalization and subfunctionalization within such group. For example, among the 197 groups displaying one-to-two orthology (one Arabidopsis gene with two maize orthologs), for 59 groups one maize gene showed divergent expression (Table S4), suggesting a different role for the two orthologs. For instance, for Arabidopsis ALCOHOL DEHYDROGENASE (ADH), with high expression during proliferation, only ZmADH1 had conserved expression, whereas ZmADH2 had high expression during expansion. This example illustrates the duplication of a progenitor gene, after which the expression, and possibly the function, of ZmADH2 could have deviated (Dennis et al., 1985).

Next, the conservation between DEGs acting as TFs or involved in chromatin-remodeling (CR), and their DE orthologs, which are expected to regulate the expression of numerous downstream genes, were analyzed. There were 467 orthologs (encoding 211 Arabidopsis TFs and 256 maize TFs), which formed 106 orthologous groups corresponding to 47 TF -and transcriptional regulator (TR) families (Table S5). These 47 families could be divided into four groups based on the number of orthologs with conserved expression: complete conservation; more than 50% conservation; less than 50% conservation; and complete divergence (Table S6). There were 19 TF -and TR families, including 45 Arabidopsis genes and 56 maize orthologs, showing complete expression conservation. Among these families, two TF -and TR groups were noticeable: a first group was related to chromatin-mediated transcription regulation (among which HIGH MOBILITY GROUP (HMG) genes, GNAT family members, SNF2 genes, SET genes, and the ULTRAPETALAL1 (ULT1) TF), while a second group consisted of more specific regulators of gene expression such as YABBY, GRF, E2F and NAC TF families (Figure 4). YABBY TFs are involved in the determination of the adaxial/abaxial leaf polarity and correspondingly, all DE TFs of this family showed high expression during proliferation (Stahle et al., 2009). The major cell cycle regulators, the E2F/DP TFs, which regulate the G1 to S phase transition (De Veylder et al., 2002; Berckmans and De Veylder, 2009), and the GRF TFs (Gonzalez et al., 2012) involved in the regulation of cell proliferation, had similar expression profiles corresponding to high expression during proliferation.

We observed 15 gene families, among which members of the TEOSINTE BRANCHED1/CYCLOIDEA/PCF (TCP) TF family, for which the expression conservation
was higher than 50% (Figure S3). In Arabidopsis, class I and class II work antagonistically by activating and suppressing growth respectively, balancing leaf developmental processes (Nicolas and Cubas, 2016). TCP orthologs with conserved expression included the class II AtTCP3, AtTCP4, and ZmTCP23, and class I AtTCP21 and ZmTCP4 with high expression during expansion (Figure 5a). Divergent expression was observed in the group comprising AtTCP19, ZmTCP19, and ZmTCP21. AtTCP19 and ZmTCP19 had conserved expression, whereas AtTCP19 and ZmTCP21 differed in expression, which identifies ZmTCP19 as the most likely functional ortholog of AtTCP19 (Figure 5a). Other TF families with more than 50% of the orthologs with conserved expression included, but were not limited to, the CCAAT, AUX/IAA, and the C2H2 families.

Members of the MYB and bHLH family were categorized in the group with less than 50% conserved expression between orthologs. The rather low degree of conservation, 36.36% and 41.67% for Arabidopsis and maize respectively, can be explained by multiple duplication events in both species followed by functional divergence (Carretero-Paulet et al., 2010; Dubos et al., 2010; Feller et al., 2011). Only three TF families, i.e. LIM, TAZ, and BRI1-EMS-SUPPRESSOR 1 (BES1), included orthologs lacking expression conservation, suggesting that these proteins might operate differently during leaf development in both species.

We conclude that TF families playing key roles in cellular processes required for leaf growth frequently have similar expression profiles in both conserved and divergent orthology groups, indicating that they might also regulate the expression of similar target genes in both organisms.

**Integration of transcription factor binding site information to identify conserved regulators involved in leaf development**

The presence of enriched TFBSs in the promoters of orthologs with conserved expression suggests the preservation of the mode of action of the TFs and their targets during leaf growth in dicots and monocots. Therefore, we investigated whether known Arabidopsis TFBSs linked with cell cycle (E2F) or leaf maturation (TCP) were overrepresented in the promoters of Arabidopsis DEGs and maize orthologs with conserved expression. The studied binding sites were previously determined using protein binding microarrays (Weirauch et al., 2014). First, the proof of concept was established using the well-characterized Arabidopsis E2F TFs known to bind the degenerative WTTSSCSS motif and of which many Arabidopsis targets are already known and described (Vandepeoele et al., 2005). One specific instance of the degenerative E2F TFBSs, TTTCCCGC, was enriched in the promoters (500bp upstream from the start codon) of A1 (47 genes, enrichment fold 2.04, p-value 2.2E-06) and Z1 (71 genes, enrichment fold 2.36
and p-value 1.26E-11 (Figure S4), both containing genes with high expression during proliferation. Within this pool of genes, 12 Arabidopsis genes had 16 maize orthologs, indicating the presence of conserved E2F target genes (Table S7). Among these common targets were MCM2 and MCM5, involved in chromatin organization, as well as REPLICATION PROTEIN A 2B (RPA2), involved in DNA replication, repair, and recombination (Elmayan et al., 2005). These results were benchmarked using two previously published studies that identified E2F-regulated genes (Naouar et al., 2009), and E2F-bound genes (Verkest et al., 2014). The overlap between both studies resulted in 393 possible E2F targets. All 12 Arabidopsis genes were confirmed as previously described E2F targets, which validates our approach to identify putative, new TF targets in both species jointly.

To identify additional regulators and target genes involved in leaf development, the conservation of the known TCP TFBSs in both Arabidopsis and maize DEGs was studied. To do so, eight experimentally identified Arabidopsis TFBSs representing three class I TCPs (TCP16, TCP19, and TCP20) and five class II TCPs (TCP2, TCP3, TCP4, TCP5, and TCP24) were employed (Figure 5b) (Weirauch et al., 2014). After motif enrichment analysis, two TFBSs representing TCP3 and TCP24, both classified as JAW-TCPs responsible for leaf curvature (Efroni et al., 2008), were enriched in clusters A5 and Z5 with high expression during expansion and A1 with proliferation-specific expression (Table S8). Among the set of 229 A1 and 272 A5 genes with an enriched TCP TFBS in their promoters, 34 genes shared 34 orthologs from a total of 279 Z5 genes, of which 25 Arabidopsis genes and 25 maize orthologs showed conserved expansion-specific expression (Figure 5c). We can thus consider these 68 genes (i) to be differentially expressed during leaf growth with a peak in expression during cell proliferation or expansion, (ii) to be orthologs in Arabidopsis and maize, and (iii) to have a TCP binding site in their first 500bp that is enriched in their respective clusters. From these three observations, 68 genes could thus be identified as putative TCP TF targets. Among the set of targets, LOX2, a known target of TCP4 and TCP20 (Danisman et al., 2012) and its maize ortholog, GRMZM5G822593, were identified as TCP3 and/or TCP24 targets (Figure 5c).

In conclusion, through the integration of conserved TFs active during leaf development and TF binding site information, detailed gene regulatory sub-networks could be identified offering new insights into target genes for specific TFs in both species. For the E2F TFs we could show strong conservation of the TFBSs in both Arabidopsis and maize while for the TCP family, the TFBS conservation was not that pronounced.
DISCUSSION

Leaf growth has been studied extensively both in Arabidopsis and in maize and there are remarkable similarities, at cellular and molecular level, between the seemingly very different outcomes (for a review see Nelissen et al., 2016). Here, we aimed at answering the question to what extend there is a likely functional conservation between all orthologous genes expressed during leaf development in these distinct species.

With our analysis, we show that comparative analysis of orthologous gene expression profiles can help identifying potential new players involved in leaf growth. We found 217 one-to-one orthology groups with conserved expression, of which the gene function is likely to be similar in both species. Maize and/or Arabidopsis candidates could therefore be selected for functional characterization. Proof of concept is demonstrated with LOW PSII ACCUMULATION 1 (LPA1), involved in photosystem II -and chloroplast maintenance (Peng et al., 2006). It seems that the LPA1 expression is conserved as it takes part in a one-to-one orthology with GRMZM2G383154, previously described as a chloroplast maintenance gene as well (Huang et al., 2013). Among these 217 orthology groups, 41 with unknown genes in both species were identified which could be potential leaf growth regulators (Table S1).

From the one-to many orthology combined with the expression information, possible sub- or neofunctionalization of orthologous genes in both species could be shown by studying orthologous groups with divergent expression. This is best shown in one-to-two orthology, in which one of the two maize orthologs changed expression as a consequence of a duplication event, causing reduced selection to maintain the ancestral expression pattern and/or function, resulting in a tailored or new function for this ortholog. We were able to identify 59 one-to-two orthology groups, of which one of the two maize orthologues with divergent expression could have a new function (Table S4). Clearly, correctly sorting out these orthologs showing different levels of expression conservation is of key importance to select valid gene candidates when translating trait information from Arabidopsis to maize.

Furthermore, integration of transcriptomics data could help untangling many-to-many orthology and narrow down the genuine functional orthologs. For instance, REGULATORY PARTICLE AAA-ATPASE 2A (RPT2a), involved in the regulation of cell expansion duration (Ueda et al., 2004), is part of a complex orthologous group with 59 Arabidopsis genes and 77 maize orthologs. By including DEGs during leaf development into this orthology network, the most likely functional maize ortholog of RPT2a, GRMZM2G056569 could be identified since
it exhibits the same expression profile. Another example involves CYCLIN D3;1 (CYCD3;1) regulating the proliferation duration (Dewitte et al., 2007). CYCD3;1 is part of an orthologous group including ten Arabidopsis genes and 23 maize orthologs. After integration of the expression data, two maize orthologs, GRMZM2G140633 and GRMZM2G133413 with the same expression pattern, were found. These two genes could be more interesting candidates to study their role in cell proliferation in maize, compared with the other 21 maize orthologs (Supplemental network file). Narrowing down the orthologs to a set of potential functional orthologs by integrating gene expression data helps prioritizing and selecting potential functional orthologs simultaneously in Arabidopsis and maize for further functional studies.

To further aid the selection of potential new growth-regulating genes, we conducted a phenotypic analysis that revealed a stronger involvement of genes with orthologs in leaf-related phenotypes compared to specific-specific genes. Furthermore, this analysis yielded nine putative Arabidopsis candidates that are likely to increase leaf size when overexpressed (Table S3). Additionally, the phenotypic analysis allowed us to list four potential Arabidopsis-specific leaf-regulating genes that do not have a maize ortholog, but which do alter the leaf phenotype when mutated. Moreover, differences in leaf size and shape between both species could be partially shown as we identified eight Arabidopsis genes with nine maize orthologs exhibiting divergent expression involved in leaf growth.

Transcription factors and regulators that tightly control gene expression play an important role in plant development including leaf growth. Here, for 19 TF/TR families with DEGs, orthologs showed expression profiles that are similar during leaf growth in Arabidopsis and maize, suggesting that their transcriptional regulation is conserved. Some of the TFs have a well-known role in leaf development, such as the previously mentioned GRF, YABBY, and E2F family members. Other genes having orthologs with conserved expression are involved in CR such as the Arabidopsis ATXR6 gene and its maize ortholog, PHD7, members of the SET regulators (Baumbusch et al., 2001). ATXR6 interacts with the PROLIFERATING CELL NUCLEAR ANTIGEN (PCNA) and is upregulated during the S-phase of the cell cycle (Raynaud et al., 2006). The expression profile of ATXR6 and PHD7, showing high expression during the cell proliferation phase, coincides therefore with that of the known SET genes.

In addition, we show evidence for expression conservation among putative orthologous targets of TFs, based on similar, enriched TFBSs in the promoters of these orthologs. This level of conservation is interesting, knowing the evolutionary distance between both species and suggests possible conservation of TFBSs in other plant species. For the E2F TFs, 12
Arabidopsis and 16 maize genes were identified as putative targets, of which all of the Arabidopsis targets were previously described as E2F targets. Even in more complex TF families such as the TCP family, conservation of the TFBS could be shown, albeit less evident. This low conservation might be because the TFBSs used in this study are Arabidopsis-centered and maize promoters might harbor adapted or even completely different TFBSs recruiting conserved TFs. Secondly, the size of the studied promoters (500bp upstream from the start codon) was kept constant in both species while maize promoters are considered larger and more diffuse as compared to Arabidopsis promoters since the majority of the maize genome is covered by transposable elements (Lisch, 2013).

In conclusion, through the cross-species analysis of gene expression profiles functional orthologous candidates with conserved expression could be identified for further functional characterization in both species during leaf development, bridging the knowledge between different model species for flowering plants.

**METHODS**

**Selection of differentially expressed genes and clustering**

Normalized expression values from the transcriptomics analysis of Arabidopsis (Andriankaja et al., 2012) and maize (Nelissen et al., 2018) were used to select DEGs. The data sets for both species were processed in a similar manner. First, fold changes and false discovery rates (FDRs) for each gene were calculated by comparing each datapoint to every other datapoint. For the Arabidopsis data set, with six datapoints, 15 comparisons \((6 \times (6 - 1))/2\) were done per gene, and for the maize data set including nine datapoints, 36 comparisons were done. Next, DEGs were selected based on \(-2 \geq \text{fold change} \geq 2\) and \(\text{FDR} \leq 0.05\). The resulting DEGs in both datasets corresponded to the union of all comparisons. Both DE data sets were clustered according to hierarchical clustering in R version 3.4.0 (R Core Team, 2017). Distances between genes were calculated with the Euclidean distance measure and complete linkage was used as agglomeration method. All cluster plots were generated with the ggplot2 R-package (Wickham, 2016).

**GO enrichment**

GO enrichment analysis was done using the PLAZA 3.0 Dicots platform. GO enrichment results were filtered based on \(p\text{value} \leq 1E^{-3}\) and \(\text{fold change} \geq 2\). Only biological process (and molecular function in case of separate clusters) terms were considered. The
heatmaps that display the GO enrichment results were created in R version 3.4.4 with the ComplexHeatmap package (Gu et al., 2016). The rows with GO terms were clustered with hierarchical clustering (Euclidean distance measure, clustering method complete linkage). Finally, the top ten GO terms with enrichment in at least one cluster or cluster overlap were displayed.

**Ortholog identification between Arabidopsis and maize**

PLAZA was used to determine the orthologous genes between Arabidopsis and maize (Van Bel et al., 2012; Proost et al., 2015). This platform includes nine species: four dicots (*Arabidopsis thaliana* (TAIR10), *Arabidopsis lyrata* (JGI v1.0), *Populus trichocarpa* (JGI assembly release v3.0, annotation v3.0), and *Solanum lycopersicum* (ITAG 2.4 on genome build SL2.50)), four monocot species (*Zea mays* (AGPv3.0), *Sorghum bicolor* (v3.1), *Oryza sativa ssp. Japonica* (MSU RGAP 7), and *Brachypodium distachyon* (v3.1)), and additionally, an out-group species, *Amborella trichopoda* (Amborella v1.0), to root the phylogenetic tree of those species. This PLAZA version allows to identify orthologs with four methods, corresponding to orthologous gene families inferred through sequence-based clustering with OrthoMCL (Li et al., 2003), phylogenetic trees, multispecies best hits and inparalogs families (Linard et al., 2011) and synteny information, although we did not use the latter method. In order to establish an orthologous relation between two genes, we used a threshold of one prediction method.

**Significant overlap between Arabidopsis and maize clusters**

We used a random sampling analysis to calculate the significant overlap. Each Arabidopsis cluster was compared to each maize cluster. The overlap between two clusters was determined as the group of orthologous Arabidopsis and maize genes present in those clusters. In order to determine significant overlaps between Arabidopsis and maize clusters, the p-value was calculated by comparing randomly sampled maize clusters to each existing Arabidopsis cluster and counting the overlap again. This was done 1,000 times so the p-value could be determined at the 0.001 level.

**In silico leaf-related phenotypic analysis**

TAIR (Berardini et al., 2015) and RARGE II (Akiyama et al., 2013) were consulted in order to perform the phenotypic analysis. The TAIR germplasm list contained phenotypic information for 1841 Arabidopsis mutant lines. This gene list was filtered using the description of the phenotype based on the presence of the ‘leaf’ or ‘leaves’ keywords, and then manually curated
for overexpressed or knocked out genes that lead to a leaf-related phenotype. This process resulted in a curated list of 291 genes. The RARGEII database included 17,809 Arabidopsis genes of which 2,726 showed a phenotype. 453 genes were selected since they were annotated with rosette leaf terms according to the Plant Ontology structure: wrinkled, upturned, epinastic, coiled, increased or decreased in length, size, and width, present in greater or fewer numbers in organism, and abnormal. Genes involved in a pale, dark, or variegated coloring of the rosette leaves were left out of the analysis. Combining curated data from both databases resulted in a list of 738 unique genes with leaf-related phenotypes when overexpressed or knocked out. For each gene group the enrichment and significance of the leaf-related phenotypes was calculated using the hypergeometric score and corrected for multiple testing with the Benjamini & Hochberg method (Benjamini and Hochberg, 1995).

Network visualization

Cytoscape 3.1 (Shannon et al., 2003) was used to visually represent the integrated comparative network. Orthologous information was used to create the network: genes were represented as nodes and their orthology relation as edges. Gene expression data together with species information and gene aliases were added as attribute format and visually represented by differently colored nodes in the network. The .cys file is made available as supporting file.

Transcription factor binding site conservation

Cis-BP was consulted for all directly inferred TCP binding sites in position weight matrix (PWM) format of Arabidopsis (Weirauch et al., 2014). All TFBSs were mapped on the first 500 bp (including the 5’ UTR) of the promoter using the Cluster-Buster algorithm with settings: background model based on 3-kb input sequences, cluster score threshold 0, and pseudo-count 0.375) (Frith et al., 2003). All mappings were ranked according to p-value and the top 3,000 hits for unique genes were used for calculating the enrichment. The enrichment of all TFBSs in each Arabidopsis and maize cluster was calculated employing the hypergeometric score and corrected for multiple testing with the Benjamini & Hochberg method (Benjamini and Hochberg, 1995).

ACKNOWLEDGEMENTS

We thank Annick Bleys for proofreading and preparing the manuscript. This work was supported by the European Research Council under the European Community's Seventh
Framework Programme [FP7/2007-2013] under ERC grant agreement n° [339341-AMAIZE]11 and by Ghent University (“Bijzonder Onderzoeksfonds Methusalem project” no. BOF08/01M00408 and Multidisciplinary Research Partnership “Biotechnology for a Sustainable Economy” Grant 01MR0510W).

AUTHOR CONTRIBUTIONS
J.V., K.V., N.G., and D.I. designed the research methodology, performed data analysis, and wrote the article. M.V.B. assisted in computing orthologs. S.R.K. helped with analysis of the PWMs and C.O-C generated the heatmaps. V.S. preprocessed the transcriptomics data sets. H.N. assisted in the biological interpretation of the maize genes.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

REFERENCES

Akiyama, K., Kurotani, A., Iida, K., Kuromori, T., Shinozaki, K., and Sakurai, T. (2013). RARGE II: an integrated phenotype database of Arabidopsis mutant traits using a controlled vocabulary. Plant and Cell Physiology. 55: e4-e4.

Andriankaja, M., Dhoudt, S., De Bodt, S., Vanhaeren, H., Coppens, F., De Milde, L., Mühlenbock, P., Skirycz, A., Gonzalez, N., Beemster, G.T.S., et al. (2012). Exit from proliferation during leaf development in Arabidopsis thaliana: a not-so-gradual process. Dev. Cell. 22: 64-78.

Avramova, V., AbdElgawad, H., Zhang, Z., Fotschki, B., Casadevall, R., Vergauwen, L., Knapen, D., Taleisnik, E., Guisez, Y., and Asard, H. (2015). Drought induces distinct growth response, protection and recovery mechanisms in the maize leaf growth zone. Plant physiology. pp. 00276.02015.

Baerenfaller, K., Massonnet, C., Walsh, S., Baginsky, S., Bühlmann, P., Hennig, L., Hirsch-Hoffmann, M., Howell, K.A., Kahlau, S., and Radziejwoski, A. (2012). Systems-based analysis of Arabidopsis leaf growth reveals adaptation to water deficit. Mol. Syst. Biol. 8: 606.

Baumbusch, L.O., Thorstensen, T., Krauss, V., Fischer, A., Naumann, K., Assalkhou, R., Schulz, I., Reuter, G., and Aalen, R.B. (2001). The Arabidopsis thaliana genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. Nucl. Acids Res. 29: 4319-4333.

Baute, J., Herman, D., Coppens, F., De Block, J., Slabbinck, B., Dell’Acqua, M., Pé, M.E., Maere, S., Nelissen, H., and Inzé, D. (2016). Combined large-scale phenotyping and transcriptomics in maize reveals a robust growth regulatory network. Plant Physiol. 170: 1848-1867.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological). 57: 289-300.

Berardini, T.Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E., and Huala, E. (2015). The Arabidopsis information resource: making and mining the “gold standard” annotated reference plant genome. genesis. 53: 474-485.

Berckmans, B., and De Veylder, L. (2009). Transcriptional control of the cell cycle. Curr. Opin. Plant Biol. 12: 599-605.
Berthet, S., Demont-Caulet, N., Pollet, B., Bidzinski, P., Cézard, L., Le Bris, P., Borrega, N., Hervé, J., Blondet, E., and Balzergue, S. (2011). Disruption of LACCASE4 and 17 results in tissue-specific alterations to lignification of Arabidopsis thaliana stems. The Plant Cell. 23: 1124-1137.

Brauner, K., Hörmiller, I., Nägele, T., and Heyer, A.G. (2014). Exaggerated root respiration accounts for growth retardation in a starchless mutant of Arabidopsis thaliana. The Plant Journal. 79: 82-91.

Carretero-Paulet, L., Galetyan, A., Roig-Villanova, I., Martínez-García, J.F., Bilbao-Castro, J.R., and Robertson, D.L. (2010). Genome-wide classification and evolutionary analysis of the bHLH family of transcription factors in Arabidopsis, poplar, rice, moss, and algae. Plant Physiol. 153: 1398-1412.

Chaumont, F., Barrieu, F., Herman, E.M., and Chrispeels, M.J. (1998). Characterization of a maize tonoplast aquaporin expressed in zones of cell division and elongation. Plant Physiology. 117: 1143-1152.

Chaw, S.-M., Chang, C.-C., Chen, H.-L., and Li, W.-H. (2004). Dating the monocot–dicot divergence and the origin of core eudicots using whole chloroplast genomes. J. Mol. Evol. 58: 424-441.

Cheng, N.-H., Pittman, J.K., Shigaki, T., Lachmansingh, J., LeClere, S., Lahner, B., Salt, D.E., and Hirschi, K.D. (2005). Functional association of Arabidopsis CAX1 and CAX3 is required for normal growth and ion homeostasis. Plant Physiology. 138: 2048-2060.

Chiang, H.-H., Hwang, I., and Goodman, H.M. (1995). Isolation of the Arabidopsis GA4 locus. The Plant Cell. 7: 195-201.

Cho, H.-T., and Cosgrove, D.J. (2000). Altered expression of expansin modulates leaf growth and pedicel abscission in Arabidopsis thaliana. Proc. Natl. Acad. Sci. USA. 97: 9783-9788.

Cosgrove, D.J. (2005). Growth of the plant cell wall. Nat. Rev. Mol. Cell Biol. 6: 850-861.

Danisman, S., van der Wal, F., Bhondt, S., Waite, R., de Folter, S., Bimbo, A., van Dijk, A.D.J., Muino, M., Chang, C.-P., and Cosgrove, D.J. (2000). Altered expression of expansin modulates leaf growth and pedicel abscission in Arabidopsis thaliana. Plant Physiol. 123: 1360-1368.

Dennis, E.S., Sachs, M.M., Gerlach, W.L., FinneGAN, E.J., and Peacock, W.J. (1985). Molecular analysis of the alcohol dehydrogenase 2 (Adh2) gene of maize. Nucl. Acids Res. 13: 727-743.

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

De Veylder, L., Beeckman, T., Beemster, G.T.S., de Almeida Engler, J., Ormenese, S., Maes, S., Naudts, M., Van Der Schueren, E., Jacqmard, A., Engler, G., et al. (2002). Control of proliferation, endoreduplication and differentiation by the Arabidopsis E2Fa-DPa transcription factor. EMBO J. 21: 1360-1368.

Donnelly, P.M., Bonetta, D., Tsukaya, H., Dengler, R.E., and Dengler, N.G. (1999). Cell cycling and cell enlargement in developing leaves of Arabidopsis. Dev. Biol. 215: 407-419.

Dubois, M.,Claeys, H., Van den Broeck, L., and Inzé, D. (2017). Time of day determines Arabidopsis transcriptionite and growth dynamics under mild drought. Plant Cell Environ. 40: 180-189.

Dubos, C., Stracke, R., Grotewold, E., Weishaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in Arabidopsis. Trends Plant Sci. 15: 573-581.

Efroni, I., Blum, E., Goldshmidt, A., and Eshed, Y. (2008). A Protracted and Dynamic Maturation Schedule Underlies Arabidopsis Leaf Development. The Plant Cell. 20: 2293-2306.

Elmayan, T., Proux, F., and Vacheret, H. (2005). Arabidopsis RPA2: A Genetic Link among Transcriptional Gene Silencing, DNA Repair, and DNA Replication. Current Biology. 15: 1919-1925.

Feller, A., Machemer, K., Braun, E.L., and Grotewold, E. (2011). Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. Plant J. 66: 94-116.

Fleet, C.M., Yamaguchi, S., Hanada, A., Kawai, H., David, C.J., Kamiya, Y., and Sun, T.-p. (2003). Overexpression of AtCPS and AtKS in Arabidopsis confers increased ent-kaurene production but no increase in bioactive gibberellins. Plant physiology. 132: 830-839.

Frith, M.C., Li, M.C., and Weng, Z. (2003). Cluster-Buster: finding dense clusters of motifs in DNA sequences. Nucleic Acids Res. 31: 3666-3668.

Gardiner, J., Sherr, I., and Scarpella, E. (2010). Expression of DOF genes identifies early stages of vascular development in Arabidopsis leaves. The International Journal of Developmental Biology. 54: 1389-1396.

Geiß, C.-M., Ober, D., Eichacker, L.A., Mühling, K.H., and Zörb, C. (2015). Down-regulation of ZmEXP6 (Zea mays β-expansin 6) protein is correlated with salt-mediated growth reduction in the leaves of Z. mays L. J. Biol. Chem. 290: 11235-11245.
Gong, H., Jiao, Y., Hu, W.-w., and Pua, E.-C. (2005). Expression of glutathione-S-transferase and its role in plant growth and development in vivo and shoot morphogenesis in vitro. Plant molecular biology. 57: 53-66.

Gonzalez, N., Vanhaeren, H., and Inzé, D. (2012). Leaf size control: complex coordination of cell division and expansion. Trends Plant Sci. 17: 332-340.

Gu, Z., Els, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 32: 2847-2849.

Hefer, C.A., Mizrachi, E., Myburg, A.A., Douglas, C.J., and Mansfield, S.D. (2015). Comparative interrogation of the developing xylem transcriptomes of two wood-forming species: Populus trichocarpa and Eucalyptus grandis. New Phytol. 206: 1391-1405.

Hepworth, J., and Lenhard, M. (2014). Regulation of plant lateral-organ growth by modulating cell number and size. Current Opinion in Plant Biology. 17: 36-42.

Hobbie, L., and Estelle, M. (1995). The axr4 auxin-resistant mutants of Arabidopsis thaliana define a gene important for root gravitropism and lateral root initiation. The Plant Journal. 7: 211-220.

Hu, J., Manduzio, S., and Kang, H. (2019). Epitranscriptomic RNA Methylation in Plant Development and Abiotic Stress Responses. Front Plant Sci. 10: 500.

Huang, M., Frigo, G., Nishimura, K., Qu, X., Oliinares, P.D.B., Majeran, W., Sun, Q., and van Wijk, K.J. (2013). Construction of Plastid Reference Proteomes for Maize and Arabidopsis and Evaluation of Their Orthologous Relationships; The Concept of Orthoproteomics. Journal of Proteome Research. 12: 491-504.

Karidas, P., Challal, K.R., and Nath, U. (2015). The tarani mutation alters surface curvature in Arabidopsis leaves by perturbing the patterns of surface expansion and cell division. Journal of experimental botany. 66: 2107-2122.

Kim, J.H., and Kende, H. (2004). A transcriptional coactivator, AtGIF1, is involved in regulating leaf growth and morphology in Arabidopsis. Proc. Natl. Acad. Sci. USA. 101: 13374-13379.

Kim, K.-C., Fan, B., and Chen, Z. (2006). Pathogen-induced Arabidopsis WRKY7 is a transcriptional repressor and enhances plant susceptibility to Pseudomonas syringae. Plant physiology. 142: 1180-1192.

Lee, B.H., Ko, J.-H., Lee, S., Lee, Y., Pak, J.-H., and Kim, J.H. (2009). The Arabidopsis GRF-INTERACTING FACTOR gene family performs an overlapping function in determining organ size as well as multiple developmental properties. Plant Physiol. 151: 655-668.

Li, L., Stecockt Jr., C.J., and Roos, D.S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. 13: 2178-2189.

Li, P., Ponnala, L., Gandotra, N., Wang, L., Si, Y., Tausta, S.L., Kebrom, T.H., Provart, N., Patel, R., Myers, C.R., et al. (2010). The developmental dynamics of the maize leaf transcriptome. Nat Genet. 42: 1060-1067.

Libault, M., Farmer, A., Joshi, T., Takahashi, K., Langley, R.J., Franklin, L.D., He, J., Xu, D., May, G., and Stacey, G. (2010). An integrated transcriptome atlas of the crop model Glycine max, and its use in comparative analyses in plants. Plant J. 63: 86-99.

Linard, B., Thompson, J.D., Poch, O., and Lecompte, O. (2011). OrthoInspector: comprehensive orthology analysis and visual exploration. BMC Bioinformatics. 12: 11.

Lintala, M., Allahverdiyeva, Y., Kidron, H., Piippo, M., Battchikova, N., Suorsa, M., Rintamäki, E., Salminen, T.A., Aro, E.M., and Mulo, P. (2007). Structural and functional characterization of ferredoxin-NADP+-oxidoreductase using knock-out mutants of Arabidopsis. The Plant Journal. 49: 1041-1052.

Liu, D., Wu, Z., and Hou, L. (2015). Loss-of-function mutation in SCY1 triggers chloroplast-to-nucleus retrograde signaling in Arabidopsis thaliana. Biologia plantarum. 59: 469-476.

Ludevid, D., Höfte, H., Himelblau, E., and Chrispeels, M.J. (1992). The expression pattern of the tonoplast intrinsic protein γ-TIP in Arabidopsis thaliana is correlated with cell enlargement. Plant physiology. 100: 1633-1639.

Mattioli, R., Marchese, D’Angeli, S., Altamura, M.M., Costantini, P., and Trovato, M. (2008). Modulation of intracellular proline levels affects flowering time and inflorescence architecture in Arabidopsis. Plant Mol Biol. 66: 277-288.

Minic, Z., and Jovanin, L. (2006). Plant glycoside hydrolases involved in cell wall polysaccharide degradation. Plant Physiol. Biochem. 44: 435-449.

Mitchum, M.G., Yamaguchi, S., Hanada, A., Kuwahara, A., Yoshioka, Y., Kato, T., Tabata, S., Kamiya, Y., and Sun, T.p. (2006). Distinct and overlapping roles of two gibberellin 3-oxidases in Arabidopsis development. The Plant Journal. 45: 804-818.

Movahedi, S., Van Bel, M., Heyndrickx, K.S., and Vandepoele, K. (2012). Comparative co-expression analysis in plant biology. Plant Cell Environ. 35: 1787-1798.
Movahedi, S., Van de Peer, Y., and Vandepoele, K. (2011). Comparative network analysis reveals that tissue specificity and gene function are important factors influencing the mode of expression evolution in Arabidopsis and rice. Plant Physiol. 156: 1316-1330.

Mustroph, A., Lee, S.C., Oosumi, T., Zanetti, M.E., Yang, H., Ma, K., Yaghoubi-Masih, A., Fukao, T., and Bailey-Serres, J. (2010). Cross-kingdom comparison of transcriptomic adjustments to low-oxygen stress highlights conserved and plant-specific responses. Plant Physiol. 152: 1484-1500.

Naouar, N., Vandepoele, K., Lammens, T., Casneuf, T., Zeller, G., Van Hummelen, P., Weigel, D., Rätsch, G., Inzé, D., Kuiper, M., et al. (2009). Quantitative RNA expression analysis with Affymetrix Tiling 1.0R arrays identifies new E2F target genes. Plant J. 57: 184-194.

Narsai, R., Rocha, M., Gegenberger, P., Whelan, J., and van Dongen, J.T. (2011). Comparative analysis between plant species of transcriptional and metabolic responses to hypoxia. New Phytologist. 190: 472-487.

Nelissen, H., Eeckhout, D., Demuyynck, K., Persiau, G., Walton, A., van Bel, M., Vervoort, M., Candaele, J., De Block, J., Aesaert, S., et al. (2015). Dynamic changes in ANGUSTIFOLIA3 complex composition reveal a growth regulatory mechanism in the maize leaf. Plant Cell. 27: 1605-1619.

Nelissen, H., Gonzalez, N., and Inzé, D. (2016). Leaf growth in dicots and monocots: so different yet so alike. Current Opinion in Plant Biology. 33: 72-76.

Nelissen, H., Rymen, B., Jikumaru, Y., Demuyynck, K., Van Lijssebettens, M., Kamiya, Y., Inzé, D., and Beemster, G.T. (2012). A local maximum in gibberellin levels regulates maize leaf growth by spatial control of cell division. Current Biology. 22: 1183-1187.

Nelissen, H., Sun, X.H., Rymen, B., Jikumaru, Y., Kojima, M., Takebayashi, Y., Abbeloos, R., Demuyynck, K., Storme, V., Vuytsteke, M., et al. (2018). The reduction in maize leaf growth under mild drought affects the transition between cell division and cell expansion and cannot be restored by elevated gibberellic acid levels. Plant Biotechnol. J. 16.

Nelson, T., and Dengler, N. (1997). Leaf vascular pattern formation. The Plant Cell. 9: 1121.

Nicolas, M., and Cubas, P. (2016). The role of TCP transcription factors in shaping flower structure, leaf morphology, and plant architecture. In Plant transcription factors: Elsevier), pp. 249-267.

Okahama-Ohtsu, N., Radwan, S., Peterson, A., Zhao, P., Badr, A.F., Xiang, C., and Oliver, D.J. (2007). Characterization of the extracellular γ-glutamyl transpeptidases, GGT1 and GGT2, in Arabidopsis. The Plant Journal. 49: 865-877.

Patel, R.V., Nahal, H.K., Brett, R., and Provant, N.J. (2012). BAR expressolog identification: expression profile similarity ranking of homologous genes in plant species. Plant J. 71: 1038-1050.

Peng, L., Ma, J., Chi, W., Guo, J., Zhu, S., Lu, Q., Lu, C., and Zhang, L. (2006). LOW PSII ACCUMULATION1 is involved in efficient assembly of photosystem II in Arabidopsis thaliana. The Plant Cell. 18: 955-969.

Proost, S., Van Bel, M., Vaneechoutte, D., Van de Peer, Y., Inzé, D., Mueller-Roeber, B., and Vandepoele, K. (2015). PLAZA 3.0: an access point for plant comparative genomics Nucleic Acids Res. 43: D974-D981.

Pyke, K.A., Marrison, J.L., and Leech, A.M. (1991). Temporal and spatial development of the cells of the expanding first leaf of Arabidopsis thaliana (L.) Heynh. Journal of Experimental Botany. 42: 1407-1416.

R Core Team (2017). R: A language and environment for statistical computing. Foundation for Statistical Computing, Vienna, Austria (http://www.R-project.org/).

Raynaud, C., Sozzani, R., Glab, N., Domenichini, S., Perennes, C., Cella, R., Kondorosi, E., and Bergounioux, C. (2006). Two cell-cycle regulated SET-domain proteins interact with proliferating cell nuclear antigen (PCNA) in Arabidopsis. Plant J. 47: 395-407.

Rensink, W.A., Lee, Y., Liu, J., Iobst, S., Ouyang, S., and Buell, C.R. (2005). Comparative analyses of six solanaceous transcriptomes reveal a high degree of sequence conservation and species-specific transcripts. BMC Genomics. 6: 124.

Sánchez-Rodríguez, C., Bauer, S., Hématy, K., Saxe, F., Ibáñez, A.B., Vodermaier, V., Konlechner, C., Sampathkumar, A., Rüggeberg, M., and Aichinger, E. (2012). Chitinase-like1/pom1 and its homolog CTL2 are glucan-interacting proteins important for cellulose biosynthesis in Arabidopsis. The Plant Cell. 24: 589-607.

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

Šimková, K., Moreau, F., Pulk, P., Vriet, C., Baruah, A., Alexandre, C., Hennig, L., Apel, K., and Laloi, C. (2012). Integration of stress-related and reactive oxygen species-mediated signals by Topoisomerase VI in Arabidopsis thaliana. PNAS. 109: 16360-16365.
Skirycz, A., Claeyss, H., De Bodt, S., Oikawa, A., Shinoda, S., Andriankaja, M., Maleux, K., Eloy, N.B., Coppens, F., Yoo, S.-D., et al. (2011). Pause-and-stop: the effects of osmotic stress on cell proliferation during early leaf development in Arabidopsis and a role for ethylene signaling in cell cycle arrest. Plant Cell. 23: 1876-1888.

Spartz, A.K., Ren, H., Park, M.Y., Grandt, K.N., Lee, S.H., Murphy, A.S., Sussman, M.R., Overvoorde, P.J., and Gray, W.M. (2014). SAUR inhibition of PP2C-D phosphatases activates plasma membrane H+-ATPases to promote cell expansion in Arabidopsis. Plant Cell. 26: 2129-2142.

Stahle, M.I., Kuehlich, J., Staron, L., Arnim, A.G.V., and Golz, J.F. (2009). YABBYs and the Transcriptional Corepressors LEUNIG and LEUNIG_HOMOLOG Maintain Leaf Polarity and Meristem Activity in Arabidopsis. The Plant Cell. 21: 3105-3118.

Storozhenko, S., Belles-Boix, E., Babiychuk, E., Hérouart, D., Davey, M.W., Slooten, L., Van Montagu, M., Inzé, D., and Kushnir, S. (2002). γ-Glutamyl transpeptidase in transgenic tobacco plants. Cellular localization, processing, and biochemical properties. Plant Physiology. 128: 1109-1119.

Ueda, M., Matsu, K., Ishiguro, S., Sano, R., Wada, T., Paponov, I., Palme, K., and Okada, K. (2004). The HALTED ROOT gene encoding the 26S proteasome subunit RPT2a is essential for the maintenance of Arabidopsis meristems. Development. 131: 2101-2111.

Uematsu, K., Suzuki, N., Iwamae, T., Inui, M., and Yukawa, H. (2012). Increased fructose 1, 6-bisphosphate aldolase in plastids enhances growth and photosynthesis of tobacco plants. Journal of Experimental Botany. 63: 3001-3009.

Van Bel, M., Proost, S., Wischnitzki, E., Movahedi, S., Scheerlinck, C., Van de Peer, Y., and Vandenpoel, K. (2012). Dissecting plant genomes with the PLAZA comparative genomics platform. Plant Physiol. 158: 590-600.

Van Dingenen, J., Blomme, J., Gonzalez, N., and Inzé, D. (2016). Plants grow with a little help from their organelle friends. J. Exp. Bot. 67: 6267-6281.

Vandenpoel, K., Vliege, K., Florquin, K., Hennig, L., Beemster, G.T.S., Gruissem, W., Van de Peer, Y., Inzé, D., and De Veylder, L. (2005). Genome-wide identification of potential plant E2F target genes. Plant Physiol. 139: 316-328.

Vercruysse, L., Verkest, A., Gonzalez, N., Heyndrickx, K.S., Eeckhout, D., Han, S.-K., Jégu, T., Archacki, R., Van Leene, J., Andriankaja, M., et al. (2014). ANGUSTIFOLIA3 binds to SWI/SNF chromatin remodeling complexes to regulate transcription during Arabidopsis leaf development. Plant Cell. 26: 210-229.

Verkest, A., Abeel, T., Heyndrickx, K.S., Van Leene, J., Lanz, C., Van De Slijke, E., De Winne, N., Eekhout, D., Persiau, G., Van Breusegem, F., et al. (2014). A generic tool for transcription factor target gene discovery in Arabidopsis cell suspension cultures based on tandem chromatin affinity purification. Plant Physiol. 164: 1122-1133.

Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K., et al. (2014). Determination and inference of eukaryotic transcription factor sequence specificity. Cell. 158: 1431-1443.

Wickham, H. (2016). ggplot2: elegant graphics for data analysis. Second edition. Switzerland: Springer.

Wittstock, U., and Burow, M. (2010). Glucosinolate breakdown in Arabidopsis: mechanism, regulation and biological significance. The Arabidopsis book/American Society of Plant Biologists. 8.

Xia, C., Wang, Y.J., Li, W.Q., Chen, Y.R., Deng, Y., Zhang, X.Q., Chen, L.Q., and Ye, D. (2010). The Arabidopsis eukaryotic translation initiation factor 3, subunit F (AteIF3f), is required for pollen germination and embryogenesis. The Plant Journal. 63: 189-202.

Xiao, J., Li, J., Ouyang, M., Yun, T., He, B., Ji, D., Ma, J., Chi, W., Lu, C., and Zhang, L. (2012). DAC is involved in the accumulation of the cytochrome b6/f complex in Arabidopsis. Plant physiology. 160: 1911-1919.

Yoshizawa, E., Kaizuka, M., Yamagami, A., Higuchi-Takeuchi, M., Matsu, M., Kakei, Y., Shimada, Y., Sakuta, M., Osada, H., and Asami, T. (2014). BPG3 is a novel chloroplast protein that involves the greening of leaves and related to brassinosteroid signaling. Bioscience, biotechnology, and biochemistry. 78: 420-429.

Yu, C.-P., Chen, S.C.-C., Chang, Y.-M., Liu, W.-Y., Lin, H.-H., Lin, J.-J., Chen, H.J., Lu, Y.-J., Wu, Y.-H., Lu, M.-Y.J., et al. (2015). Transcriptome dynamics of developing maize leaves and genomewide prediction of cis elements and their cognate transcription factors. PNAS. 112: E2477-2486.

Zhang, D.-F., Li, B., Jia, G.-Q., Zhang, T.-F., Dai, J.-R., Li, J.-S., and Wang, S.-C. (2008). Isolation and characterization of genes encoding GRF transcription factors and GIF transcriptional coactivators in Maize (Zea mays L.). Plant Sci. 175: 809-817.
FIGURES

Figure 1 | Overview of the properties of the orthologous network constructed between Arabidopsis and maize using the integrative orthology method in PLAZA. (a) Visual representation of Arabidopsis GROWTH REGULATING FACTOR 3 (GRF3) (diamond) and its orthologs in maize. Three maize orthologs (upper three triangles) are confirmed by multispecies Best-Hits-and-Inparalogs families (B) and reconciled phylogenetic trees (T) and twelve additional orthologs (lower 12 triangles) are predicted by T. Arabidopsis genes are represented by diamonds, maize genes by triangles. (b) and (c) Venn diagrams depicting the number and the percentage of Arabidopsis (b) and maize (c) genes sharing orthologs with the other species through the three evidence sources in PLAZA. (d) Pie chart showing the percentage of orthologous groups with one-to-one, one Arabidopsis gene with many maize orthologs (One-To-Many Arabidopsis), one maize gene with many Arabidopsis orthologs (One-To-Many Maize), or many-to-many orthologous links.
Figure 2 | Clustered expression profiles of DE genes in Arabidopsis and maize during leaf development together with the orthology overlap between clusters. The six selected Arabidopsis clusters are depicted horizontally (A1-A6, diamonds) and the seven selected maize clusters are ordered vertically (Z1-Z7, triangles), all of them sorted according to their visual trend. Clusters A1, A2, and Z1 have high expression in the early stages, corresponding to the proliferation phase, and a decreasing expression as leaf development advances (red). Clusters A3, A4, Z2, Z3, and Z4 have high expression during transition and low expression at the beginning and later stages of leaf development (blue). Clusters A5, A6, Z5, and Z6 have an upward gene expression trend coinciding with the onset of the expansion phase (green). Cluster Z7 has high expression during proliferation and expansion, and low expression during the transition phase (yellow). The total number of genes in each cluster and the number of DE genes with an DE ortholog are shown under or right of the cluster name for Arabidopsis or maize respectively. For each comparison between two clusters, the overlap is shown both for the Arabidopsis (upper number) and maize genes (lower number). Eight significant overlaps (p<.001) are shown in dark grey and two significant overlaps (p<0.01) are shown in light grey. At the bottom right, two representative clusters of the Arabidopsis DE gene set (left) and the maize DE gene set (right) are shown. Normalized gene expression is depicted on the x-axis according to the differential expression. Points on the y-axis which correspond to 0-13DAS for the Arabidopsis dataset and to distance from the leaf base for the maize dataset. Ath. Arabidopsis thaliana, Zma. Zea mays, Nr. number, sp. species, DE differentially expressed, DAS Days after stratification.
Figure 3 I Heatmap showing the top 10 enriched GO terms (BP) per overlap with enrichment in at least one overlap for the four significant overlaps (p<0.001) with conserved expression showing a peak in expression during cell proliferation (red) and for the one significant overlap (p<0.001) with conserved expression with an expression peak during cell expansion (green). BP, biological process.
Figure 4 | Overview of the DE genes from four TF families, YABBY, GRF, E2F, and NAC, showing conserved expression patterns in both Arabidopsis and maize. In each panel, the top image represents the integrated orthology network in which Arabidopsis TFs (diamonds) and maize orthologs (triangles) are linked. The color of the nodes corresponds to an increasing (green) or a decreasing (red) expression trend during leaf development. The thickness of the edge corresponds to the number of evidences found for that particular orthology link (B: Best-Hits-and-Paralogs families, O: OrthoMCL, and T: Reconciled Phylogenetic Trees). The color of the edges represents a significant orthology overlap between the respective Arabidopsis and maize genes. The heatmap in each panel represents the normalized expression values of each DE TF throughout leaf development. *At* Arabidopsis thaliana, *zma* Zea mays, DE differentially expressed.
Figure 5 | Overview of the conserved expression among TCP TFs, their orthologs, and their putative targets. a Overview of the Arabidopsis (diamonds) class I and class II TCP DE genes with their DE maize (triangles) orthologs during leaf growth. The heat-map represents the normalized expression values of each DE TF throughout leaf development in both species. b Visual representation of eight experimentally verified TFBSs corresponding to eight TCP TFs, three class I TFs and five class II TFs. The TFBSs of the underlined TFs TCP3 and TCP24, were found enriched in promoters of Arabidopsis and maize gene clusters (Table S8). Four different colors represent the four bases (yellow: G, red: T, blue: C, and green: A) and the relative size of the bases corresponds to the frequency of occurrence. c Visual representation of the 34 Arabidopsis and 34 maize putative TCP3 and/or TCP24 targets that were identified based on the TFBS enrichment analysis. Genes (nodes) are colored according to their expression trend: increasing profile (green) or decreasing profile (red) during leaf development, while color of the node edge shows the gene to be a target of TCP3 (blue), TCP24 (yellow), or both (purple). The thickness of the edge corresponds to the number of evidences found for that particular orthology link (B: Best-Hits-and-Inparalogs families, O: OrthoMCL, and T: Reconciled Phylogenetic Trees). The color of the edges represents a significant orthology overlap between the respective Arabidopsis and maize genes. Ath. Arabidopsis thaliana, Zma. Zea mays, DE differentially expressed.
SUPPLEMENTAL INFORMATION

Figure S1 | Visual representation of the two datasets and overview of the hierarchical clustering of the Arabidopsis and maize DE genes. Overview of the sampling strategy leading to the generating of the maize (top image) and Arabidopsis (lower image) transcriptomics dataset. Leaves in Arabidopsis or zones of the maize leaf that are proliferating are colored red, while leaves or leaf zones are colored green when expanding. Leaves or leaf zones in transition have a red-green gradient. DAS Days after stratification (a), b, c Dendrograms for the Arabidopsis (b) and maize (c) representing the hierarchical clustering. d, e Hierarchical clustering of the Arabidopsis differentially expressed genes (a) from two until eleven clusters and hierarchical clustering of the maize differentially expressed genes (e) from two until eight clusters. The cluster numbers are depicted above each cluster image. DE differentially expressed.
Figure S2.1 Heatmap showing the top 10 enriched GO terms (BP and MF) per cluster with enrichment in at least one cluster. The colors under the cluster names represent the heatmap trend: high expression during proliferation (red), high expression during expansion and transition (blue), high expression during expression (green), and high expression during proliferation and expansion (yellow). BP: biological process, MF: molecular function.
Figure S3 | Visual representation of the Arabidopsis class I (left) and class II (right) TCP gene family with the maize orthologs. Among the class I TCP TFS, only two DE Arabidopsis genes (diamonds: TCP11 and TCP19) have three DE maize orthologs (triangles: ZmTCP4, ZmTCP19 and ZmTCP21). In class II, DE Arabidopsis genes TCP3 and TCP4 share one DE maize ortholog, ZmTCP23. Grey nodes represent genes not DE, while white nodes represent genes not included in the DEG analysis. DE differentially expressed, DEG differentially expressed gene.
Figure S4 | Schematic overview of the identified putative E2F targets benchmarking. The E2F TFRS, TTTCCCGC, was found to be enriched in both A4 and Z3 clusters that share 12 Arabidopsis and 16 maize genes and include genes with a decreasing expression during leaf development. Ath. Arabidopsis thaliana, Zma. Zea mays.
SUPPLEMENTAL TABLE LEGENDS

Table S1 | Overview of the 217 one-to-one orthology groups with conserved expression and part of one of the significant overlaps (A1-Z2, A5-Z5, A5-Z6, A6-Z5, A6-Z6). Unknown genes in both Arabidopsis and maize were found in 41 groups, while 28 groups contained known Arabidopsis and unknown maize genes. For the rest of the groups, genes from both species were annotated. NA Not Available.

Table S2 | Calculation of percentage, enrichment, and FDR of leaf-related phenotype data for the groups with conserved or divergent expression, and the Arabidopsis group with no orthologs of clusters A1, A5, and A6. DE differentially expressed, Nr Number, HGS hypergeometric score, FDR false discovery rate, TFs transcription factors, OE overexpressed, KO knocked out.

Table S3 | Group, ID, name, TF information, and phenotype description resulting from a knockout (KO) or an overexpression (OE) described in TAIR and or RARGE and confirmed by a publication for 49 genes, 10 genes, and 48 genes possibly involved in a decrease or increase in leaf size or another leaf-related phenotype, respectively. A1CON, conserved group in cluster A1; A1DIV, divergent group in cluster A1; A1NO, no orthologs group in cluster A1. Likewise, A5CON, A5DIV, A5NO, A6CON, A6DIV, and A6NO are defined.

Table S4 | Overview of the 59 one-to-two orthology groups exemplifying possible neo- or subfunctionalization of the maize ortholog having divergent expression.

Table S5 | Overview of 211 DE Arabidopsis TFs and their 256 DE maize orthologs either with conserved or divergent expression throughout leaf development. For each entry: TF family, ID, gene name, expression trend during leaf growth, cluster number, maize ortholog with ID, gene name expression trend, and cluster number, and conservation or divergence in expression between Arabidopsis and maize ortholog.

Table S6 | Overview of the DE TF families including Arabidopsis genes with their maize orthologs that are conserved, partially conserved, or divergent in expression throughout leaf development. Each entry corresponds to a different TF family with the number of DE
Arabidopsis and maize genes, the number of orthology relations between them, and the percentage of expression conservation.

Table S7 | Overview of the identified E2F targets with conserved expression in Arabidopsis (12 genes) and maize (16 genes). For each Arabidopsis gene: gene ID, name, cluster ID, expression trend during leaf development, maize orthologs with cluster ID and expression trend during leaf development are shown with the type and number of evidences for the orthology link. O sequence-based clustering with OrthoMCL, T phylogenetic trees, B multispecies best hits and inparalogs families.

Table S8 | Overview of the Arabidopsis and maize clusters in which TCP3 and TCP24 TF binding sites are enriched based on the hypergeometric score (HGS). Each entry corresponds to either an Arabidopsis or a maize cluster in which the TCP3 or TCP24 TFBSs are enriched.

Supplemental network file with the full integrated network as well as the case-studies handled in this manuscript.