Insights into Populus XIP aquaporins: evolutionary expansion, protein functionality, and environmental regulation.

David Lopez, Gisèle Bronner, Nicole Brunel, Daniel Auguin, Sylvain Bourgerie, Franck Brignolas, Sabine Carpin, Colette Tournaire-Roux, Christophe Maurel, Boris Fumanal, et al.

Journal of Experimental Botany, Vol. 63, No. 5, pp. 2217–2230, 2012
doi:10.1093/jxb/err404 Advance Access publication 5 January, 2012

Journal of Experimental Botany
www.jxb.oxfordjournals.org

RESEARCH PAPER

Insights into *Populus* XIP aquaporins: evolutionary expansion, protein functionality, and environmental regulation

D. Lopez[1,2], G. Bronner[3], N. Brunel[1,2], D. Auguin[4,5], S. Bourgerie[4,5], F. Brignolas[4,5], S. Carpin[4,5], C. Tournaire-Roux[6], C. Maurel[6], B. Fumanal[1,2], F. Martin[7], S. Sakr[8], P. Label[5], J.L. Julien[1,2], A. Gousset-Dupont[1,2] and J.S. Venisse[1,2,*]

[1] Université Blaise Pascal UMRA 547 PIAF, BP 10448, F-63000 Clermont-Ferrand, France
[2] INRA, UMRA 547 PIAF, F-63100 Clermont-Ferrand, France
[3] Université Blaise Pascal, UMR LMGE, BP 10448, F-63000 Clermont-Ferrand, France
[4] Université d'Orléans, Laboratoire de Biologie des Ligneux et des Grandes Cultures, UPRES EA 1207, F-45067 Orléans, France
[5] INRA, USC1328 Arbres et Réponses aux Contraintes Hydriques et Environnementales (ARCHE), BP 6759, F-45067, France
[6] Biochimie et Physiologie Moléculaire des Plantes, Institut de Biologie Intégrative des Plantes, UMR 5004 CNRS/UMR 0386 INRA/Montpellier SupAgro/Université Monptellier 2, F-34060 Montpellier Cedex2, France
[7] Unité Mixte de Recherche 1136, Institut National de la Recherche Agronomique/Nancy Université, Interactions Arbres/Micro-organismes, Centre de Nancy, 54280 Champenoux, France
[8] Agrocampus Ouest, Centre d'Angers, UMR SAGAH, IFR QUASAV 149, 2 rue le Nôtre, F-49045 Angers, France

* To whom correspondence should be addressed. E-mail: j-stephane.venisse@univ-bpclermont.fr

Received 1 September 2011; Revised 14 November 2011; Accepted 16 November 2011

Abstract

A novel category of major intrinsic proteins which share weak similarities with previously identified aquaporin subfamilies was recently identified in land plants, and named X (for unrecognized) intrinsic proteins (XIPs). Because XIPs are still ranked as uncharacterized proteins, their further molecular characterization is required. Herein, a systematic fine-scale analysis of XIP sequences found in flowering plant databases revealed that XIPs are found in at least five groups. The phylogenetic relationship of these five groups with the phylogenetic organization of angiosperms revealed an original pattern of evolution for the XIP subfamily through distinct angiosperm taxon-specific clades. Of all flowering plant having XIPs, the genus *Populus* encompasses the broadest panel and the highest polymorphism of XIP isoforms, with nine *Pt*XIP sequences distributed within three XIP groups. Comprehensive *PtXIP* gene expression patterns showed that only two isoforms (*PtXIP2;1* and *PtXIP3;2*) were transcribed in vegetative tissues. However, their patterns are contrasted, *PtXIP2;1* was ubiquitously accumulated whereas *PtXIP3;2* was predominantly detected in wood and to a lesser extent in roots. Furthermore, only *Pt*XIP2;1 exhibited a differential expression in leaves and stems of drought-, salicylic acid-, or wounding-challenged plants. Unexpectedly, the *Pt*XIPs displayed different abilities to alter water transport upon expression in *Xenopus laevis* oocytes. *Pt*XIP2;1 and *Pt*XIP3;3 transported water while other PtXIPs did not.

Key words: Aquaporin, evolution, *in situ* hybridization, molecular physiology, *Populus*, XIP.

Introduction

Water flux across biological membranes can occur directly through the lipid bilayer, but predominantly happens through water channel proteins named aquaporins (AQPs) (Agre *et al.*, 1993). AQPs are integral membrane proteins of the larger major intrinsic protein (MIP) superfamily. These channels are found in all living organisms, from archea and eubacteria to fungi, animals, and plants (Engel and Stahlberg, 2002). Looking at this global distribution, AQPs are

especially abundant in plants, with expression in virtually all cell types (Kjellbom *et al.*, 1999; Wallace and Roberts, 2004). AQPs constitute a large and divergent family with 35 members identified in *Arabidopsis thaliana* (Johanson *et al.*, 2001), 36 in *Zea mays* (Chaumont *et al.*, 2001), 33 in *Oryza sativa* (Sakurai *et al.*, 2005), 37 in *Solanum lycopersicon* (Sade *et al.*, 2009), and 65 members in *Populus trichocarpa* (Gupta and Sankararamakrishnan, 2009). Some AQPs form channels facilitating bidirectional water fluxes, whereas others can conduct a wide range of small neutral solutes including urea, lactic acid, glycerol, hydrogen peroxide, silicic acid, metalloids, and gases such as ammonia and carbon dioxide. AQPs play a central role in plant–soil water relations, seed germination, cell elongation, drought resistance, salt tolerance, and fruit ripening. Based on sequence similarities, AQPs of most plant species can typically be divided into four subfamilies: the plasma membrane intrinsic proteins (PIPs), the tonoplast intrinsic proteins (TIPs), the nodulin 26-like intrinsic proteins (NIPs), and the small basic intrinsic proteins (SIPs). Most of these subfamilies are being extensively investigated with respect to their structural signatures, expression patterns, subcellular localizations, and substrate specificities. Phosphorylation, pH, Ca^{2+}, and osmotic gradients were also reported to affect their water channel activities. Moreover, recent insights into their co- and post-translational modifications and cellular trafficking have led to promising models that explain the regulation of their transport activity in various physiological processes, including photosynthesis, osmotic regulation, tree embolism recovery, and water use efficiency. Recent reviews have summarized increasing knowledge about these processes (Chaumont *et al.*, 2005; Kaldenhoff and Fischer, 2006; Maurel, 2007; Maurel *et al.*, 2009).

Three additional AQP subfamilies were recently described in the non-vascular moss *Physcomitrella patens*: GlpF-like intrinsic proteins (GIPs) homologous to some glycerol channels of Gram-positive bacteria (Gustavson *et al.*, 2005), the hybrid intrinsic proteins (HIPs), and uncategorized members designated X intrinsic proteins (XIPs) (Danielson and Johanson, 2008). Members of the latter subfamily are found in protozoa, fungi, and land plant kingdoms. Since XIPs are not encoded by the sequenced genomes of certain plants such as *Arabidopsis*, *Oryza* or *Zea*, partial knowledge together with hypothetical functions of plant XIPs are starting to emerge (Danielson and Johanson, 2008; Sade *et al.*, 2009; Gupta and Sankararamakrishnan, 2009; Shelden *et al.*, 2009; Park *et al.*, 2010; Bienert *et al.*, 2011).

Plant XIP sequences have substantially diverged from those of fungi and protozoa. Strikingly, such sequence divergence is encountered even within the plant kingdom itself, as observed when comparing the Lycopodiophyta *Selaginella moellendorffii*, the Bryophyta *Physcomitrella patens*, and various Magnoliophyta (angiosperms) XIP sequences. Within this last plant division, the XIP subclass seems to be exclusive to *Magnoliopsida* (dicots) and its members have been phylogenetically clustered into two groups (XIP1 and XIP2). As these proteins harbour amino acid variations in regions corresponding to the primary selection filter and the major checkpoint for solute permeability [i.e. the first Asn-Pro-Ala (NPA) box and aromatic/arginine region (ar/R)] (Beitz *et al.*, 2006; Tornroth-Horsefield *et al.*, 2006; Mitani-Ueno *et al.*, 2011), it was first hypothesized that XIPs are non-functional as water channels but instead are dedicated to the transport of hydrophobic solutes (Danielson and Johanson, 2008; Gupta and Sankararamakrishnan, 2009). These predictions have been recently validated by Bienert and his collaborators (2011) on XIPs from three Solanales: *Nicotiana tabacum*, *Solanum lycopersicon*, and *Solanum tuberosum*.

Although substantial progress has been made regarding the XIP molecular structure, subcellular localization, transcriptional regulation, and functional gating, the reason for their availability in plants is not understood, and several questions remain to be answered. For instance, (i) to what extent did the evolutionary expansion of XIP occur in plants? The evolutionary history of the XIPs was partially explained in a broad phylogenetic survey including some sequences of flowering plants (angiosperms), protozoa, fungi, and mosses (Gupta and Sankararamakrishnan, 2009). This was paradoxically done without considering the angiosperm evolutionary story and this Viridiplantae division encompasses the largest number of XIP sequences. It follows that previous XIP phylogenetic data alone were not conclusive. (ii) Can some XIPs different from those studied in Solanales be considered as multifunctional channels, being able to transport water? Indeed, many members in each MIP subclass present differential selectivity for substrates, and this is particularly true in the case of plants which comprise the most remarkable MIP diversity.

In the study, XIP sequences from a wide variety of plants were first retrieved and compared, making full use of the available expressed sequence tag (EST) as well as plant genome sequences databases. This data set was used to analyse their phylogenetic relationship, together with an integrative amino acid-based evolutionary partitioning of this subfamily during angiosperm evolution. Among this set, poplar, with nine sequences, is to date the species with the largest number of XIPs with significant amino acid diversity, making it a suitable model to re-examine this AQPs subfamily molecularly. Expression analyses using quantitative real-time PCR revealed that only two *PtXIP* genes (*PtXIP2;1* and *PtXIP3;2*) were expressed in vegetative tissues, and only *PtXIP2;1* was differentially and strongly modulated in response to abiotic stresses. Supplementary *in situ* hybridization experiments showed that the *PtXIP2;1* and *PtXIP3;2* genes were highly expressed in well-defined cell types or poplar tissues. Finally, *Pt*XIP2;1 and *Pt*XIP3;3 were functionally characterized as being the poplar XIPs able to transport water in *Xenopus* oocyte expression assays.

Materials and methods

Plant material and experimental design

Populus trichocarpa (Torr. & Gray ex Hook), clone INRA 101-74 kindly provided by Dr Catherine Bastien (INRA, Orléans, France), was selected for the experiments as it is increasingly being

used as the male parent for genetic research programmes. Homogeneous 25 cm long woody stem cuttings were planted in 10.0 l pots filled with a commercial substrate (40% black, 30% brown, and 30% blond peat moss, pH 6.1, DUMONA-RN 75-3851 Arandon, The Netherlands) and placed in a controlled-environment greenhouse under a 16 h/8 h light/dark photoperiod, at 18/22 °C (day/night), with relative humidity set at 70±10% and a daytime photosynthetic photon flux ⩾350 µmol m^{-2} s^{-1}. Three treatments were conducted on 3-month-old plants. Phytohormone treatment was carried out on fully expanded leaves with 1 mM salicylic acid (SA) dissolved in 100 µl of absolute ethanol before concentration adjustment in distilled water and then pulverization. Drought conditions were produced by withholding water until the first basal leaves showed signs of severe dehydration (soil water potential $\Psi_s \approx -1.2 \pm 0.15$ MPa). Rehydration conditions were obtained by re-watering the soil. For mechanical wounding treatments, 40% of the leaf area was lightly crushed using large clamps. At each sampling time, harvested tissues were immediately frozen in liquid nitrogen and stored at –80 °C until RNA extraction.

Bioinformatics analysis

The plant XIP gene subfamily was explored using XIP sequences previously described (Danielson and Johanson, 2008; Gupta and Sankararamakrishnan, 2009) as initial queries. Basic Local Alignment Search Tool (BLAST) search tools, tBLASTn and BLASTp (Altschul *et al.*, 1997), were used against Viridiplantae taxon non-redundant generalist databases National Center for Biotechnology information (NCBI, http://www.ncbi.nlm.nih.gov/) and Department of Energy's Joint Genome Institute (JGI, http://www.jgi.doe.gov/), and from plant-dedicated assembly genome projects: Coffee Genome Project (http://www.lge.ibi.unicamp.br/cafe), Genome Database for Rosacae (http://www.rosaceae.org), Genoscope (http://www.genoscope.cns.fr), Public *Eucalyptus* Genome resource (http://eucalyptusdb.bi.up.ac.za), and Sol Genomics Network (http://solgenomics.net). Incomplete sequences showing conserved regions were manually assembled for further analysis, whereas redundant entries, including sequences with putative point mutations or polymorphism (similarities >98%), and single incomplete sequences displaying a hypothetical length <80% compared with its putative poplar orthologue were discarded.

The poplar MIP gene family was uncovered via two bioinformatics explorations: *P. trichocarpa* whole genome shotgun (WGS) sequences expanded with GenBank *Populus* EST collections and full-length cDNAs, and the recent release of the *P. trichocarpa* genome from JGI database assembly version 2.2 [JGI v2.2 annotation of the v2.2 assembly (JGIv2.2) available via Phytozome version 7.0 (www.phytozome.net)] in comparison with the older JGI annotation of the v1.1 assembly (JGIv1.1). A complementary analysis was carried out on the Molecular Genbank non-redundant sequence databases using the tBLASTn search tool server in order to identify potential poplar ESTs related to each MIP sequence previously identified. This made it possible to redefine any sequence with false predictions. The nomenclature of Gupta and Sankararamakrishnan (2009) was used for naming *Populus* MIPs, except for the *Pt*XIP subfamily, for which an updated nomenclature was proposed based on the percentage of amino acid similarity between members and the phylogenetic studies presented herein.

Percentages of amino acid similarity and identity were calculated using the NCBI bl2seq algorithm. Multiple sequence alignment analysis was performed using PRALINE-PSI (Simossis and Heringa, 2005) with a five PSI-BLAST iteration search (with an *e*-value cut-off of 0.01) within the non-redundant database. Jalview (Clamp *et al.*, 2004) was used with the ClustalX colour code to illustrate the alignment results.

The unrooted phylogenetic trees were constructed using the maximum likelihood (ML) method implemented in the PhyML program v3.0 (Guindon and Gascuel, 2003). Specifically, PhyML analyses were conducted with the Jones–Taylor–Thornton (JTT) substitution matrix, 1000 bootstrap replicates to assess the reliability for degree of support for each internal branch on the phylogenetic trees, estimated proportion of invariable sites, four rate categories, and estimated gamma distribution parameter. Trees were viewed and edited with TreeDyn (Chevenet *et al.*, 1997), and bootstrap values >50% were reported. Except for *Pt*XIP3;1b (ProteinID 829126), all other related pseudogenes and truncated amino acid sequences were eliminated from further analysis.

Total RNA isolation, cDNA synthesis, and amplification

Total RNA was extracted from 150 mg of roots, the apical part of the stem, bark (including phloem) and wood from the intermediate part of the stem, buds, and leaves using cetyltrimethylammonium bromide (CTAB) extraction buffer according to Chang *et al.* (1993). First-strand cDNA was synthesized from 1 µg of total RNA using SuperScript III (Invitrogen, Carlsbad, CA, USA) following the manufacturer's instructions. Real-time PCR amplification was performed using a MyiQ thermocycler (Bio-Rad, Hercules, CA, USA) with MESA GREEN qPCR MasterMix Plus (Eurogentec, Belgium) containing 2 µl of 40-fold diluted cDNA according to the manufacturer's protocol. The thermal profile of the reaction was: 94 °C for 3 min and 40 cycles of 94 °C for 20 s, 54–58 °C for 20 s, and 72°C for 20 s. The specificity of amplicons was routinely verified by melting curve analysis, and checked by gel electrophoresis. PCR efficiencies were deduced for each gene according to Pfaffl's (2001) procedures. The crossing cycle number (C_t) was automatically determined for each reaction by the iCycler iQ v2.0 software with default parameters. For normalization of the target gene abundance, the software application *BestKeeper* v1 (http://www.gene-quantification.info; Pfaffl *et al.*, 2004) was used according to the developers' manual to determine the best suited reference genes from nine widely used housekeeping genes (Czechowski *et al.*, 2005; Xu *et al.*, 2011), and then to estimate a *BestKeeper Index*. The reference genes selected (gene locus; GM, geometric mean of C_t ±SD; CV, BestKeeper coefficient of Pearson correlation; and *r*, variance from all sample at all time points) were *Actin1* (POPTR_0001s31700; GM, 24.78±0.59; CV, 2.40; *r*, 0.864), *SAND* (POPTR_0009s01980; GM, 29.74±0.59; CV, 1.99; *r*, 0.870), and *TIP41a-like* (POPTR_0001s30580; GM, 27.12±0.77; CV, 2.66; *r*, 0.787), chosen from different protein families in order to reduce the possibility of co-regulation. Constitutive MIP transcript steady-state levels in each vegetative tissue were assessed by comparison of the mean of the C_t after normalization with the *BestKeeper Index*. The final results from each statistical output were graphically represented after distribution and assignation of a score value between 0 and 100, an arbitrary scoring range. A C_t of 20 was assigned an arbitrary value of 100 (corresponding to the highest MIP expression level in *Populus*), while a C_t of 40 was assigned an arbitrary value of 0 (no expressed gene). Then all other C_ts were assigned values between 0 and 100, scaled based on their relative distributions. As for differential accumulation of MIP transcripts in response to stresses, the relative changes were calculated with the equation $2^{-\Delta\Delta Ct}$ according to Pfaffl procedure's (2001). Circadian *PtXIP* expression under control conditions was first monitored allowing 't_0' untreated samples to be used as controls (corresponding to leaves sampled just before experiments). For legibility, and especially for down-regulated genes, values were graphically normalized to 0, this course representing no change in gene expression. Each unit on both sides of 0 corresponds to a 2-fold increase and a 2-fold decrease. The mean C_t value was determined from three independent biological replicates for each sample, and every PCR was carried out in triplicate. Values are given as means ±SD. Primers were designed using the Primer3plus application (http://www.bioinformatics.nl/primer3plus; Rozen and Skaletsky, 2000), and are listed in Supplementary Table S3 available at *JXB* online. A concomitant analysis of genomic DNA and cDNA was carried out in order to evaluate the robustness of each primer set (data not shown).

In situ hybridization

Fresh stems and leaves were harvested, cut, and immediately fixed in 4% (w/v) paraformaldehyde + 2.5% (v/v) glutaraldehyde overnight at 4 °C. Fixed samples were dehydrated and embedded in methacrylate resin (methyl methacrylate–butyl methacrylate resin, EMS, Mündelsheim, France) according to the manufacturer's instructions. Polymerization occurred in gelatin capsules overnight at 60 °C. Transverse sections of 3–4 µm thickness were cut with a rotary microtome, mounted on SuperFrost Plus slides (Fisher Scientific, Elancourt, France), and dried at 42 °C for 2 d. A 5 min incubation in pure acetone removed the methacrylate resin. *In situ* hybridization (ISH) was performed as described in Leblanc-Fournier *et al* (2008). Gene-specific RNA probes were designed to be located in the variable 3′-untranslated region (UTR) of *Pt*XIP2;1 and *Pt*XIP3;2 transcripts with an average size of 250 ribonucleotides (primers listed in Supplementary Table S3 at *JXB* online). DNAs encoding the probes were cloned in pGEM® T-Easy vector (Promega, Madison, WI, USA). Briefly, sense and antisense digoxigenin (DIG)-labelled RNA probes were synthesized as described in Leblanc-Fournier *et al* (2008) using an *in vitro* transcription kit [DIG RNA labeling kit (SP6/T7), Roche Diagnostics, Mannheim, Germany] according to the manufacturer's instructions. The antisense and sense probes were transcribed from SP6 or T7 polymerase promoters (after vector linearization with *Nco*I or *Spe*I, respectively). Stem and leaf sections were incubated overnight at 50 °C with 1.5 ng μl^{-1} or 3 ng μl^{-1}, respectively, of sense and antisense probes. They were then washed with 2× SSC/50% formamide. Detection was performed using anti-dioxigenin–alkaline phosphatase conjugate, followed by colorimetric detection of phosphatase activity (Bio-Rad). After suitable colour development, the reaction was stopped by rinsing in water, and sections were dried and mounted in Eukitt (Euromedex, Mündelsheim, France). Observations were performed under an Axioplan 2 microscope (Zeiss, Jena, Germany). Data were recorded on a digital camera (AxioCam HR, Zeiss) using Axiovision digital imaging software.

Expression in Xenopus laevis *oocytes and osmotic water transport assay*

The *PtXIP* cDNAs cloned in a pGEM®-T Easy vector were amplified by PCR (primers listed in Supplementary Table S3 at *JXB* online). Upstream of each *PtXIP* primer specific for the 5′-UTR and 3′-UTR, the *Bgl*II and *Spe*I restriction sites were introduced, respectively. The amplicons were digested by *Bgl*II and *Spe*I (Promega) and the resulting fragments were cloned in the corresponding sites of a T7Ts vector, fused with the 5′- and 3′-UTR sequences of a *Xenopus* γ-globin gene, to favour stability of the derived complementary RNA (cRNA). The integrity of the two constructs was checked by sequencing. Plasmids were linearized by *Eco*RI before *in vitro* transcription, and functional expression of aquaporins in *X. laevis* oocytes was as described by Maurel *et al.* (1993). The statistical analysis was performed on data pooled from all experiments using a one-way analysis of variance (ANOVA; $P < 0.001$), and post-hoc multiple comparisons were run using a Tukey test on the same set of data for which normality was checked. In the figures, the bars indicate the SEM and different letters denote statistically significant ($P < 0.001$) differences.

Results

Phylogenetic analysis of the angiosperm XIP group

An evolutionary lineage fine analysis of all available angiosperm XIP-like proteins was performed, based on the current angiosperm phylogeny. To accomplish this, all XIP sequences previously reported by Danielson and Johanson (2008) and by Gupta and Sankararamakrishnan (2009) were used as queries against various genomic and EST angiosperm collections. From slightly less than 200 XIP-related sequences retrieved, 55 non-redundant representative sequences were selected (Supplementary Table S1 at *JXB* online). As previously mentioned, except for *Liriodendron tulipefera* (Magnoliales), all XIP sequences came from the dicotyledonous phylum. In addition, an unrooted tree constructed from alignments of protein sequences confirmed that XIPs split into two independent clusters, identified as XIP-A and XIP-B in this study (Fig. 1; sequence alignments in Supplementary Fig. S3). These two subgroups correspond to XIP2 and XIP1, respectively, in the former nomenclature (Gupta and Sankararamakrishnan, 2009). A fine reading of the phylogenetic XIP distribution led to further refinement of the former nomenclature. The XIP-A cluster encompassed 15 orthologous sequences from mesangiosperms (Magnolianae superorder with *L. tulipefera*), and various core eudicots within the Rosanae taxon exclusively. These XIP-A sequences shared an average amino acid similarity of 77%, which fell to ~60% with XIP-B members. The second cluster, XIP-B, reflected the emergence of four major clades: XIP2, XIP3, XIP4, and XIP5 with well-supported to strong bootstrap support (82, 95, 99, and 95%, respectively). Interestingly, XIP-Bs were exclusive to eudicots and, unlike XIP-As, distinctly clustered into plant taxon-specific clades. Thus, XIP2 and XIP3 were spread into the Rosideae superorder: XIP2 encompassed sequences from Malpighiales (*Euphorbia*, *Gossypium*, *Manihot*, *Populus*, and *Ricinus*), Brassicales (*Carica*), and Rosales (*Malus* and *Prunus*), while XIP3 members were encountered in Malpighiales (*Manihot* and *Populus*) and Sapindales (*Citrus*). The XIP4 and XIP5 clades were, in contrast, exclusively linked to the Asteranae superorder: XIP4 members were linked to Lamiids (*Ipomea*, *Mimulus*, *Nicotiana*, *Solanum*, and *Triphysaria*), while XIP5 members were linked to Campanulids (*Centaurea*, *Cichorium*, *Helianthus*, *Guizotia*, and *Lactuca*). Members within each XIP-B clade shared a minimum of 85% amino acid sequence similarity, and an average between-clade similarity of 79%. Finally, as reported for the monophyletic *Liliopsida* taxon (Danielson and Johanson 2008), if the absence of XIP sequences is true and not an artefact due to limited data sets, plants belonging to ferns, ANITA (the basal angiosperms, Amborella, Nymphaeales, Illiciales, Trimeniaceae, and Austrobaileya), and coniferophyte monophyletic taxa lack the XIP subfamily.

A more integrative view of this phylogenetic XIP clade distribution (Fig. 2) was also revealed when considering the current angiosperm phylogeny (Angiosperm Phylogeny Website, V.9 http://www.mobot.org/MOBOT/research/APweb/, complemented with the Angiosperm Phylogeny Group III system, Chase and Reveal 2009). This analysis showed a XIP-A branch in which members diverged parallel to the plant species divergence, whereas the XIP-B branch would reflect clade divergences in a plant taxon-specific manner, both with their own evolutionary rate.

Features of the XIP *gene subfamily in* Populus

The molecular characterization of the *Populus XIP*-like genes was updated using the XIP protein sequences retrieved from

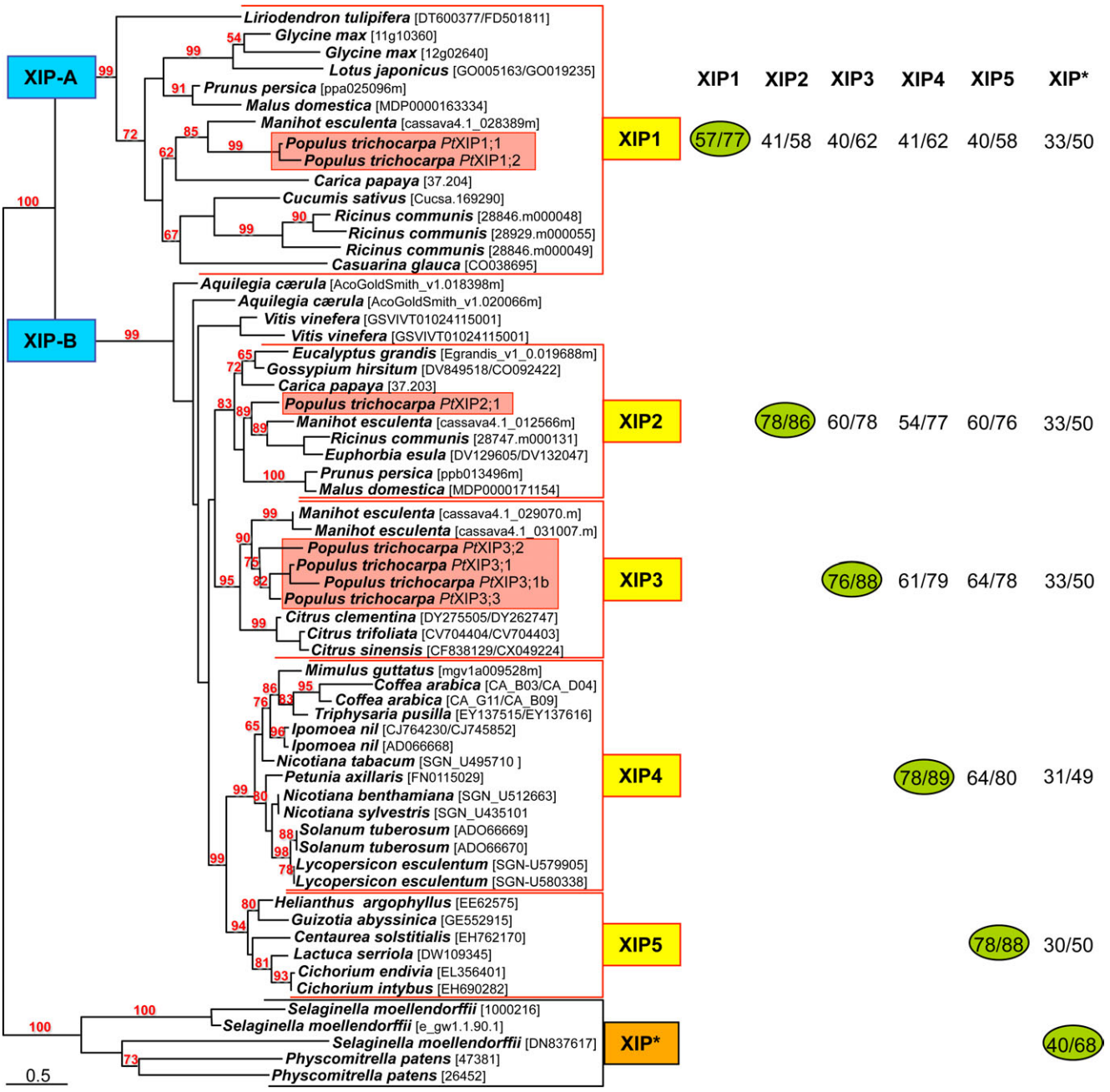


Fig. 1. Evolutionary relationship between Viridiplantae plant XIPs. The unrooted phylogeny of the XIP protein sequence was inferred using maximum likelihood. The tree was produced using PhyML with a genetic distance calculated by the JTT model of amino acid change. The numbers at the nodes represent the percentage bootstrap values (>50%) based on 1000 reassembling. The distance scale denotes the number of amino acid substitutions per site. Green circled numbers represent the percentage amino acid sequence identity and similarity intragroup; other numbers are amino acid sequence identity and similarity intergroups. *Populus* XIP sequences are highlighted in red. Species and accession numbers are listed in Supplementary Table S1 at *JXB* online. *Selaginella moellendorffii* and *Physcomitrella patens* XIP sequences (XIP*) were edited for information purposes, and were used as angiosperm taxon outgroups.

a reiterative search against the Phytozome Version 7.0. This upgraded version includes the last JGI v2.2 annotation of the *P. trichocarpa* assembly (JGIv2.2). The opportunity to retrieve the whole poplar *MIP* gene complement and to re-examine their sequences was also taken. Sixty-five sequences are by default annotated as putative '*Aquaporin (major intrinsic protein family)*', leading to 54 full-length open reading frames (ORFs) as initially reported (Gupta and Sankararamakrishnan, 2009, Almeida-Rodriguez *et al*, 2010). This research has led to the retrieval of two other sequences belonging to the XIP subfamily: *Pt*XIP1;2 (POPTR_0009s13105) and a pseudogene at the nucleotide position 10 525 867–10 526 066 located on the scaffold IX and not annotated. Sequences and final Phytozome v7.0 labelling are given in Supplementary Table S2 at *JXB* online. In agreement with previous works, the 56 ORFs

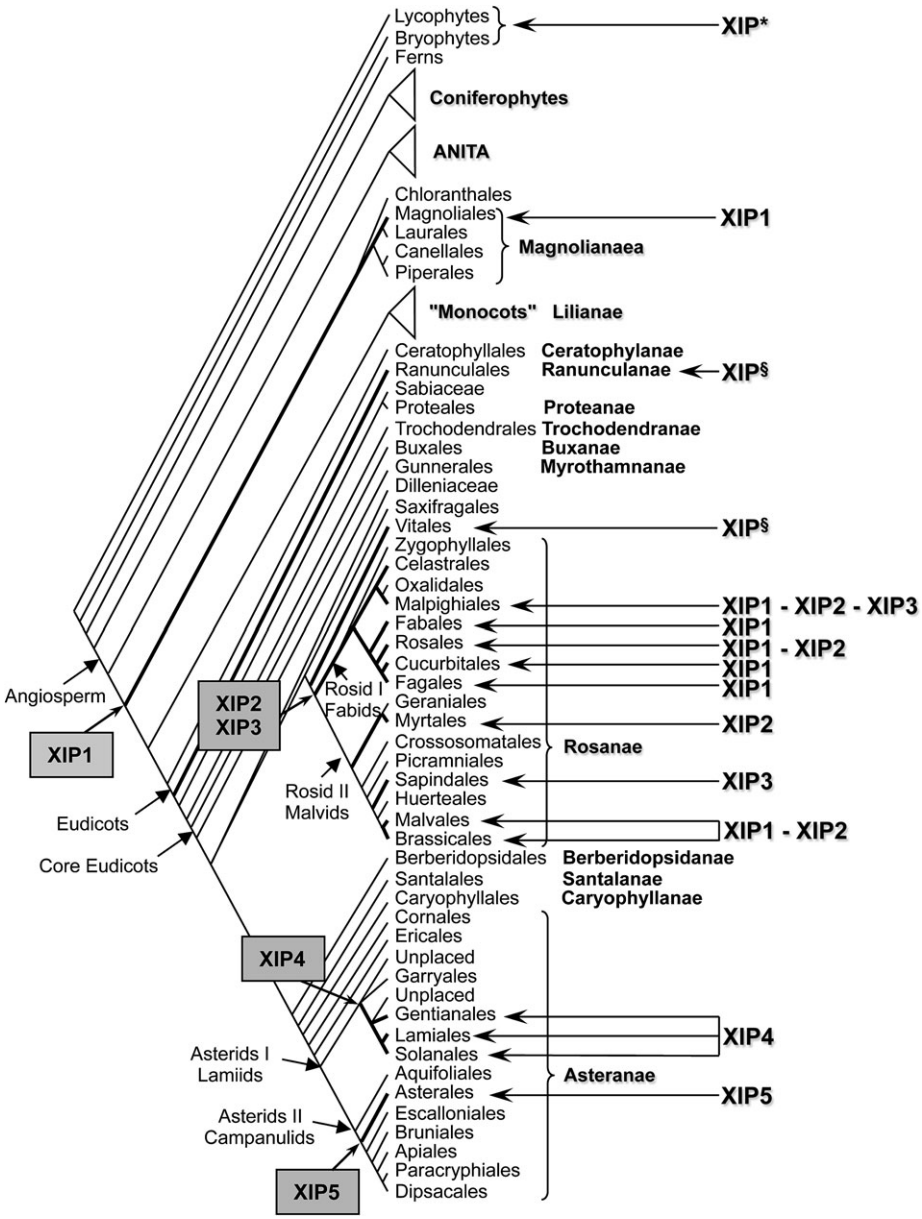


Fig. 2. Evolutionary relationship of XIP sequence divergence and consensus angiosperm phylogeny. The rooted tree proposed here was a compilation of APGIII (2009) and Missouri Botanical Garden (MBG; www.mobot.org) trees. Arrows indicate the presence of XIP groups in taxa. *Selaginella moellendorffii* (Lycopodiophyta) and *Physcomitrella patens* (Bryophyta) XIP sequences (XIP*), and *Aquilegia* (Ranunculanae) and *Vitis* (Vitales) XIP sequences (XIP§) were placed for information purposes, but are not discussed further.

were phylogenetically assigned to five distinct subfamilies (phylogenetic distribution and sequence alignments detailed in Supplementary Figs S1 and S2): PIPs (with 15 members subdivided into five PIP1s and 10 PIP2s), TIPs (17 members), NIPs (11 members), and SIPs (six members). For consistency, gene names specified by Gupta and Sankararamakrishnan (2009) were used for this work. These groups will not be discussed further. The fifth subfamily defines the as yet uncharacterized XIPs, and constitutes the main issue of the present discussion.

The poplar *XIP* subfamily appears to comprise nine putative members with seven full-length protein sequences (DbXrefJGI 557138, 821124, 557139, 759781, 829126, 767334, and POPTR_0009s13105) and two pseudogene sequences (DbXrefJGI 579650 and *PtXIP*-nd). The 829126 sequence (*PtXIP3;1b*) was a truncated sequence due to an ATA triplet insertion within the coding sequence, generating a stop codon nine bases downstream of the initiation codon in relation to its paralogue 557139 (*PtXIP3;1*). This insertion was validated by sequencing multiple clones.

Finally, the full-length *Pt*XIP-related proteins were classified into three groups: *Pt*XIP1 (with two members: *Pt*XIP1;1-557138 and PtXIP1;2-POPTR_0009s13105), *Pt*XIP2 (with one member: *Pt*XIP2;1-821124), and *Pt*XIP3 (with three members: *Pt*XIP3;1-557139, *Pt*XIP3;2-767334, and *Pt*XIP3;3-759781). This new nomenclature is supported by the following three arguments.

(i) Conventionally, plant aquaporins are divided into subfamilies (i.e. NIPs, PIPs, SIPs, and TIPs) that are further divided into groups of related proteins. To maintain consistency with the AQP nomenclature principles of Johanson *et al.* (2001), the *Pt*XIP subfamily should be divided into three groups. Additionally, this repartition was confirmed by the phylogenetic distribution of *Pt*XIPs alone (Fig. 3A). Proposed names for *Pt*XIPs consisted of the subfamily name followed by a number indicating the XIP master group and a second number characterizing the individual XIP member within the group. Intergroup amino acid sequence similarities between *Pt*XIP1 and *Pt*XIP2 or *Pt*XIP3 were ~59%, and ~81% between *Pt*XIP2 and *Pt*XIP3 (Fig. 3B). *Pt*XIP1 members were highly conserved, with a similarity of 94%, as well as *Pt*XIP3s with an average of 90% similarity. The pseudogene sequences and the truncated protein 829126-related sequence (*Pt*XIP3;1b in Fig. 3) were not classified, although they seem to be paralogues of *Pt*XIP2;1 and *Pt*XIP3;1, respectively.

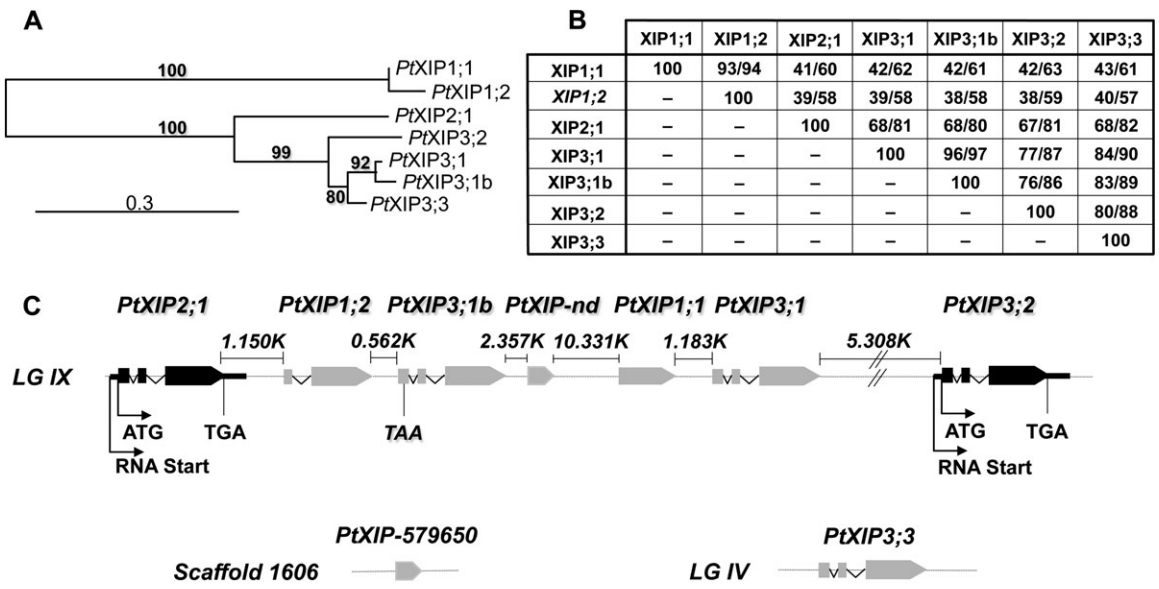
(ii) *Pt*XIP1;2 was distinguished by a gene structure with a unique intron, whereas *Pt*XIP1;1 was an intronless gene. Both exhibited a longer N-terminus as compared with other MIP homologues, a hydrophilic C-loop region deleted of eight amino acid residues specific to the dicot XIP subfamily, and the substitution of the first NPA motif by an SPV motif (Supplementary Fig. S3 at *JXB* online). This substitution was also encountered in *Lotus* and *Glycine max* sequences, while the other XIP1 clustered sequences showed a substitution of the alanine residue of the NPA motif by a valine, isoleucine, or cysteine residue. Remarkably, *Pt*XIP2;1 had both NPA motifs strictly conserved. *Pt*XIP3 members shared in their first NPA motif a substitution of the alanine by an isoleucine or leucine residue.

(iii) Lastly, the evolutionary relationships of the poplar *Pt*XIPs in a compilation of all the angiosperm XIP-like proteins characterized so far backed up overall the choice of nomenclature (Fig. 1).

The new JGIv2.2 nomenclature showed that *Populus MIP* genes were evenly spread over 18 of the 19 haploid 'chromosomes' constituting the poplar genome (Supplementary Table S2 at *JXB* online). Yet, seven out of the nine *Pt*XIP sequences were arranged head-to-tail on linkage group IX (in order, and taking the new nomenclature into account: *PtXIP2;1–PtXIP1;2–PtXIP3;1b–PtXIP*-nd–*PtXIP1;1–PtXIP3;1–PtXIP3;2*); one was located on linkage group IV (*Pt*XIP3;3), and the final one (*PtXIP*-579750) was located on a scaffold that had not yet been assigned to a specific linkage group (Fig. 3C). Such scaffolds were reported to be heterochromatic or derived from substantially divergent haplotypes in the sequenced clone (Tuskan *et al.*, 2007; Kelleher *et al.*, 2008).

Expression pattern of PtXIP *transcripts*

The transcriptional expression of the *PtXIP* gene subfamily and two *PtPIP* genes as references, *PtPIP1;2* (mentioned as



	XIP1;1	XIP1;2	XIP2;1	XIP3;1	XIP3;1b	XIP3;2	XIP3;3
XIP1;1	100	93/94	41/60	42/62	42/61	42/63	43/61
XIP1;2	–	100	39/58	39/58	38/58	38/59	40/57
XIP2;1	–	–	100	68/81	68/80	67/81	68/82
XIP3;1	–	–	–	100	96/97	77/87	84/90
XIP3;1b	–	–	–	–	100	76/86	83/89
XIP3;2	–	–	–	–	–	100	80/88
XIP3;3	–	–	–	–	–	–	100

Fig. 3. Valuer's *Pt*XIP subfamilly. (A) Phylogenetic trees of the *Pt*XIP subfamily. The truncated sequence *PtXIP3;1b* (829126) was included in the analysis, whereas the pseudogenes Scaffold 1606 (*PtXIP-579650*) and XIP-nd related sequences were discarded. Full-length amino acid sequences were analysed by the maximum likelihood method with genetic distance calculated by the JTT model of amino acid change. The numbers at the nodes represent the percentage bootstrap values based on 1000 replicates. The distance scale denotes the number of amino acid substitutions per site. (B) Amino acid identity and similarity percentages between *Populus trichocarpa* cv. Nisqually XIP protein sequences. Percentages were calculated using the NCBI bl2seq algorithm. (C) Schematic representation of the genomic distribution of the nine *PtXIP* genes according to Phytozome v6.0 information (*Populus trichocarpa* cv. Nisqually). Black box regions refer to untranslated mRNA sequences and coding exon sequences transcribed from *P. trichocarpa* cultivar INRA no.101-74, and cloned for this work. Introns are represented by dotted lines, and the numbers above indicate the number of base pairs. The TAA codon stop in *PtXIP3;1b* was corollary generated to the ATA insertion event. LG, linkage group.

PtPIP1;1 in Secchi *et al.*, 2009) and *PtPIP2;2*, was traced in various vegetative organs and in plants exposed to different stresses. The steady-state level of constitutive transcript accumulation relative to the XIP subfamily was monitored in different vegetative tissues including dormant buds (apical and axillary buds mixed), leaves at different maturation stages, apical growing stems, wood and bark (including phloem) of 2-year-old mature stems, and roots (Fig. 4). Of the nine members, only *PtXIP2;1* and *PtXIP3;2* showed detectable expression in *P. trichocarpa* 101-74. No transcripts of *PtXIP1* genes *PtXIP3;1*, or *PtXIP3;3* were detected. *PtXIP2;1* was ubiquitously expressed in vegetative tissues, with high accumulation in buds and immature organs (stems and leaves). It decreased significantly in mature leaves and dropped drastically in senescent leaves. A moderate transcript accumulation was monitored in roots. As regards *PtXIP3;2*, its expression was high in wood, and to a lesser extent, in buds and roots. In all other organs, *PtXIP3;2* expression was very low (mature leaves and bark) or not detectable. *PtXIP1* genes, *PtXIP3;1*, and *PtXIP3;3* did not show significant transcript accumulation under the observed conditions. *PtPIP1;2* and *PtPIP2;2* were highly expressed in all the vegetative tissues of *P. trichocarpa* studied.

PtXIP2;1 and *PtXIP3;2* gene expression was assessed in leaves and branches of plants subjected to severe drought followed by soil re-watering when the most basal leaves showed signs of physiological dehydration. Expression of *PtXIP* genes was also monitored in leaves exposed to SA or mechanical wounding (Fig. 5). Regardless of treatments, *PtXIP2;1* showed a marked transcriptional modulation, whereas *PtXIP3;2* showed limited transcriptional modulation including in stems where its constitutive accumulation appears substantial. No circadian variations in *Pt*XIP mRNA expression patterns were observed (data not shown). During water stress and re-watering (Fig. 5A), *PtXIP2;1* gene expression was greater in leaves than in stems. *PtXIP2;1* was quickly and drastically down-regulated in response to drought. Interestingly, *PtXIP2;1* expression increased 30 min after induction by re-watering and then dropped to its lowest level at 6 h. Gene expression recovered to the basal level 12 h after re-watering. Under SA and wounding treatments, transcripts of *PtXIP2;1* were transiently up-regulated (Fig. 5B, C). Although modulation occurred earlier under SA than wounding, transcripts peaked at 12 h and returned to a steady-state expression level at 24 h in both cases.

Similar to *PtXIP2;1*, *PtPIP* modulations in response to drought were more marked in leaves than in stems. *PtPIP1;2* transcript abundance showed an early and marked up-regulation, whereas *PtPIP2;2* transcript abundance was down-regulated. Both returned to basal levels only 6 h after re-watering. Lastly, transcript levels of *PtPIP1;2* were down-regulated by SA and transiently up-regulated by wounding, whereas transcript levels of *PtPIP2;2* were up-regulated by both SA and wounding.

In situ hybridization

ISH experiments performed on cross-sections of non-stressed *Populus* stems and leaves showed that *PtXIP2;1* transcripts were highly abundant in most vegetative tissues (Fig. 6A–D, I–N), while *PtXIP3;2* expression seemed to be restricted to stems (Fig. 6E–H; Supplementary Fig. S4 at *JXB* online). In stems, *PtXIP2;1* expression was detected in bark, phloem, and wood with a relatively high intensity (Fig. 6A, C). *PtXIP3;2* transcripts co-located well with those of *PtXIP2;1* (Fig. 6E, G). In leaves, due to the cell turgor level and a predominantly vacuolar cell volume, coloration is located in the cell periphery. *PtXIP2;1* expression was detected in almost all parts of the midrib except the cambium (Fig. 6I). In the lamina (Fig. 6K, M) *PtXIP2;1* was present in the epidermis, spongy parenchyma, and vascular tissues of lateral veins (Fig. 6M, arrow lv).

Water transport properties of PtXIPs

To gain further insights into the function of *Pt*XIPs, cRNAs encoding the related proteins were transcribed *in vitro* and injected into *Xenopus* oocytes. After 2–3 d, oocytes were

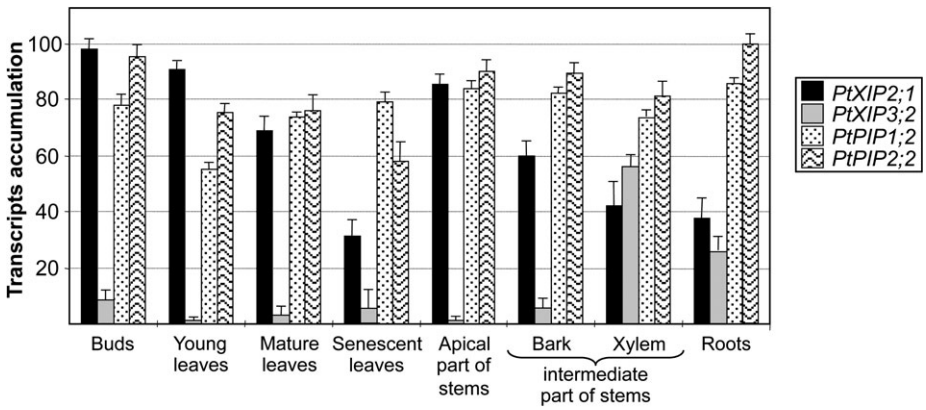


Fig. 4. Real-time quantitative RT-PCR analyses of the constitutive *PtXIP2;1* and *PtXIP3;2* transcript accumulation in vegetative tissues. The expression level of *PtActin 1, PtSAND,* and *PtTIP41-like* gene was used as the normalization internal control. *PtPIP1;2* and *PtPIP2;2* expression was followed for comparison purposes. Bars represent standard deviations of at least three technical repetitions from three independent biological experiments. Four plants were pooled for this biological analysis.

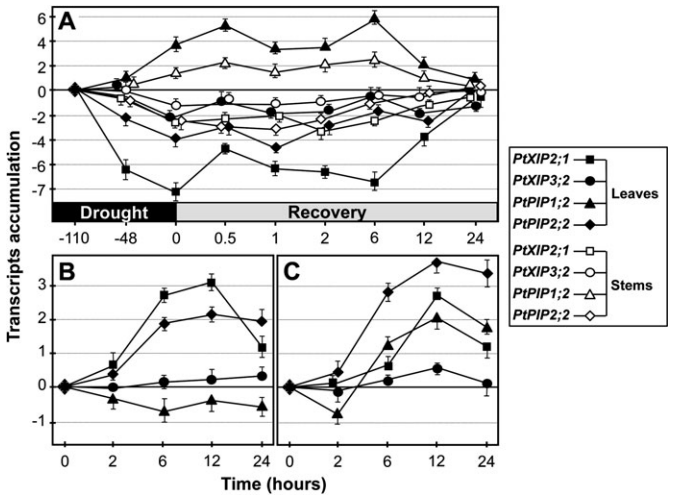


Fig. 5. Real-time quantitative RT-PCR analyses of *PtXIP2;1* and *PtXIP3;2* transcript accumulation kinetics in stress-challenged plants: (A) in leaves and intermediate parts of the stems during the drought stress time-course followed by recovery (re-watering noted at time 0), (B) in 1 mM SA-treated leaves, and (C) in wounded leaves. The expression level of *PtActin 1, PtSAND,* and *PtTIP41-like* genes was used as the normalization internal control. *PtPIP1;2* and *PtPIP2;2* genes were followed for comparison purposes. Bars represent the standard deviations of at least three technical repetitions from three independent biological experiments. Three plants were pooled per biological assay.

transferred into an hypo-osmotic solution and their osmotic water permeability (P_f) was deduced from the swelling kinetics of individual oocytes. Oocytes that did not receive any cRNA or that expressed the highly active aquaporin *Pt*PIP2;2 were taken as negative and positive controls, respectively. Figure 7 shows that whereas injection of *Pt*XIP1s, *Pt*XIP3;1, and *Pt*XIP3;2 cRNA did not induce any change in oocyte P_f, expression of *Pt*XIP2;1 and *Pt*XIP3;3 increased P_f by ~2.5-fold. In parallel experiments, oocytes expressing *Pt*PIP2;2 showed an even more pronounced (~8-fold) increase in P_f. The standard errors are very small because of the high number of cell replicates, which is much higher than in other similar studies. Checks were carried out to verify that similar results (i.e. enhanced P_f in *Pt*XIP2;1 oocytes) were obtained in several independent oocyte batches and cRNA preparations. The present data establish a significant water channel activity for *Pt*XIP2;1 and *Pt*XIP3;3 but not for *Pt*XIP1s, *Pt*XIP3;1, and *Pt*XIP3;2.

Discussion

XIP diversification in the light of angiosperm evolutionary expansion

To date, the comparison of XIP sequences from fungi and plant phyla has been limited to residues that constitute loops and ar/R selectivity filters, and to exon–intron organization (Gupta and Sankararamakrishnan, 2009). In this previous work, it was asserted that during evolution amino acids of the ar/R filter became more hydrophobic in dicot XIPs than in their homologues of fungi and moss. This could reflect a distinct spectrum of solutes transported by these XIPs, which would entail a notable evolutionary divergence between various living kingdoms. Although this previous study sketched interesting evolutionary features, it did not provide insights into the angiosperm taxon by itself which, however, encompasses the great majority of XIP members. By pointing to a notable sequence diversity in angiosperms, the present work highlights evolutionary dynamics of XIPs within this taxon. Without an *a priori* use of previously described XIP sequences, a new query against plant EST collections resulted in the retrieval of 55 putative XIP sequences. All came from a very wide range of flowering plants which, except for *Liriodendron*, were land dicotyledonous plants. Yet, the most striking observation lies in the relationship between the phylogenetic XIP distribution and the currently established phylogeny of angiosperms (Figs 1, 2). The present data emphasizes a prospective evolutionary scheme where XIP members clustered into at least five significantly distinct groups, including four which paralleled the divergence of angiosperm taxa. Although it is likely that the XIPs from angiosperms have originated from a common ancestral gene and have duplicated early after the emergence of this plant clade, it appears clearly that they have then substantially diverged into two clusters (XIP-A and XIP-B). The XIP-A (or XIP1) cluster might include the oldest angiosperm XIP sequences. Indeed, it encompasses the only sequences found in the early-diverging Magnoliales taxa. In core eudicots, except for a clade belonging to Rosanae (Fabibs) in which XIP1s are expressed, the XIP1 evolutionary pathway marks genes for silencing in all other clades. Furthermore, no XIP1 homologues appear in the Asteranae sister clade. As for the second XIP-B cluster, it represents the most interesting expanded subset, with at least four groups. As these groups are only found after core eudicot emergence, they may result from a more recent expansion compared with XIP-A. Thus, XIP2s and XIP3s are exclusively linked to the Rosanae superorder. XIP4s and XIP5s are represented exclusively in Asteranae plants. Finally, this discriminative organization contrasts with the homogenous PIP subfamily and the more heterogeneous NIP, TIP, and SIP subfamilies in which members do not spilt in such a plant taxa-classified manner.

Following gene duplication events, paralogous genes can take on different fates: this includes loss of some of paralogues, divergence and functional differentiation, or maintenance of partially overlapping functions (Conant and Wolfe, 2008). The various diversification scenarios displayed by the XIP subfamily offer a great opportunity to tackle four fundamental questions. (i) As XIP-A members showed a significant divergence and loss of expression in several core eudicots clades, would this branch be subject to loss of function over time? (ii) Would the XIP-B members be newly evolved sequences, suggesting specialized functions? (iii) Would natural selection pressure under *XIP* gene divergence ultimately be involved in speciation processes? (iv) Taken as a whole, despite the actual impossibility of

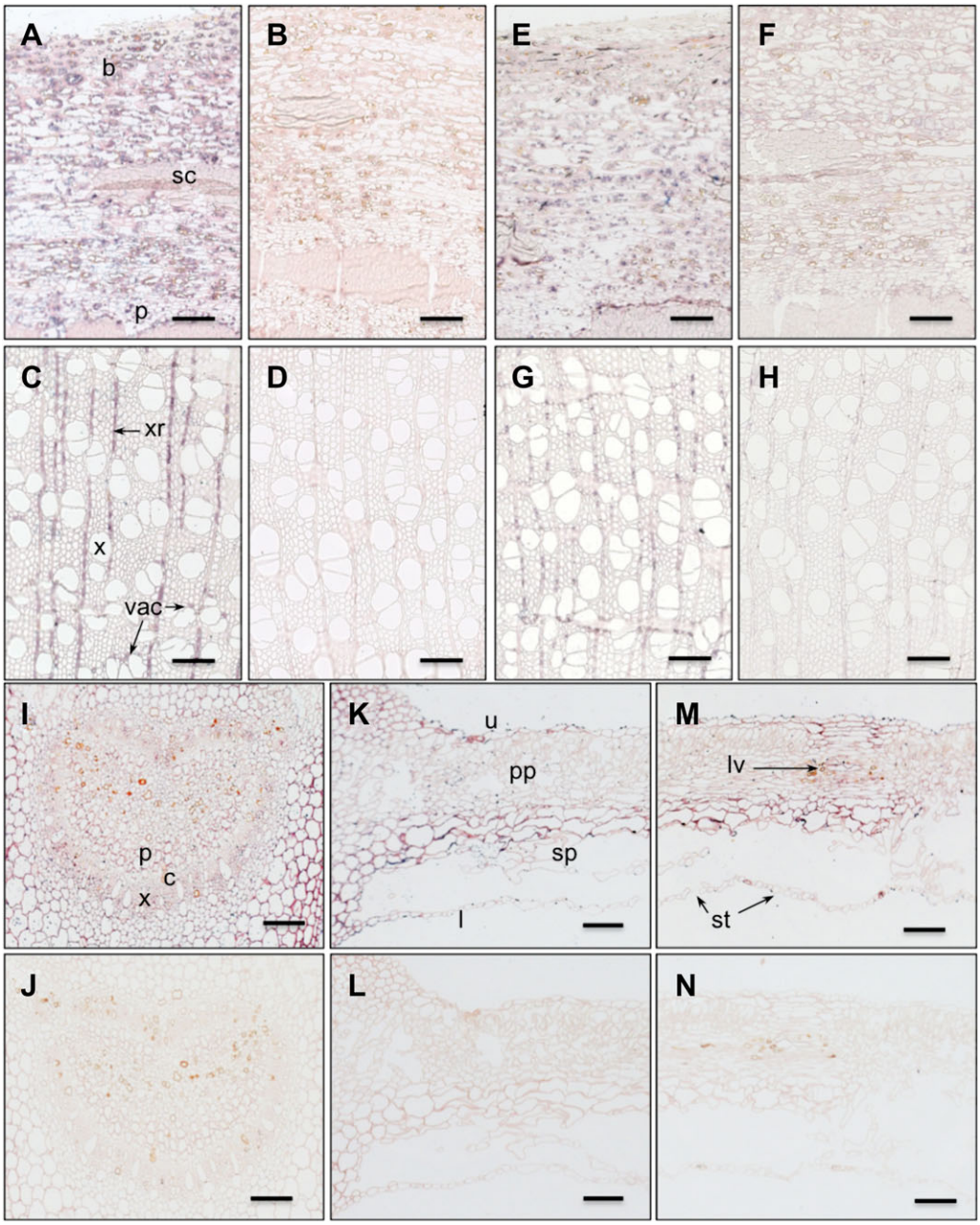


Fig. 6. *In situ* localization of *PtXIP2;1* mRNA in stems (A–D) and in leaves (I–N), and *PtXIP3;2* mRNA in stems (E–H). Methacrylate-embedded *Populus trichocarpa* transversal sections of leaf and stem sample hybridized with specific antisense probes (positive) A, C, E, G, I, K, M; or with sense probes as negative controls (B, D, F, H, J, L, N). b, bark; sc, sclerenchyma; p, phloem; c, cambium; x, xylem vessels; xr, xylem rays; vac, vessel-associated cells; u, upper epidermis; pp, palisade parenchyma; lv, lateral vein; sp, spongy parenchyma; st, stomata; l, lower epidermis. Positive hybridization signals are visualized by violet staining using a DIG-labelled RNA immunodetection system as described in the Material and methods. Arrows indicates alkaline phosphatase staining. Bar indicates 100 μm.

interpreting the presence of XIPs in numerous plant taxa which differ fundamentally (Judd *et al.*, 2002; Soltis *et al.*, 2004), could this unusual phylogenetic repartition be rationalized through some ecological and/or phenotypic explanations? Some of these questions will be answered once more extensive knowledge about ecological and phenotypic relationships, and larger genomic collections are available especially for plants belonging to early-diverging angiosperms and Asteranae.

The vastly expanded XIP subfamily is a unique feature of Populus *species*

Of the embryophytes analysed, poplar is unique in that is is the sequenced land plant that to date contains the highest amount and degree of diversity of XIP copies. The *P. trichocarpa* XIP subfamily (*Pt*XIPs) is composed of nine sequences: among which are six full-length coding sequences (CDS) [leaving aside two pseudogenes and one truncated

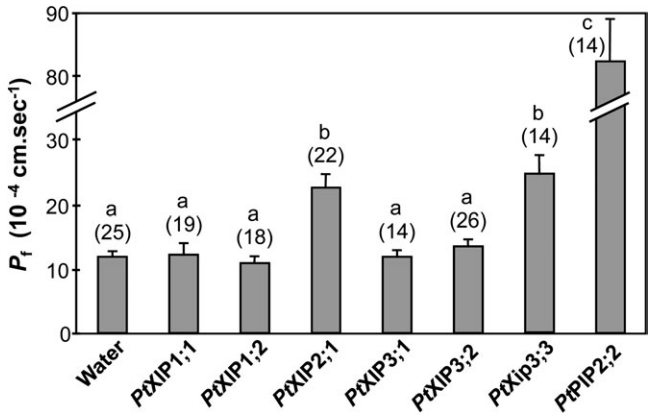


Fig. 7. Osmotic water permeability (P_f) of oocytes expressing aquaporin isoforms. P_f (± SE, with number of cells in parentheses) was measured in oocytes, injected either with water or with cRNAs encoding the indicated aquaporins. Data are from two representative experiments with different oocyte batches and cRNA preparations. Different letters above each bar represent statistically significant differences ($P < 0.05$; one-way ANOVA, Newman–Keuls test).

related product sequence (*Pt*XIP3;1b)]. A fine analysis of the *Pt*XIP subfamily indicated that its members have phylogeneticaly diverged into three branches: *Pt*XIP1 (*Pt*XIP1;1 and *Pt*XIP1;2), *Pt*XIP2 (*Pt*XIP2;1), and *Pt*XIP3 (*Pt*XIP3;1, *Pt*XIP3;2, and *Pt*XIP3;3).

An interesting outcome was revealed during AQP sequence compilation from the two JGI assembly versions: *Pt*XIPs present a significant degree of polymorphism, contrasting with other subfamilies that showed less variation (Supplementary Table S2 at *JXB* online). As previously observed, XIP loci may result from high haplotype variations (Kelleher *et al.*, 2007). Moreover, such amino acid substitutions can reveal potential adaptive evolutionary events, which generally come about with related pressure selection modalities (purifying, neutral, or positive/diversifying). Because the genetic code is redundant, any mutation may or may not be synonymous. Despite a limited number of *Populus* XIP-related ESTs (13 partial sequences phylogenetically homologous to *Pt*XIP2;1 and four partial sequences homologous to *Pt*XIP3;1), the ratio of non-synonymous versus synonymous substitutions was calculated ($\omega = dN/dS$) which is an indicator of the history of selection acting on a gene. Although expressed *Pt*XIPs had high ω values (*Pt*XIP2;1=0.42 and *Pt*XIP3;2=0.37) compared with *Pt*PIPs (*Pt*PIP1s=0.11 and *Pt*PIP2s=0.05), they remained <1. This suggests that in the course of their evolution, expressed *PtXIP* genes underwent a purifying pressure selection, thus illustrating their functional role in the *Populus* genus. A last argument in favour of such purifying pressure concerns the cluster organization of this XIP subfamily. Seven of the nine *Pt*XIP sequences are arranged head-to-tail on chromosome IX. Gene organization typically leads either to a tandem array of reiterated units (e.g. *PtPIP2;5*/*PtPIP2;6*) or to a cluster (such as *PtXIP* genes) when conversion and divergence events occurred (Graham 1995). Such a clustering feature is considered to facilitate the expansion in gene quantity through recombination, and may reflect an adaptive mechanism originating in genome-selective pressure and selection. However, following gene duplication events, paralogous genes can take on alternative fates (Conant and Wolfe, 2008). Within this *Pt*XIP cluster, the duplicated sequences *PtXIP3;1* (expressed in *P. tremula* and *P. tomentiglandulosa*) and *PtXIP3;1b* (non-expressed) are an interesting case study of formation of a gene family in which paralogous copies differentially evolved. Indeed, rather than an acquisition of new adaptive functions through 'positive' mutation (neo-/subfunctionalization), *PtXIP3;1b* may be incapacitated by the occurrence of a deleterious premature stop codon insertion (non-functionalization signature), and then becomes a functionless pseudogene.

A literature review highlights genetic redundancy as a salient feature of living organisms, and gene duplication events are considered as a primary driving force, providing raw material for evolutionary novelty (Taylor and Raes, 2004; Freeling, 2009; Kafri *et al.*, 2009). However, besides the evidence that overlapping functions between duplicate genes manifest as synthetic aggravating interactions between paralogues, they could lead to major repercussions on related steady-state mRNA and/or protein pools and potentially on the regulatory mechanisms controlling various physiological processes, with far-reaching phenotypic effects (Gu *et al.*, 2003; DeLuna *et al.*, 2008). In the light of these data, it is intuitive to note that high non-synonymous variations and major mutations such as premature stop codons may reflect a significant excess of polymorphisms that substantially affect XIP protein structure and function. In other words, this supports and extends these evolutionary results showing that the *Pt*XIP loci and the plant XIP subfamily as a whole could be under a strong selection force associated with duplication and significant differentiations.

PtXIP gene expression and related protein function assessments

Previous works have shown that, when over-represented in tandemly duplicated arrays, plant genes usually respond to environmental constraints (Hanada *et al.*, 2008). Despite poplar *XIP* divergence and expansion, the expressed *XIP* genes are under-represented in poplar databases, with only 17 ESTs, and because most sequences were isolated from pooled tissues no general conclusion could be drawn regarding their expression pattern. *XIP* expression was evaluated first using microarray databases available in public repositories (Gupta and Sankararamakrishnan, 2009). Although microarray methods provide an unprecedented capacity for whole genome profiling, their limits are well characterized, and accurate normalization through quantitative PCR (qPCR) remains a somewhat unavoidable step for obtaining reliable and conclusive results. This is particularly true for the Affymetrix poplar genome array with only five *XIP* genes represented and for which related probesets data should be interpreted with substantial precaution. Here, *XIP* expression was reappraised *in planta* using qPCR. In addition, ISH was performed to determine cell-specific *PtXIP* gene expression patterns. In accordance with expressed sequences

databases, only two members (*PtXIP2;1* and *PtXIP3;2*) were found to be expressed in *P. trichocarpa* 101-74. Surprisingly, they differ substantially from each other by their own accumulative patterns. *PtXIP2;1* was one of the most expressed poplar *MIP* genes. It was ubiquitously expressed in all vegetative tissues with a distinct developmental gradient. It peaked in leaves, with a marked cellular expression in vascular tissues, spongy parenchyma, and epidermis. To a lesser extent, *PtXIP2;1* accumulated in wood, was uniformly distributed in bark and phloem, but was absent in the cambial region where cells have not reached a high level of specialization. As regards *PtXIP3;2*, it preferentially exhibited transcript accumulation in wood, and, although sublocalized in all part of stems (vessel-associated cells, xylem rays, bark, and phloem), its expression reached a maximum in xylem. Similarly, *PtXIP2;1* and *PtXIP3;2* were also expressed differently in challenged plants, where *PtXIP2;1* showed the most contrasted transcript accumulation. A drastic down-regulation of *PtXIP2;1* expression was observed under severe drought stress, whereas *PtXIP2;1* was transiently up-regulated by SA and wounding. Furthermore, the oocyte expression experiments clearly revealed that *Pt*XIP2;1 promotes a significant water channel activity. Considering that the plasma membrane localization of Solanales XIP could be extrapolated to *Pt*XIP, these data may suggest different roles for XIPs in membrane transport at specific plant sublocalizations, and under stable or fluctuating environmental conditions. *Pt*XIP2;1 could play a leading role in this, and, as suggested for *Pt*PIPs and MIPs from various plant species (Heinen *et al.*, 2009; Secchi and Zwieniecki, 2010; Almeida-Rodriguez *et al.*, 2010; this study), *Pt*XIP2;1 can be reasonably considered as a physiological co-actor contributing to regulate the cellular osmotic equilibrium.

Whereas the oocyte expression experiments did not allow any definitive conclusion about the putative transport activity of *Pt*XIP1s, *Pt*XIP3;1, and *Pt*XIP3;2, this approach revealed a significant water permeability for *Pt*XIP2;1 and *Pt*XIP3;3. Previous studies have shown that several sequence regions within XIPs present notable differences compared with typical AQPs. Some of these concern substitutions in the ar/R selectivity filter, forming a quite hydrophobic pore environment predominantly allocated to the conduction of hydrophobic and bulky solutes (Bienert *et al.*, 2011). In the present assays and in line with its water transport functionality, only *Pt*XIP2;1 clearly differs in sequence from the other *Pt*XIPs, with a particular ar/R selectivity filter and canonical NPA boxes. This is also relevant comparing *Pt*XIP2;1 with XIPs of other land plants. A significant water channel activity for *Pt*XIP3;3 was also observed. This ability is quite puzzling as it was not possible to link it with any singular structural motif. Indeed, *Pt*XIP3;3 harboured ar/R residues and NPA motifs similar to those of*Pt*XIP1 and of other members of the *Pt*XIP3 clade for which no water transport capacity was observed. These observations suggested that the transport water ability of *Pt*XIP3;3 is not simply controlled by NPA motifs and the ar/R selectivity filter, and that other structural features could also be involved in this control. Further studies are needed to identify the residue(s) beyond those in the ar/R and NPA regions which are important for this water permeability. As for *Pt*XIP1s, *Pt*XIP3;1 and *Pt*XIP3;2 were apparently devoid of water transport, and they exhibited several ar/R residue similarities with Solanale*s* XIPs which suggested probable common channel activities. Their solute specificity will have to be established.

To conclude, XIP-related data illustrate the potential transport of alternative solutes between members from an MIP subclass, suggesting differential but complementary functional specialization in plant environmental adaptive responses. To conclude definitively on the water and solute transport functions of *Pt*XIPs and their physiological relevance in a plant's life processes, suppression of *XIP* gene function will have to be addressed in future experiments. Moreover, further studies are needed to determine the role, if any, of structural motifs in *Pt*XIP-related functions and those involved in this XIP water permeability.

Supplementary data

Supplementary data are available at *JXB* online.

Figure S1. Phylogenetic trees of all the full-length members of the major membrane intrinsic AQP protein family of *Populus trichocarpa* cv. Nisqually.

Figure S2. Sequence alignment of all the full-length members of the major membrane intrinsic AQP protein family of *Populus trichocarpa* cv. Nisqually.

Figure S3. Protein sequence alignment of the 50 Viridiplantae XIPs.

Figure S4. *In situ* localization of *PoptrXIP3;2* mRNA in leaves.

Table S1. List of the non-redundant representative Viridiplantae *XIP* gene sequences used in this work.

Table S2. Excel spreadsheet of the *Populus trichocarpa* cv. Nisqually *MIP* gene family including all available genomic annotations from JGI.

Table S3. Primers used for qPCR, *in situ* hybridization, and oocyte experiments.

References

Agre P, Sasaki S, Chrispeels MJ. 1993. Aquaporins: a family of water channel proteins. *American Journal of Physiology* **265,** F461.

Almeida-Rodriguez A, Cooke JEK, Yeh F, Zwiazek JJ. 2010. Functional characterization of drought-responsive aquaporins in *Populus balsamifera* and *Populus simonii×balsamifera* clones with different drought resistance strategies. *Physiologia Plantarum* **140,** 321–333.

Altschul S, Madden T, Schaffer A, Zhang J, Zhang Z, Miller W, Lipman D. 1997. Gapped BLASt and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25,** 3389–3401.

Beitz E, Wu B, Holm LM, Schultz JE, Zeuthen T. 2006. Point mutations in the aromatic/arginine region in aquaporin 1 allows

passage of urea, glycerol, ammonia, and protons. *Proceedings of the National Academy of Sciences, USA* **103,** 269–274.

Bienert GP, Bienert MD, Jahn TP, Boutry M, Chaumont F. 2011. *Solanaceae* XIPs are plasma membrane aquaporins that facilitate the transport of many uncharged substrates. *The Plant Journal* **66,** 306–317.

Chang S, Puryear J, Cairney J. 1993. A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* **11,** 113–116.

Chaumont F, Barrieu F, Wojcik E, Chrispeels MJ, Jung R. 2001. Aquaporins constitute a large and highly divergent protein family in maize. *Plant Physiology* **125,** 1206–1215.

Chaumont F, Moshelion M, Daniels MJ. 2005. Regulation of plant aquaporin activity. *Biology of the Cell* **97,** 749–764.

Chase MW, Reveal JL. 2009. A phylogenetic classification of the land plants to accompany APG III. *Botanical Journal of the Linnean Society* **161,** 122–127.

Chevenet F, Brun C, Banuls AL, Jacq B, Chisten R. 2006. TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics* **7,** 439.

Clamp M, Cuff J, Searle SM, Barton GJ. 2004. The Jalview Java alignment editor. *Bioinformatics* **20,** 426–427.

Conant GC, Wolfe KH. 2008. Turning a hobby into a job: how duplicated genes find new functions. *Nature Reviews Genetics* **9,** 938–950.

Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR. 2005. Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiology* **139,** 5–17.

Danielson JAH, Johanson U. 2008. Unexpected complexity of the aquaporin gene family in the moss *Physcomitrella patens*. *BMC Plant Biology* **8,** 45.

De Luna A, Vetsigian K, Shoresh N, *et al*. 2008. Exposing the fitness contribution of duplicated genes. *Nature Genetics* **40,** 676–681.

Engel A, Stahlberg H. 2002. Aquaglyceroporins: channel proteins with a conserved core, multiple functions, and variable surfaces. *International Review of Cytology* **215,** 75–104.

Freeling M. 2009. Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. *Annual Review of Plant Biology* **60,** 433–453.

Graham GJ. 1995. Tandem genes and clustered genes. *Journal of Theoretical Biology* **175,** 71–87.

Gu ZL, Steinmetz LM, Gu X, *et al*. 2003. Role of duplicate genes in genetic robustness against null mutations. *Nature* **421,** 63–66.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* **52,** 696–704.

Gupta AB, Sankararamakrishnan R. 2009. Genome-wide analysis of major intrinsic proteins in the tree plant *Populus trichocarpa*: characterization of XIP subfamily of aquaporins from evolutionary perspective. *BMC Plant Biology* **20,** 134.

Gustavsson S, Lebrun A-S, Norden K, Chaumont F, Johanson U. 2005. A novel plant major intrinsic protein in *Physcomitrella patens* most similar to bacterial glycerol channels. *Plant Physiology* **139,** 287–295.

Hanada K, Zou C, Lehti-Shiu MD, Shinozaki K, Shiu SH. 2008. Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. *Plant Physiology* **148,** 993–1003.

Heinen RB, Ye Q, Chaumont F. 2009. Role of aquaporins in leaf physiology. *Journal of Experimental Botany* **60,** 2971–2985.

Johanson U, Karlsson M, Johansson I, Gustavsson S, Slovall S, Fraysse L, Weig AR, Kjellbom P. 2001. The complete set of genes encoding major intrinsic proteins in Arabidopsis provides a framework for a new nomenclature for major intrinsic proteins in plants. *Plant Physiology* **126,** 1358–1369.

Judd WS, Campbell CS, Kellogg EA, Stevens PF, Donoghue MJ. 2002. *Plant systematics: a phylogenetic approach*. Sunderland, MA: Sinauer Associates.

Kaldenhoff R, Fischer M. 2006. Functional aquaporin diversity in plants. *Biochimica and Biophysica Acta* **1758,** 1134–1141.

Kafri R, Springer M, Pilpel Y. 2009. Genetic redundancy: new tricks for old genes. *Cell* **136,** 389–392.

Kelleher CT, Chiu R, Shin H, *et al*. 2007. A physical map of the highly heterozygous *Populus* genome: integration with the genome sequence and genetic map and analysis of haplotype variation. *The Plant Journal* **50,** 1063–1078.

Kjellbom P, Larson C, Johansson I, Karlsson M, Johanson U. 1999. Aquaporins and water homeostasis in plants. *Trends in Plant Sciences* **4,** 308–314.

Leblanc-Fournier N, Coutand C, Crouzet J, Brunel N, Lenne C, Moulia B, Julien JL. 2008. Jr-ZFP2, encoding a Cys2/His2-type transcription factor, is involved in the early stages of the mechano-perception pathway and specifically expressed in mechanically stimulated tissues in woody plants. *Plant, Cell and Environment* **31,** 715–726.

Maurel C. 2007. Plant aquaporins: novel functions and regulation properties. *FEBS Letters* **581,** 2227–2236.

Maurel C, Reizer J, Schroeder JI, Chrispeels MJ. 1993. The vacuolar membrane protein γ-TIP creates water specific channels in *Xenopus* oocytes. *EMBO Journal* **12,** 2241–2247.

Maurel C, Santoni V, Luu DT, Wudick MM, Verdoucq L. 2009. The cellular dynamics of plant aquaporin expression and functions. *Current Opinion in Plant Biology* **12,** 690–698.

Mitani-Ueno N, Yamaji N, Zhao FJ, Ma JF. 2011. The aromatic/arginine selectivity filter of NIP aquaporins plays a critical role in substrate selectivity for silicon, boron, and arsenic. *Journal of Experimental Botany* **62,** 4391–4398.

Park W, Scheffler BE, Bauer PJ, Campbell BT. 2010. Identification of the family of aquaporin genes and their expression in upland cotton (*Gossypium hirsutum* L.). *BMC Plant Biology* **10,** 142.

Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research* **29,** 2003–2007.

Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP. 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: *BestKeeper*—Excel-based tool using pair-wise correlations. *Biotechnology Letters* **26,** 509–515.

Rozen S, Skaletsky HJ. 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology* **132,** 365–386.

Sade N, Vinocur BJ, Diber A, Shatil A, Ronen G, Nissan H, Wallach R, Karchi H, Moshelion M. 2009. Improving plant stress tolerance and yield production: is the tonoplast aquaporin SlTIP2;2 a key to isohydric to anisohydric conversion? *New Phytologist* **181,** 651–661.

Sakurai J, Ishikawa F, Yamlaguchi T, Uemura M, Maeshima M. 2005. Identification of 33 rice aquaporin genes and analysis of their expression and function. *Plant and Cell Physiology* **46,** 1568–1577.

Secchi F, Maciver B, Zeidel ML, Zwieniecki MA. 2009. Functional analysis of putative genes encoding the PIP2 water channel subfamily *Populus trichocarpa*. *Tree Physiology* **29,** 1467–1477.

Secchi F, Zwieniecki MA. 2010. Patterns of PIP gene expression in *Populus trichocarpa* during recovery from xylem embolism suggest a major role for the PIP1 aquaporin subfamily as moderators of refilling process. *Plant, Cell and Environment* **33,** 1285–1297.

Shelden MC, Howitt SM, Kaiser BN, Tyerman SD. 2009. Identification and functional characterization of aquaporins in the grapevine, *Vitis vinifera*. *Functional Plant Biology* **36,** 1065–1078.

Simossis VA, Heringa J. 2005. PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. *Nucleic Acids Research* **33,** W289–W294.

Soltis PS, Soltis DE, Chase MW, Endress PK, Crane PR. 2004. The diversification of flowering plants. In: Cracraft J, Donoghue MJ, eds. *The tree of life*. Oxford: Oxford University Press, 154–167.

Taylor JS, Raes J. 2004. Duplication and divergence: the evolution of new genes and old ideas. *Annual Review of Genetics* **38,** 615–643.

Tornroth-Horsefield S, Wang Y, Hedfalk K, Johanson U, Karlsson M, Tajkhorshid E, Neutze R, Kjellbom P. 2006. Structural mechanism of plant aquaporin gating. *Nature* **439,** 688–694.

Tuskan GA, DiFazio S, Jansson S, *et al*. 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313,** 1596–1604.

Wallace IS, Roberts DM. 2004. Homology modelling of representative subfamilies of Arabidopsis major intrinsic proteins. Classification based on the aromatic/arginine selectivity filter. *Plant Physiology* **135,** 1059–1068.

Xu M, Zhang B, Su X, Zhang S, Huang M. 2011. Reference gene selection for quantitative real-time polymerase chain reaction in Populus. *Analytical Biochemistry* **408,** 337–339.