

Functional evolution of nodulin 26-like intrinsic proteins: from bacterial arsenic detoxification to plant nutrient transport

Benjamin Pommerrenig^{1,2} , Till A. Diehn¹ , Nadine Bernhardt³ , Manuela D. Bienert¹ ,
Namiki Mitani-Ueno⁴ , Jacqueline Fuge¹ , Annett Bieber¹ , Christoph Spitzer¹ , Andrea Bräutigam⁵ ,
Jian Feng Ma⁴ , François Chaumont⁶  and Gerd P. Bienert¹ 

¹Department of Physiology and Cell Biology, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), 06466, Gatersleben, Germany; ²Division of Plant Physiology, University Kaiserslautern, 67663, Kaiserslautern, Germany; ³Genebank Department, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), 06466, Gatersleben, Germany; ⁴Institute of Plant Science and Resources, Okayama University, 710-0046, Kurashiki, Japan; ⁵Computational Biology, Faculty of Biology, Bielefeld University, 33615, Bielefeld, Germany; ⁶Louvain Institute of Biomolecular Science and Technology, UC Louvain, 1348, Louvain-la-Neuve, Belgium

Summary

Author for correspondence:

Gerd P. Bienert

Tel: +49 39482 5385

Email: bienert@ipk-gatersleben.de

Received: 15 May 2019

Accepted: 17 September 2019

New Phytologist (2019)

doi: 10.1111/nph.16217

Key words: aquaporin, arsenic, boron, evolution, metalloids, nodulin 26-like intrinsic proteins, nutrient transport.

- Nodulin 26-like intrinsic proteins (NIPs) play essential roles in transporting the nutrients silicon and boron in seed plants, but the evolutionary origin of this transport function and the co-permeability to toxic arsenic remains enigmatic. Horizontal gene transfer of a yet uncharacterised bacterial AqpN-aquaporin group was the starting-point for plant NIP evolution.
- We combined intense sequence, phylogenetic and genetic context analyses and a mutational approach with various transport assays in oocytes and plants to resolve the transorganismal and functional evolution of bacterial and algal and terrestrial plant NIPs and to reveal their molecular transport specificity features.
- We discovered that *aqpN* genes are prevalently located in arsenic resistance operons of various prokaryotic phyla. We provided genetic and functional evidence that these proteins contribute to the arsenic detoxification machinery. We identified NIPs with the ancestral bacterial AqpN selectivity filter composition in algae, liverworts, moss, hornworts and ferns and demonstrated that these archetype plant NIPs and their prokaryotic progenitors are almost impermeable to water and silicon but transport arsenic and boron. With a mutational approach, we demonstrated that during evolution, ancestral NIP selectivity shifted to allow subfunctionalisations.
- Together, our data provided evidence that evolution converted bacterial arsenic efflux channels into essential seed plant nutrient transporters.

Introduction

The metalloids boron (B) and silicon (Si) are fundamental for the development of vascular plants not only because they ensure proper differentiation, structural support and elasticity of plant cell walls, but also because they contribute to pathogen defence and general stress tolerance (Ma *et al.*, 2006; Bienert & Chaumont, 2011). Consequently, corresponding transport mechanisms for these elements are essential for sufficient plant B and Si nutrition. Nodulin 26-like intrinsic proteins (NIPs) mediate B and Si transport and thereby sustain growth, fertility and yield of terrestrial plants. NIPs belong to the major intrinsic protein (MIP) superfamily (also termed as aquaporins, AQPs), which comprises channels for the diffusion of small neutral and mostly polar molecules across various biological membranes in all kingdoms of life. MIP channels possess six transmembrane helices connected by five loops and the two termini facing the cytoplasm. The so-called aromatic/arginine (ar/R) constriction region acts as an MIP selectivity filter (SF) and comprises four amino acids of

transmembrane helices 2 and 5 (positions R1 and R2) and loop E (positions R3 and R4) representing the narrowest part of the channel pore and forming a size exclusion barrier conferring selectivity to particular substrates (Murata *et al.*, 2000).

Phylogenetic analyses revealed the existence of four phylogenetically well supported prokaryotic MIP clades (Finn *et al.*, 2014; Finn & Cerdà, 2015). These are aquaporin Z-like proteins (AqpZ), aquaporin M-like proteins (AqpM), glycerol uptake facilitator-like proteins (GlpF), and aquaporin N-like proteins (AqpN). AqpZ proteins permeate water, AqpMs both water and glycerol, and GlpFs water, glycerol, urea and metalloids (Fu *et al.*, 2000; Jensen *et al.*, 2001; Kozono *et al.*, 2003; Lee *et al.*, 2005). Bacterial AqpN channels have not been functionally characterised to date. Eukaryotic MIPs evolved from bacterial and/or archaeal ancestor isoforms (Pao *et al.*, 1991; Park & Saier, 1996; Zardoya *et al.*, 2002). MIPs of higher plants group phylogenetically with AqpZs and the yet nonstudied AqpNs.

Sequences for NIP subfamily members have not been found in animal and fungal genomes and it has, therefore, been

hypothesised that the first plant NIP was the result of a horizontal *aqpN* gene transfer (HGT) event during the origin of plant life forms (Finn & Cerdà, 2015; Zardoya *et al.*, 2002; Danielson & Johanson, 2010; Abascal *et al.*, 2014). However, the exact origin of the plant NIP subfamily and the evolution and diversification of functions within early plant lineages are completely unclear and unresolved.

NIPs form one of the isoform-richest subfamily in plants with a high diversity regarding their amino acid sequences and substrate specificities. Besides their vast substrate spectrum (Bienert & Chaumont, 2011; Bienert & Bienert, 2017), the physiological relevance of NIPs *in planta* may be restricted to the transport of the essential and beneficial metalloids B and Si, but also to the extrusion of the toxic minerals arsenic (As), antimony (Sb), and germanium (Ge) (Pommerrenig *et al.*, 2015; Bienert & Bienert, 2017). NIPs subdivide into three groups termed NIP-I, NIP-II and NIP-III (Wallace & Roberts, 2004; Mitani *et al.*, 2008). This grouping is based on common substrate selectivities, sequence similarities and an amino acid composition consistency of their ar/R SF. The NIP-I and NIP-II subgroups are present in all higher plants, while NIP-III isoforms are largely, but not exclusively, confined to Liliopsida species (Ma & Yamaji, 2015; Trembath-Reichert *et al.*, 2015; Deshmukh & Bélanger, 2016).

Multiple NIP isoforms belonging to the NIP-II group (such as AtNIP5;1, AtNIP6;1, ZmNIP3;1, and OsNIP3;1) have been unambiguously demonstrated to be essential for the root uptake and translocation of B in various monocot and dicot plant species and therefore for B nutrition in seed plants (reviewed in Pommerrenig *et al.*, 2015; Bienert & Bienert, 2017). NIP-III group isoforms are mainly found in plant species that are biosilicifying or which benefit from high Si concentrations in their plant body such as *Poaceae* or *Cucurbitaceae* species. Accordingly, NIP-III group members (such as OsNIP2;1, OsNIP2;2, HvNIP2;1, HvNIP2;2 and CmNIP2;1) were shown to be crucial for an efficient root uptake and translocation of Si and therefore for Si homeostasis in such plant species (reviewed in Ma & Yamaji, 2015; Bienert & Bienert, 2017).

NIPs differing from these three subgroups have been identified in the moss *Physcomitrella patens* and the lycophyte *Selaginella moellendorffii* (Danielson & Johanson, 2008; Anderberg *et al.*, 2012). Those NIPs form a fourth NIP group (NIP-IV) and have a SF composition consisting of F_{R1}-A_{R2}-A_{R3}-R_{R4} residues. As this SF layout is restricted to bacterial AqpNs and a few NIP isoforms found in extant members of early land plant lineages, it might represent the archetype SF layout of all plant NIPs and the original NIP channel selectivity (Finn & Cerdà, 2015; Danielson & Urbanson, 2010; Abascal *et al.*, 2014; Trembath-Reichert *et al.*, 2015). It is unlikely that boric acid and silicic acid transport, the essential functions of modern seed plant NIP-IIs and NIP-IIIs, respectively, represent the ancestral function of bacterial AqpN proteins as neither B nor Si essentiality is common in bacteria. Initial analyses of seed plant NIPs identified glycerol as the first common substrate and suggested a role for bacterial AqpNs and NIPs in early nonvascular terrestrial plants in glycerol uptake and distribution (Zardoya *et al.*, 2016; Roberts & Routray, 2017). However, the exact origin of plant NIPs and the

evolution and diversification of functions within early plant lineages are completely unclear and neither bacterial AqpNs nor plant NIPs with a F-A-A-R-type SF composition have been characterised so far.

The present work sheds light into the functional evolution of NIPs. We combined intense sequence- and genetic context analyses with transport assays to resolve the functional evolution of NIPs within plants and reveal molecular features of bacterial AqpNs and algal and terrestrial plant F-A-A-R-type NIP channels. AqpN channels are permeable to arsenous acid and are frequently part of arsenic (As) resistance operons (*ars* operons) of diverse bacterial phyla. This suggests that an intrinsic function of the ancestors of plant NIPs namely the bacterial AqpNs resides in As detoxification processes and that modern NIP channels underwent subfunctionalisation and neofunctionalisation and turned from As effluxers into essential and beneficial plant nutrient channels. Permeability of ancestral plant NIPs was a premise for the ability of vascular plants to efficiently take up and translocate B and Si into and within the plant body. This transport ability first allowed plants to use these elements efficiently, a crucial prerequisite for terrestrial upright growth and stress tolerance.

Materials and Methods

Arabidopsis

Arabidopsis thaliana Col-0 wild-type (NASC: N60000) and *Atnip5;1* knock-down mutant (SALK_122287; NASC: N622287) seeds were obtained from The European Arabidopsis Stock Centre. For *in vitro* culture, seeds were sterilised with 70% ethanol plus 0.05% Triton X-100 followed by three washes of 99% ethanol. Seeds were plated on half-strength Murashige and Skoog (½MS) medium (2.2 g l⁻¹ MS salts with minimal organics (Sigma), 1% sucrose, 0.7% agar, pH 5.8 (KOH)) with or without antibiotics and vernalised for 2 d at 4°C. *In vitro* cultures were grown in a photoperiod of 10 h : 14 h, 22°C : 19°C, light : dark (120 µmol m⁻² s⁻¹). For soil growth experiments, seedlings were transferred into pots after 2.5 wk of *in vitro* growth and grown in a phytochamber at 18°C and a 17 h light period for another 2.5 wk.

Physcomitrella

Physcomitrella patens (Gransden, IMSC no. 40001) was obtained from the International Moss Stock Center and grown at 22°C under a 14 h : 10 h, light : dark cycle and a photosynthetic photon flux density of 90 µmol m⁻² s⁻¹. Moss cultures were cultivated on a modified Knop medium (Reski & Abel, 1985). Final nutrient concentrations were: KH₂PO₄ (250 mg l⁻¹), KCl (250 mg l⁻¹), MgSO₄ (250 mg l⁻¹), Ca(NO₃)₂ · 4H₂O (1000 mg l⁻¹), FeSO₄ · 7H₂O (12.5 mg l⁻¹). 1% Phytagel was used as solidifying agent. Before sterilisation of the medium, the pH was adjusted to 5.8 with KOH. To remove traces of B, the B chelating agent Amberlite IRA-743 (Sigma) was used. Amberlite IRA-743 was washed three times with MilliQ water and incubated together with the Knop medium at a concentration of

3 g l⁻¹ overnight. Amberlite IRA-743 was removed before autoclaving the medium. Metalloids were added to the Knop growth medium at 500 µM sodium arsenate dibasic heptahydrate (As(V)), 350 µM sodium arsenite (As(III)), and 10 µM boric acid.

Sequence retrieval, gene context analysis and phylogenetic analysis

Multiple public-accessible gene, operon and genome databases have been used for sequence retrievals and gene context analyses. Bayesian phylogenetic analyses and tree computation have been performed with curated protein alignments. Detailed information on these procedures is provided in Supporting Information Methods S1. Sequences which have been used in this study are given in Dataset S1.

Cloning and vector construction

Detailed information about vector constructions and the procedures for molecular cloning techniques is provided in Methods S1.

Subcellular localisation and confocal microscopic imaging

Detailed information on transient transformation protocols and subsequent subcellular localisation analyses of YFP-tagged PpNIP5 proteins using a Zeiss LSM 780 confocal laser scanning microscope can be found in Methods S1.

Oocyte transport assays

In vitro cRNA synthesis, oocyte handling procedures, and various oocyte uptake assays with subsequent determination of permeability coefficients or element levels of oocytes are described in detail in Methods S1.

Complementation analysis of *Atnip5;1* T-DNA insertion mutants

Detailed information on the T-DNA insertion line, vector constructions, the procedures for transgenic Arabidopsis generation and selection is provided in Methods S1.

RNA extraction, cDNA synthesis and real-time quantitative PCR

Various experimental information related to the different working steps necessary for a RT-qPCR as well as RT-qPCR-related specification details are described in detail in Methods S1.

Statistical analysis

The two-sample Student's *t*-test was applied to find out which independent sample sets were significantly different from each other. Two-sided testing was used. In all cases, statistical significance was defined as: *, *P* < 0.05; **, *P* < 0.01; or ***,

P < 0.001. Tukey's test was used to compare means at a probability level of 5%. In this case, levels of significance are represented by letters as indicated in the captions.

Results

Phylogenetic analysis of NIP-type aquaporins from prokaryota and eukaryota

Our phylogenetic analyses using 366 bacterial, archaeal and plant NIP-type MIPs showed that the major prokaryotic MIP groups, which are AQPZs, AQPMs, AQPNs and GLPFs, formed well supported clades (Fig. 1). Six major AQPN clades (I–VI) with a full node support were detected. Two of these AQPN clades do not encode F–A–A–R-type SFs but display a wide variety of SF residue combinations: AQPN-I is mainly composed of archaeal and AQPN-II of bacterial sequences (Fig. 1).

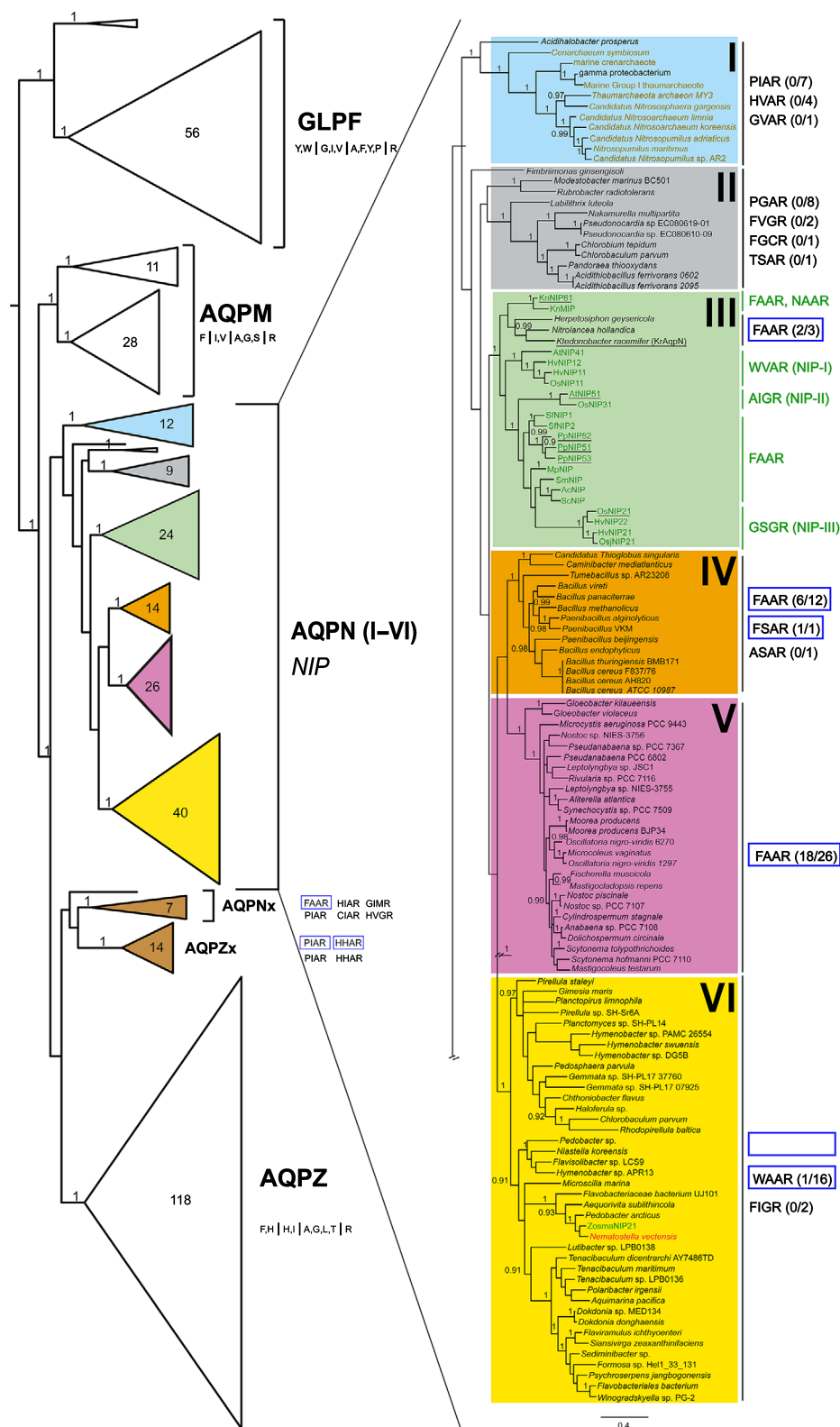
The remaining four AQPN clades are characterised by a predominant occurrence of F–A–A–R-type SFs. All plant NIP-type sequences as well as three bacterial AqpNs formed together the AQPN-III clade (Figs 1, S1). Within this clade, algal NIPs (KnNIP6;1 and KnMIP) can be clearly separated from the NIP-I, NIP-II, NIP-III groups and the F–A–A–R-type NIPs of land plants (Figs 1, S1). Interestingly, the F–A–A–R-type NIPs of land plants cluster closest with the NIP-III group members, which are crucial for Si transport in seed plants.

AQPN-IV clade is formed by sequences deriving mainly from *Firmicutes* but also *Proteobacteria* that are F–A–A–R-type SF-dominated (Fig. 1). The AQPN-V clade is exclusively formed by cyanobacterial sequences, which all encode F–A–A–R-type SFs (Fig. 1). The member-rich AQPN-VI clade is formed by sequences from various bacterial phyla either encoding F–A–A–R-type or an W–A–A–R-type SFs (Fig. 1).

Two additional MIP groups, AQP_{Nx} and AQP_{Zx} cluster apart from other AqpZ and AqpN-like sequences. These sequences derive from various different bacterial and archaeal species, and encode MIPs with very diverse SFs and form a phylogenetically loose cluster. The overall backbone of the phylogenetic tree is unresolved.

Bacterial AqpNs with an F–A–A–R-type selectivity filter are abundantly located in arsenic resistance operons

Very few *aqpNs* have previously been identified in the genomes of few bacterial phyla (Finn & Cerdà, 2015; Danielson & Johanson, 2010; Abascal *et al.*, 2014). To obtain more information on the uncharacterised bacterial AQPN clade, which is of high significance for the evolution of plant nutrient NIP channels, we studied AqpNs sequences in more detail. The phylogenetic analysis revealed that only 57 (56 bacterial and one archaeal) AQPN-like sequences are present in the 4576 prokaryotic genomes deposited in the KEGG database (Fig. 1). All of these AqpNs are classified as 'AQPZs' or 'MIPs' in the database. Therefore, only 1.3% of the prokaryote genomes of the database encode for AqpNs (Fig. 2).



By contrast, 1879 and 2893 sequences were classified as *aqpZ* and *glpF* genes, respectively. Identified AqpNs derived from 10 and one bacterial and archaeal phyla, respectively, with the bacteria themselves sharing no distinct lifestyle or habitat. The SFs of AqpNs are variable but many of these

are composed of the amino acid residues $F_{R1}-A_{R2}-A_{R3}-R_{R4}$ (Fig. 1; Dataset S1).

In order to infer a potential role of AqpNs, we investigated their genetic context providing information on operons and functional units. Remarkably, 19 out of the 57 *aqpNs* (33.3%) were

Fig. 1 Phylogeny of bacterial major intrinsic proteins (MIPs) and plant nodulin 26-like intrinsic proteins (NIPs). Consensus tree derived from archaeal, bacterial, plant and putative metazoan amino acid sequences using Bayesian phylogenetic inference. Numbers beside the nodes indicate the posterior probability values if larger than 0.9. For the presentation of the tree midpoint-rooting was applied. To the left, fully supported clades were collapsed. The size of the triangles corresponds to the number of sequences (given in the triangle if larger than three) comprised in the clade. Prokaryotic GLPF, AQPm, AQPn and AQPz clades are labelled and the amino acid residue compositions which comprise the aromatic/arginine selectivity filter (SF) in the corresponding clade are given below in the one-letter-code. Six well supported AQPn/NIP clades (I–VI) were formed. These clades are coloured differently AQPn-I: blue – archaeal sequences, AQPn-II: grey – bacterial sequences, AQPn-III: green – mainly plant and three bacterial sequences, AQPn-IV: orange – *Firmicutes* and *Proteobacteria* sequences, AQPn-V: purple – cyanobacterial sequences, AQPn-VI: yellow – sequences from various different prokaryotic phyla. AQPnx and AQPzx clades (brown triangles) neither clearly belong to the AQPz nor AQPn clades. A few genes that belong to these clades have been found in arsenic resistance operons (*ars* operons). SF compositions found in these clades are displayed next to the clade name. Blue boxed SFs occur in *ars* operons. To the right, a detailed representation of the AQPn/NIP clade is depicted. Instead of the prokaryotic AqpN protein name/identifier, the organism by which the MIP is encoded is displayed. Different font colours indicate different taxonomic groups: brown, archaea; black, bacteria; red, metazoan; green, plants. If multiple individuals for a taxon were included then they are distinguished by number. Different vertical line colours indicate whether the proteins of this AQPn clade are classified as AqpN (black) or plant NIP (green) sequences. For each clade, the amino acid residues that determine the SFs are shown in the one-letter-code (green and black writing indicates NIP and AqpN SF compositions, respectively). Blue boxed SF combinations are found in *ars* operons. The first digit in parentheses indicates how many genes of the labelled clade with the indicated SF are located in *ars* operons while the second digit is the total number of genes with the indicated SF composition within the labelled AQPn clade. AQPn-III clade sequences that have been functionally analysed in this study are underlined. Sequences, identifiers and SF motifs can be found in Supporting Information Dataset S1.

located in arsenic resistance (*ars*) operons, while the remaining *aqpN* isoforms do not seem to be associated to any other non-metalloid regulatory genetic unit (Dataset S1). All except one of the 19 *ars* operon-encoded AqpNs contained the F–A–A–R SF layout. AqpNs were part of diversely constituted *ars* operons, consisting of genes encoding at least an ArsR transcription activator, an arsenate reductase (ArsC), and sometimes an arsenite *S*-adenosylmethyltransferase (ArsM) or an arsenical resistance protein (ArsH). The observation that AqpNs were frequently located in *ars* operons suggests that they play a physiologically important role in As detoxification as effluxers of As out of bacterial cells.

To further investigate the abundance of *aqpNs* in *ars* operons, a protein BLAST search retrieved 36 AqpN sequences that clustered with AqpN isoforms, for which the genetic context could be accessed, and which were not present in the KEGG database (Fig. 1; Dataset S1). In this case, 50% of the *aqpN* sequences (18 out of 36) were found in *ars* operons, and all of them code for channels having the F–A–A–R SF-motif. The finding that *aqpNs* are present with a high frequency in *ars* operons was further supported by the reanalysis of 18 MIPs that have been identified in 685 *ars* operons (Yang *et al.*, 2015). We assigned six AqpNs (all encoding for an F–A–A–R-type SF) and 12 GlpFs (Fig. S2a). Moreover, in an operon database covering 2072 prokaryotic genomes, we identified 31 MIP gene containing *ars* operons. Amongst them, our analyses identified three *aqpNs* encoding an F–A–A–R-type SF and 28 *glpF* genes (Fig. S2b).

Occurrence of F–A–A–R-type NIPs in land plants

The prevalence of F–A–A–R residues constituting the SF in bacterial AqpNs and in early diverging plant NIPs such as *S. moellendorffii* and *P. patens* suggests that the last common ancestor to the plant NIPs encoded a F–A–A–R-type SF. To study the frequency and evolution of this SF composition, we screened various plant genome databases for NIPs with this SF composition. We identified a NIP with an F–A–A–R-type SF in the genome of the charophyte *Klebsormidium nitens* and thereby discovered the presence of NIPs also in green algae (Fig. 1). This

finding suggests that the HGT of an *aqpN* happened before plants invaded the land and associate with the estimation of Zardoya *et al.* (2002). Additionally, we found such F–A–A–R-type NIPs in liverwort (*Marchantia polymorpha*), moss (*Sphagnum fallax*) and fern (*Salvinia cucullata* and *Adiantum capillus-veneris*) taxa (Figs 1, 2). These findings closed the knowledge gap on the NIP lineage throughout the phylogeny from bacteria, green plants, over land plants to vascular plants. Seed plants, both gymnosperms and flowering plants, encode various NIPs with SFs typical for the NIP-I, NIP-II and NIP-III group. However, the F–A–A–R SF-group has not been identified in seed plants in any of the analysed genome-rich databases.

Identification of F–A–A–R-type MIPs allowing studying the functional evolution of NIP-type aquaporins

We subjected several MIPs to further analyses to resolve the functional evolution of F–A–A–R-type MIPs from bacteria along the plant lineage. KrAqpN is found in the aerobic, filamentous, non-motile, Gram-positive bacterium *Ktedonobacter racemifer* belonging to the *Chloroflexi* phylum from which an *aqpN* gene may have been transferred via a HGT to plants (Fig. 1). Moreover, this F–A–A–R SF-encoding *aqpN* is located in a typical *ars* operon (Fig. 2). The F–A–A–R SF-carrying KnNIP6;1 is encoded by the filamentous terrestrial alga *K. nitens*, which belongs to the *Charophyta* green algae group from which all terrestrial plants, the Embryophyta, emerged (Hori *et al.*, 2014). The name KnNIP6;1 was adopted according to a detailed phylogenetic analysis (U. Johanson, per. commun.: H. I. Anderberg & U. Johanson, unpublished). PpNIP5;1, PpNIP5;2, and PpNIP5;3 are encoded by the moss *P. patens* and represent members of the F–A–A–R-type NIPs of early land plant lineages.

Bacterial, algal and moss F–A–A–R-type NIPs do not conduct water in *Xenopus* oocytes

The water channel activity of KrAqpN, KnNIP6;1, PpNIP5;1, PpNIP5;2, and PpNIP5;3 was tested by

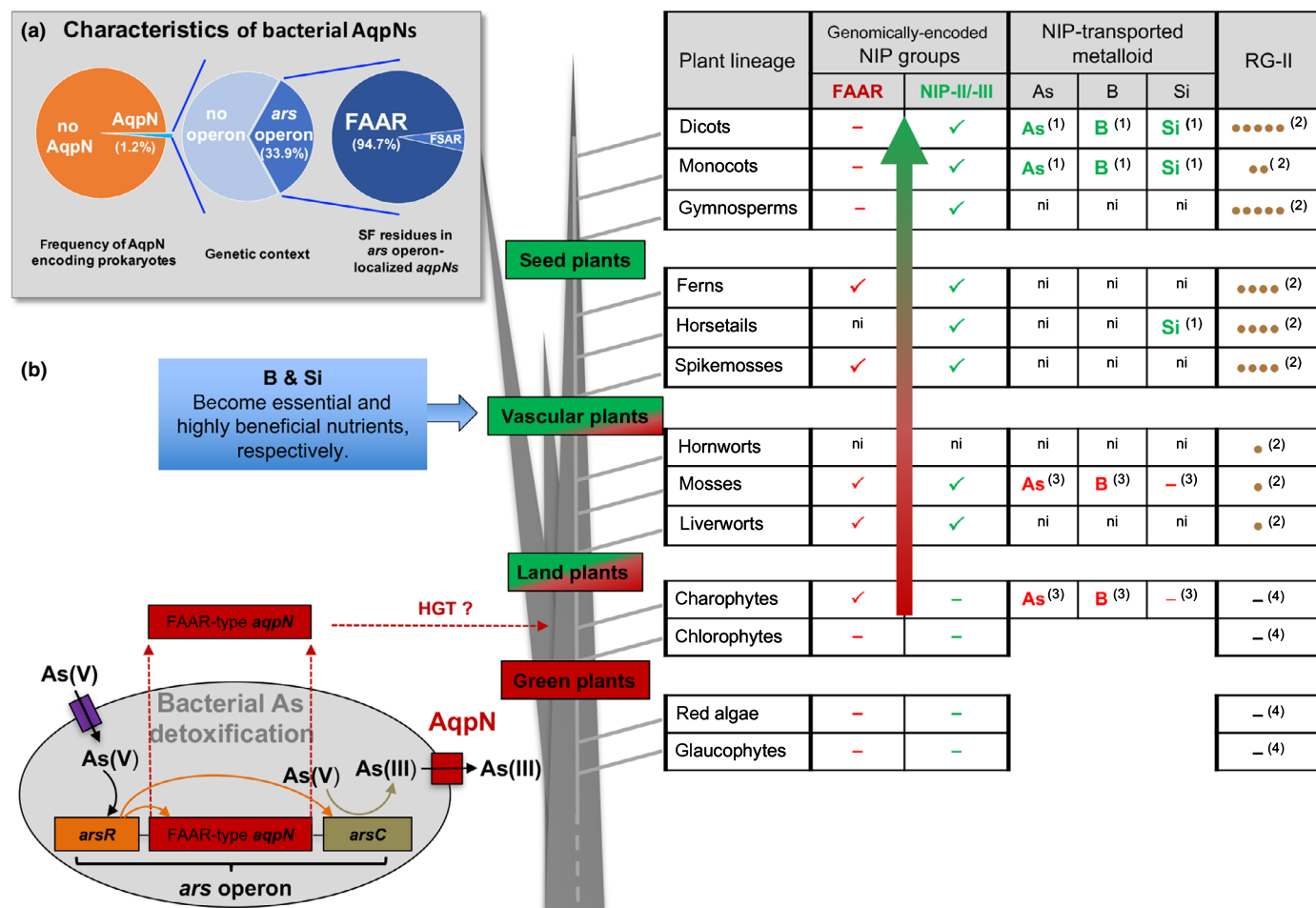


Fig. 2 Origin, evolution and functions of the plant metalloid-transporting nodulin 26-like intrinsic protein (NIP) clade. (a) Characteristics of bacterial AqpNs. 1.3% of the prokaryotic genomes deposited in the KEGG database possess *aqpN* genes. 33.3% of *aqpNs* are part of arsenic (As) resistance (*ars*) operons. 94.7% of the *ars* operon-localised *aqpNs* have a selectivity filter (SF) composed of F–A–A–R residues. (b) Prokaryotes (light grey circle) possess *ars* operons for As detoxification and export. Instead of active As efflux transporters, some bacteria possess an *ars* operon-located *aqpN* gene (red), which facilitates the transmembrane diffusion of As(III) species. In prokaryotes, arsenate (As(V)) enters the cells via phosphate transporters (purple box). Transcription of *arsR* (orange) is induced by As(V) and the protein produced, ArsR, regulates the expression of *aqpN* (red) and *arsC* (olive-green). The chemical gradient needed to cause As efflux through AqpNs is maintained by the reduction of As(V) to As(III), catalysed by the arsenate reductase ArsC. The SF of many AqpNs is composed of F–A–A–R residues. An AqpN with a F–A–A–R-type SF has been suggested to be transferred to plants by a horizontal gene transfer (HGT) event. During land plant evolution (central panel) AqpNs have been employed by plants and underwent neo- and subfunctionalisation. While in primitive green plants only NIPs with a F–A–A–R-type SF are found (red shading) other functional NIP groups (green shading) such as the NIP-II group amplified. When silicon (Si) and boron (B) became beneficial and essential elements for vascular plant groups, respectively, F–A–A–R-type NIPs disappeared from the genomes of these plants. The change of F–A–A–R-type SFs to other SF compositions in evolving NIP-II and NIP-III group NIPs of higher plants contributed to a shift in metalloid transport properties. Si permeability formed and As transport ability probably diminished in favour of B permeability. The table summarises: (1) whether F–A–A–R-type NIPs (red check mark) and/or NIP-II type NIPs (green check mark) exist in genomes of the indicated phyletic plant group; (2) experimentally determined substrate specificity for the corresponding NIP-type (green font, NIP-II type NIPs; red font, F–A–A–R-type NIPs); (3) qualitative amount of rhamnogalacturonan-II (RG-II) molecules within the cell wall of corresponding plant groups (one brown circle represents minimal RG-II detection and five brown circles represent the RG-II amount found in dicots). –, assayed but not existent; ni, present knowledge does not allow a classification as no information is available; ¹Bienert & Bienert (2017); ²O'Neill *et al.* (2004); ³this study; ⁴Popper *et al.* (2011)

heterologous expression in *Xenopus* oocytes. High osmotic water permeability coefficient (P_f) values were obtained for cells expressing the positive control ZmPIP2;5 (Fetter *et al.*, 2004). By contrast, no significant P_f increase was observed for cells expressing the different F–A–A–R-type MIPs from bacteria, algae or moss compared to control oocytes (Fig. 3a).

Bacterial, algal and moss F–A–A–R-type NIPs are permeable to arsenous acid and boric acid when expressed in *Xenopus* oocytes

We determined the transport activity for arsenous acid in oocytes expressing *KrAqpN*, *KuNIP6;1*, *PpNIP5;1*, *PpNIP5;2*, *PpNIP5;3*, or the positive control *OsLsi1* (Ma *et al.*, 2006).

Expression of all MIPs resulted in a significant increase in As transport activity compared to water-injected oocytes quantified via ICP-MS analysis (Fig. 3b).

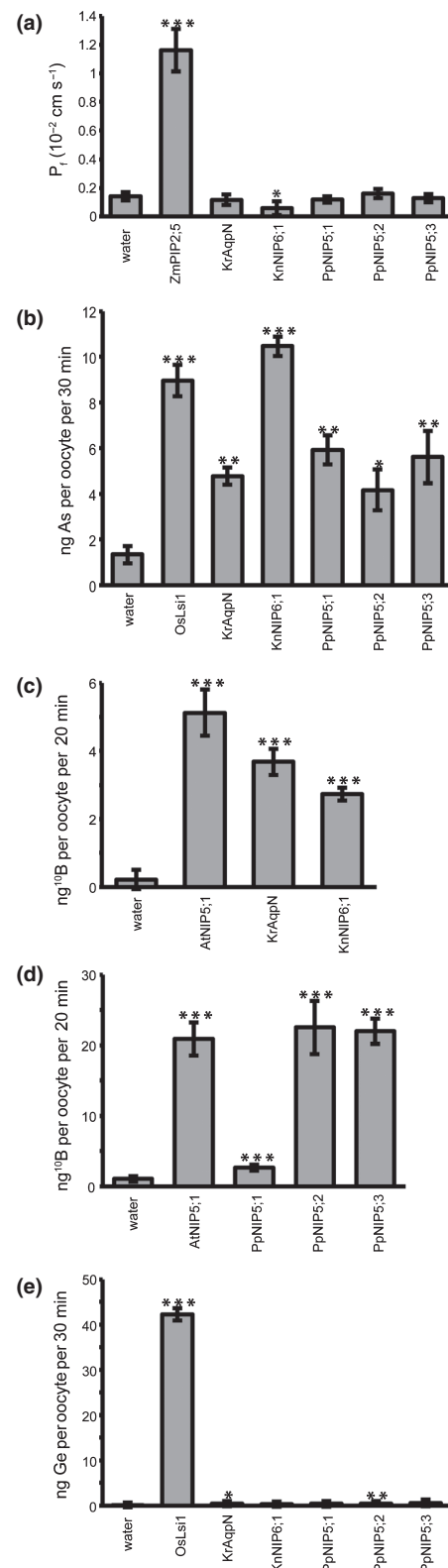
The permeability to boric acid was tested in direct uptake assays into oocytes in the presence or the absence of MIP expression. AtNIP5;1 which is physiologically important for the uptake of B into Arabidopsis roots was used as a positive control (Takano *et al.*, 2006). Quantification of B levels in oocytes showed that the bacterial, algae and moss MIPs significantly increased the B uptake over that of water-injected oocytes (Fig. 3c,d).

Bacterial, algal and moss F–A–A–R-type NIPs do not facilitate the diffusion of silicic acid in *Xenopus* oocytes

To examine the silicic acid permeability of KrAqpN, KnNIP6;1, PpNIP5;1, PpNIP5;2, and PpNIP5;3 isoforms, we measured Ge accumulation in oocytes expressing these MIPs or OsLsi1, a NIP-III group NIP, as a positive control (Fig. 3e). Germanium dioxide (GeO_2), generating germanic acid in solution, proved to be an excellent tracer for silicic acid in plant- and oocyte uptake assays (Ma *et al.*, 2006). Oocytes expressing *OsLsi1* accumulated about 300-fold more Ge than water-injected control oocytes when incubated with 1 mM GeO_2 . None of the F–A–A–R-type MIPs from bacteria, algae or moss increased the Ge

uptake in a similar manner to the positive control *OsLsi1* or to a level which would point to a physiologically significant Si channel (Fig. 3e).

Fig. 3 Transport selectivity of bacterial and plant F–A–A–R-type NIPs. (a) Water permeability coefficient (P_i) measurements of bacterial (KrAqpN) and plant F–A–A–R-type NIPs (KnNIP6;1, PpNIP5;1, PpNIP5;2, PpNIP5;3) in a hypo-osmotic *Xenopus laevis* oocyte swelling assay. *ZmPIP2;5* cRNA- and water-injected oocytes were used as positive and negative controls, respectively. Bar charts show means $\pm$ 95% CIs of $n = 9$ –16 oocytes. (b) Arsenous acid uptake rates of *Xenopus* oocytes expressing bacterial (KrAqpN) and plant F–A–A–R-type NIPs (KnNIP6;1, PpNIP5;1, PpNIP5;2, PpNIP5;3). Oocytes expressing *OsLsi1* and water-injected oocytes were used as positive and negative controls, respectively. Oocytes were exposed to a 0.1 mM NaAsO_2 containing buffer solution for 30 min. Data represent means $\pm$ SD of four pools of oocytes ($n = 5$ –8 oocytes per pool). (c, d) Uptake of ^{10}B boric acid by KrAqpN or KnNIP6;1 (c) and PpNIP5;1, PpNIP5;2, and PpNIP5;3 (d) expressing oocytes in a direct uptake assay. Oocytes expressing AtNIP5;1 and water-injected oocytes were used as positive and negative controls, respectively. Oocytes were exposed to a 5 mM ^{10}B boric acid containing buffer solution for 20 min. Data represent means $\pm$ SD of 3–10 pools of oocytes ($n = 8$ –10 oocytes per pool) (c) or 3–5 pools of oocytes ($n = 9$ –11 oocytes per pool) (d). Differences in absolute uptake rates in (c,d) are due to different oocyte batches. (e) Permeability of a bacterial AqpN and plant F–A–A–R-type NIPs to germanic acid (a chemical analogue of silicic acid) in *Xenopus* oocyte direct uptake assays. Oocytes expressing KrAqpN, KnNIP6;1, PpNIP5;1, PpNIP5;2 and PpNIP5;3 and the control *OsLsi1* as well as water-injected oocytes were exposed to 1 mM GeO_2 (which forms in solution $\text{Ge}(\text{OH})_4$, as the analogue of $\text{Si}(\text{OH})_4$) for 30 min in a modified Barth's saline. Bar charts represent means $\pm$ SD of 4 pools of oocytes ($n = 4$ –7 oocytes per pool). For metalloid quantifications (b–e), the assayed oocytes were washed, dried, digested and the As, ^{10}B and Ge concentrations were determined by ICP-MS. Significant differences in transport rates compared to water-injected negative control oocytes were assessed in (a–e) using Student's *t*-test: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. All oocyte uptake assays have been repeated two to three times with independent oocyte batches and consistent results.



Bacterial AqpN_{F-A-A-R} and plant NIP-II_{A-I-G-R} type selectivity filters are tailor-made to efficiently conduct their physiological substrates arsenous acid and boric acid, respectively

NIP-I, NIP-II and NIP-III groups developed in land plants in addition to F-A-A-R-type NIPs (Figs 1, 2). The number of non-F-A-A-R-type isoforms exceeds the number of F-A-A-R-type NIPs in vascular plants, for example in horsetails and lycophytes. No F-A-A-R-type NIP was identified in seed plant genomes. The stepwise disappearance of F-A-A-R-type NIPs along the evolution of land plants may imply an evolutionary disadvantage of that SF composition in seed plants compared to NIP-I to NIP-III group SF compositions. F-A-A-R-type NIPs disappeared when B and Si acquired a dominant nutritional role in land plants (Fig. 2), possibly because they are efficient As channels but are suboptimal to serve as efficient B/Si channels. Therefore, we tested the hypothesis whether a switch from an F-A-A-R- to a NIP-II-type SF results in a decrease in arsenous acid transport rates relative to boric acid transport rates. To this aim, SFs were mutated in KrAqpN (a bacterial *ars* operon-residing AqpN channel with permeability to As), PpNIP5;3 (a F-A-A-R-type moss NIP of unknown physiological function) and AtNIP5;1 (a physiologically important NIP-II type B channel). The SFs of KrAqpN_{FAAR} and PpNIP5;3_{FAAR} were mutated into the SF of AtNIP5;1_{AIGR} (SF of the NIP-II group) resulting in

KrAqpN_{AIGR} and PpNIP5;3_{AIGR}. The SF of AtNIP5;1_{AIGR} was mutated into AtNIP5;1_{FAAR}. Oocytes expressing these channel proteins were simultaneously exposed to an arsenous acid and boric acid containing buffer to avoid a quantitative bias on metalloid uptake levels due to potentially different active channel numbers in individual oocyte batches. The absolute values of As and B uptake rates cannot be compared with each other, given that different As and B concentration gradients have been applied for experimental reasons (see the Materials and Methods section). The expression of native AtNIP5;1_{AIGR} resulted in higher B levels in oocytes than the expression of native PpNIP5;3_{FAAR} and KrAqpN_{FAAR} channel proteins (Fig. 4a, left panel).

Inversely, the exactly same PpNIP5;3_{FAAR} and KrAqpN_{FAAR} expressing oocytes which have been analysed for B uptake,

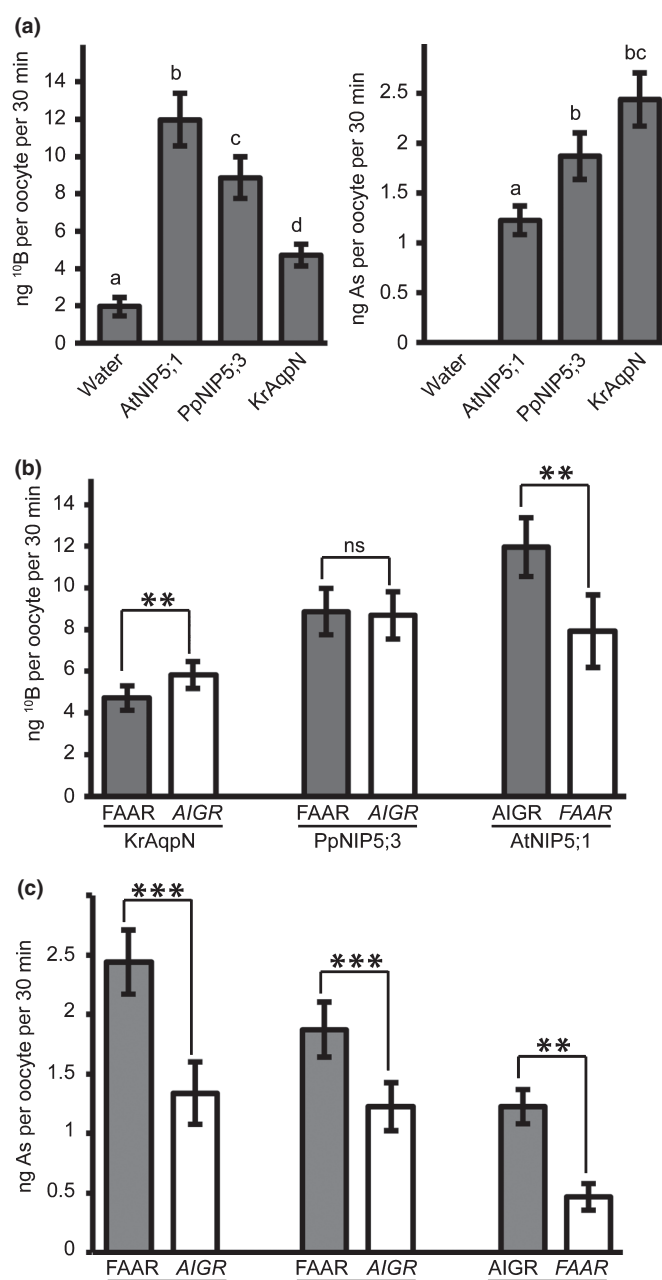


Fig. 4 Influence of the ar/R selectivity filter (SF) composition of KrAqpN, PpNIP5;3 and AtNIP5;1 on the permeability to arsenous acid and boric acid. (a) ¹⁰Boric acid (left panel) and arsenous acid (right panel) transport activity was simultaneously determined in direct uptake assays using *Xenopus laevis* oocytes expressing native AtNIP5;1, KrAqpN and PpNIP5;3 channel proteins. Water-injected oocytes were used as negative controls. Oocytes were exposed to a 5 mM ¹⁰boric acid and 0.1 mM NaAsO₂ containing buffer solution for 30 min. Oocytes were washed, dried, digested and the ¹⁰B and As content was determined by HR-ICP-MS analysis. The As content of water-injected negative control oocytes were below the detection limit of the high-resolution mass spectrometer. Grey chart bars represent the means of metalloid uptake rates ± SD of seven pools ($n = 10$ oocytes per pool), four pools ($n = 8$ oocytes per pool), nine pools ($n = 9–10$ oocytes per pool) and five pools ($n = 9–10$ oocytes per pool) of water-injected oocytes or oocytes having expressed AtNIP5;1, KrAqpN or PpNIP5;3, respectively. Significant differences in metalloid uptake rates between negative control oocytes or AtNIP5;1 and KrAqpN or PpNIP5;3 expressing oocytes were calculated using the Tukey's test to compare means at a probability level of 5%. Levels of significance are represented by lower-case letters. (b, c) ¹⁰Boric acid (b) and arsenous acid (c) transport activity was determined as in (a) for oocytes expressing native AtNIP5;1_{AIGR}, KrAqpN_{FAAR} or PpNIP5;3_{FAAR}, or SF-mutated proteins AtNIP5;1_{FAAR}, KrAqpN_{AIGR} or PpNIP5;3_{AIGR}. Grey chart bars in (b) and (c) represent the same data as in the left and right panel in (a), respectively. White chart bars represent the metalloid uptake rate means ± SD of four pools ($n = 8$ oocytes per pool), 10 pools ($n = 10–11$ oocytes per pool) and 10 pools ($n = 10$ oocytes per pool) of oocytes having expressed the mutated proteins AtNIP5;1_{FAAR}, KrAqpN_{AIGR} or PpNIP5;3_{AIGR}, respectively. Significant differences in metalloid uptake rates between native AtNIP5;1, KrAqpN or PpNIP5;3 channels and their corresponding SF mutant variants in (b,c) were calculated using Student's *t*-test: **, $P < 0.01$; ***, $P < 0.001$; ns, not significant. The uptake assay has been repeated twice with independent oocyte batches and consistent results.

possessed significantly higher As levels than the oocytes having expressed AtNIP5;1_{AIGR} (Fig. 4a, right panel). These results indicate that the significantly higher B and lower As uptake rates, respectively, of AtNIP5;1 compared to KrAqpN and PpNIP5;3 cannot be explained by different expression levels of the three MIP constructs, but must have resulted from different selectivity properties such as a differing SF composition. When the native SF of KrAqpN_{FAAR} was mutated into that of the physiological B channel AtNIP5;1 (KrAqpN_{AIGR}), significantly decreased As levels (Fig. 4c) and significantly increased B levels (Fig. 4b) were observed in oocytes suggesting a quantitative shift in the transport rates. The expression of the mutant PpNIP5;3_{AIGR} also resulted in a significant relative decrease of As levels in oocytes compared to that of the native PpNIP5;3_{FAAR} channel, while B transport rates remained similar (Fig. 4b,c). The expression of mutated AtNIP5;1_{FAAR} compared to that of AtNIP5;1_{AIGR} resulted in significantly decreased As and B levels indicating an overall negative impact of that SF composition on AtNIP5;1 metalloid transport capacity (Fig. 4b,c).

YFP:PpNIP5 fusion proteins localise to the plasma membrane

To study subcellular localisation, constructs carrying fusions of YFP and the coding sequences of PpNIP5;1, PpNIP5;2 or PpNIP5;3 were transiently expressed in epidermal tobacco leaf cells. In addition, transformed cells were stained with the fluorescent plasma membrane (PM) marker dye FM4-64. YFP- or FM4-64- dependent fluorescence appeared mainly at the borders of the cells. Overlay of YFP- and FM4-64- channels merged fluorescence signals and indicated that YFP:PpNIP5s fusion proteins localised to the PM (Fig. 5a).

F–A–A–R-type PpNIP5 isoforms partially rescue the boron deficiency phenotype of the *Atnip5;1* knock-down mutant

To test whether plant F–A–A–R-type NIPs are functional metalloidoporins in plants and to have an experimentally independent confirmation of the B transport ability, we performed an *in planta* growth complementation assay. We transformed *Atnip5;1* knock-down Arabidopsis lines with genetic constructs in which PpNIP5;1, PpNIP5;2 and PpNIP5;3 were cloned behind the control of the *AtNIP5;1* promoter sequence. The *Atnip5;1* plants were unable to take up sufficient amounts of B from the soil substrate under our standard growth conditions, resulting in a strong B-deficiency phenotype (Fig. 5b). This growth deficit was complemented when the *Atnip5;1* mutant was transformed with the native *AtNIP5;1* cDNA (Fig. 5b,c). Independent PpNIP5;1 and PpNIP5;2 expressing lines also significantly complemented the *Atnip5;1* mutant growth phenotype (Fig. 5b,c) and had a significantly higher B uptake ability, indicated by the higher shoot B concentrations, than the *Atnip5;1* knock-down mutants, unambiguously demonstrating their ability to facilitate the uptake of B into plant roots (Fig. 5d). However, the PpNIP5;1 and PpNIP5;2 expressing lines exhibited still slight B-deficiency symptoms when

compared to the wild-type (Fig. 5b). It was previously observed that the expression of functional NIP5 isoforms or the *AtNIP5;1pro:AtNIP5;1* construct do not fully restore the shoot B concentrations of *Atnip5;1* mutants compared to the wild-type (Diehn *et al.*, 2019). All generated PpNIP5;3 harbouring lines spliced the *PpNIP5;3* mRNA resulting in a *cDNA* sequence which is not encoding a full-length MIP (Fig. S3). Accordingly, *PpNIP5;3* expression failed to complement the growth of the *Atnip5;1* mutant. In *P. patens*, *PpNIP5s* transcript abundances were not regulated in response to varying metalloid feeding and exposure levels (Fig. S4).

Discussion

Bacterial *aqpNs* are rare, abundantly located in *ars* operons and facilitate the transport of arsenous acid

In accordance with previous studies considering few *aqpNs* of few bacterial phyla, our phylogenetic analyses suggested that plant NIPs descend from an ancestral *aqpN* gene of extant *Chloroflexi* bacteria (Zardoya, 2005; Danielson & Johanson, 2010; Abascal *et al.*, 2014; Finn & Cerdà, 2015; Zardoya *et al.*, 2016; Roberts & Routray, 2017). Alternatively, plant NIPs, as many other ‘plant’ genes, may derive from cyanobacteria, which represent the ancestors of plant plastids and which frequently possess *aqpNs* in their genomes. Due to low node support, the exact evolution of NIPs from their bacterial AqpN ancestors across the prokaryote–eukaryote border could not be resolved. The reason for this probably lies in the paucity of NIP sequences from nonseed plants and the long period since these taxa shared a most recent common ancestor. Our database survey illustrated that *aqpN* genes are much less abundant than genes clustering with the intensively characterised GLPF and AQPZ clades. The total number of identified *aqpN* sequences is extremely small (Fig. 1). Throughout different prokaryotic phyla, *aqpN* genes are strikingly located in *ars* operons. *GlpF* genes are also found in *ars* operons but the majority is found in carbon utilisation-related operons (Bienert *et al.*, 2013). None of the (by us) identified *aqpNs* was part of a carbon regulon and this contrasts with the genetic context of the glycerol channel-encoding *glpFs* arguing against the hypothesis that glycerol transport constitutes the physiological function of AqpNs and thereby accounts for the original function of the archetype of plant NIPs (Roberts & Routray, 2017).

In agreement with the *ars* operon localisation of *aqpNs*, we demonstrated their permeability to As as exemplified by KrAqpN. Together, functional and genetic data suggest that AqpNs play a physiological role in As efflux. AqpNs of diverse bacteria are placed together with *arsM* or *arsH* genes in *ars* operons. ArsH oxidises trivalent to relatively nontoxic pentavalent methylated As species, the latter representing potential substrates for AqpNs/NIPs given that a rice NIP can indeed transport such organic As species (Li *et al.*, 2009; Chen *et al.*, 2015). This may indicate that certain AqpNs and NIPs mediate the efflux of organic in addition to inorganic As species to detoxify cells. Based on the observation that the tested KrAqpN was permeable to boric acid, we also hypothesise that AqpNs of cyanobacterial

origins are permeable to boric acid since many cyanobacteria also require B for metabolism (Bonilla *et al.*, 1990).

As AqpNs are closest related to members of the AQPZ clade, which play key roles in bacterial water regulation, we also tested

the bacterial AqpN for water transport. Despite the phylogenetic vicinity, KrAqpN was not facilitating the diffusion of water (Fig. 3a). This is in agreement with the poor water permeability often observed for the closely related plant NIP-I to NIP-III group isoforms (Pommerrenig *et al.*, 2015).

Surprisingly, BLAST searches identified *aqpNs* in the genomes of the monocot seagrass *Z. marina* and the metazoan sea anemone *N. vectensis* (Fig. 1). The genome database information on these intron-lacking sequences does not allow a conclusion whether they are prokaryotic DNA contaminations, more likely or whether these MIPs have arisen from an unlikely, late HGT. Both scenarios can be explained by the fact that *Z. marina* and *N. vectensis* share aquatic habitats with members of the *Bacteroidetes*, which encode for *aqpN* genes (Fig. 1).

Occurrence and potential role of F–A–A–R-type NIPs in algae and early land plants

To gain a better understanding on NIP evolution, we compared archetype NIP sequences throughout different taxa of land plants. We identified F–A–A–R-type NIPs in almost all groups of plants up to ferns with the exception of horsetails and hornworts, which might be due to insufficient genome coverage and sequence accessibility for species belonging to these taxa (Fig. 2). The occurrence of NIPs comprising F–A–A–R residue SFs does not correlate with the environmental habitats of the corresponding plant taxa. There is strong evidence that the ancestors of current terrestrial plants were closely related to the ancestors of present-day charophyte taxa (Hori *et al.*, 2014). We identified F–A–A–R-type NIPs in the genome of the charophyte *K. nitens*. In our transport assays, the algal KnNIP6;1 does not transport water but is permeable to arsenous acid and boric acid, essentially displaying a similar transport profile compared with the tested bacterial KrAqpN. Comparing the existence of *NIP* genes in the genomes of diverse plant species and the amount of Si found in their tissues suggested the existence of an ancestral Si transport ability of archetype plant NIPs (Trembath-Reichert *et al.*, 2015).

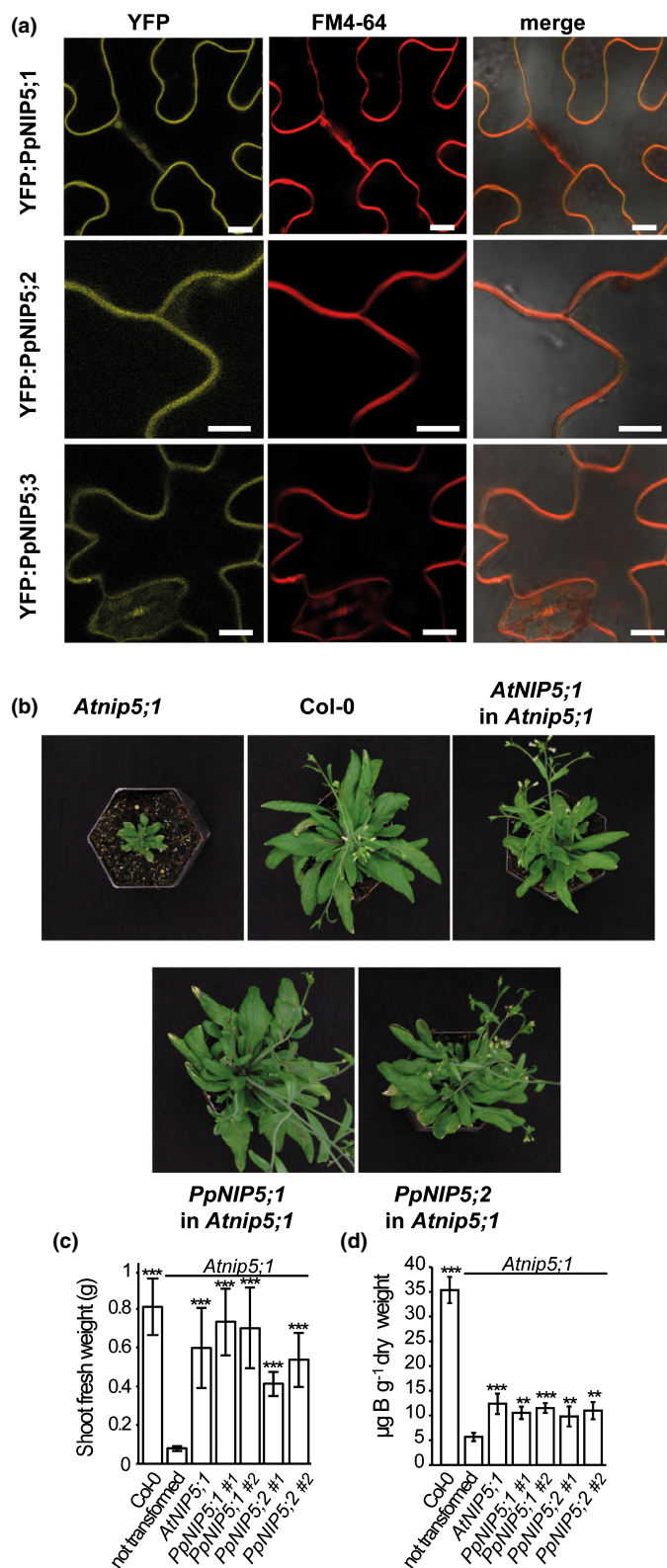


Fig. 5 Determination of *in planta* PpNIP5 properties. (a) Subcellular localisation of YFP:PpNIP5s fusion proteins. Abaxial epidermal tobacco (*Nicotiana benthamiana*) cells infiltrated with the plasma membrane marker FM4-64 transiently expressing mYFP:PpNIP5;1 (upper panel), mYFP:PpNIP5;2 (middle panel) or mYFP:PpNIP5;3 (lower panel). Bars, 20 µm for YFP:PpNIP5;1 and YFP:PpNIP5;3. Bars, 10 µm for YFP:PpNIP5;2. (b–d) *Atnip5;1* *Arabidopsis thaliana* growth complementation assay under B-limiting soil substrate conditions. *Arabidopsis Atnip5;1* knock-down plants expressing *PpNIP5;1*, *PpNIP5;2* or *AtNIP5;1* under the control of the *AtNIP5;1* promoter show a growth complementation on B-sufficient soil conditions compared to the *Atnip5;1* knock-down and display a similar growth than the *Arabidopsis Col-0* wild-type (b). Representative plant appearance at the reproductive growth stage (b), shoot fresh weight (FW) data (c) and shoot B concentrations (d) of 5-wk-old plant lines are displayed. Two independent transgenic *AtNIP5;1* *pro*:*PpNIP5;1* (#1 and #2) and *AtNIP5;1* *pro*:*PpNIP5;2* (#1 and #2) lines (T3 or T4 generation) have been assessed. Chart bars represent the means of shoot FW ($n = 7–10$ plants per line) or shoot B concentration ($n = 4–5$) data $\pm$ SD. Significant differences were determined using Student's *t*-test: **, $P < 0.01$; ***, $P < 0.001$.

Our results from the Si uptake assays, however, did not support this hypothesis. The neofunctionalisation of NIPs, namely to conduct Si, must have occurred later in evolution and was probably the driving force for the expansion of the NIP-III group in seed plants.

B is essential only for a few, but not for all, red, brown and green algae, and for diatom species (Carrano *et al.*, 2009). Constantly high concentrations of B in seawater (0.4 mM) suggests that B transport proteins may not be needed for B uptake into algae. Moreover, *Klebsormidium* species, including *K. nitens*, lack the cell wall polysaccharide RG-II, which represent the only characterised physiologically relevant B-binding molecules in green plants (Sørensen *et al.*, 2011; Domozych *et al.*, 2012). This suggests that charophytic NIPs are not needed for B uptake into charophytic algae.

With respect to As detoxification, *K. nitens* encodes an *S*-adenosylmethionine methyltransferase (AS3MT; equivalent to ArsM in bacteria), the main metabolic enzyme that methylates As in sequential steps for detoxification in chlorophyta, bacteria and humans (Palmgren *et al.*, 2017). Seed plants have lost AS3MT genes (Palmgren *et al.*, 2017). Methylated pentavalent As species such as methylarsonic acid (MMA(V)) and dimethylarsinic acid (DMA(V)) might therefore represent substrates for KnNIP6;1, as it is the case for the plant NIP OsLsi1 (Li *et al.*, 2009). *Klebsormidium nitens* also encodes an arsenate reductase and two ACR3 efflux transporters, which together represent an efficient arsenite efflux pathway and detoxification machinery. We suggest that the physiological role of KnNIP6;1 may predominantly be As efflux. NIPs or AqpNs have not been identified in sister species to land plants such as glaucophytes, red algae or chlorophytes, which are closely related to charophytes (this study; Anderberg *et al.*, 2011).

F–A–A–R-type plant NIP features suit the molecular demands to sub- and neofunctionalise into physiologically important metalloid channels

To fulfil an ancient role in metalloid uptake or efflux a localisation in the PM is necessary. Our analysis demonstrated that PpNIP5s, which are able to complement the *Atnip5;1* B-deficiency phenotype in Arabidopsis, are indeed localised to the PM (Fig. 5). Unlike seed plant NIPs, which respond to metalloid levels (Takano *et al.*, 2006; Tanaka *et al.*, 2011), *PpNIP5* transcript abundances were unresponsive to low B levels in the growth medium. Toxic externally applied arsenate concentrations did not result in an increased expression of any of the PpNIP5s, which would be expected for playing a role in detoxification processes (Fig. S4). Under highly toxic external arsenite concentrations, *P. patens* only formed protonema tissue in which all three *PpNIP5* transcript levels were very low also under control growth conditions (Chen *et al.*, 2012; Xiao *et al.*, 2011). These observations suggest that the F–A–A–R-type PpNIP5s have an impact on the metalloid status of *P. patens* but are not transcriptionally regulated by the tested metalloid treatments. The physiological function of individual PpNIP5 isoforms remains to be identified in future analyses.

F–A–A–R-type NIPs disappear in seed plants for which B and Si is nutritionally important

Bacterial type AqpN-like NIPs in liverworts, mosses, clubmosses and ferns with a F–A–A–R-type SF, remained existent up to the origin of vascular plants. In addition to F–A–A–R-type NIPs, the genomes of the liverwort *M. polymorpha*, the moss *P. patens*, the clubmoss *S. moellendorffii* and the fern *S. cucullata* possess typical NIPs (MA_158586g0010, MA_60111g0010, PHYPADRAFT_147365, SmNIP3;1 to SmNIP3;5, Sacu_v1.1_s0057.g014809, respectively), which cluster with B-allocating NIP-II group isoforms of seed plants (Fig. 2). By contrast, F–A–A–R-type NIP genes are absent in the published genomes of both angiosperms and gymnosperms. This suggests that F–A–A–R-type NIP transport selectivity and/or ability had no evolutionary advantage for these plant species. The demand for B significantly correlates with the amount of borate ester cross-linked RG-II found in the plants' pectin cell wall fraction (Matsunaga *et al.*, 2004; O'Neill *et al.*, 2004). Our data showed that the taxa-specific occurrence and proliferation of NIP-II and NIP-III isoforms, at the expenses of F–A–A–R-type-NIPs, either positively match RG-II levels of the corresponding plant taxa (Fig. 2) or the Si demand of plant taxa such as *Poaceae* or *Cucurbitaceae* species, respectively.

We tested whether the evolutionary switch from F–A–A–R-type to NIP-II-type SFs might have had an impact on pore selectivity characteristics influencing, for instance, As and B transport rates. Indeed, the mutagenesis approach corroborate the hypothesis, as it demonstrated that MIPs of bacteria and plants encoding the F–A–A–R-type SF have a higher permeability to As compared to their mutants encoding an A–I–G–R-type SF being intrinsic to typical seed plant B channels, such as AtNIP5;1. Moreover, the mutated variant, KrAqpN_{AIGR} of the *ars* operon-located native KrAqpN_{FAAR} had a relatively increased B permeability when it contained the SF residues of the physiological B channel AtNIP5;1_{AIGR}. The permeability of AtNIP5;1_{AIGR} to both As and B was reduced when mutated into AtNIP5;1_{FAAR}. This, in AtNIP5;1_{FAAR} decreased metalloid permeability was expected as another mutational study demonstrated that the native R1 position of the SF of AtNIP5;1 (A_{R1}-I_{R2}-G_{R3}-R_{R4}) is essential for its metalloid transport ability (Mitani-Ueno *et al.*, 2011).

Our mutational approach demonstrated (1) that *ars* operon-localised AqpN As channels encoding a F–A–A–R-type SF have a higher As transport capacity compared to their mutants encoding a NIP-II-type SF and (2) that physiologically important NIP-II B channels encoding a NIP-II-type SF have a higher B transport capacity compared to their mutants encoding a F–A–A–R-type SF. These results indicate that plant AtNIP5;1 and bacterial KrAqpN possess a SF pore layout which serves an optimal transport regulation of their physiological relevant substrate, namely boric acid and arsenous acid, respectively (Fig. 4a).

A high permeability to As in F–A–A–R-type NIPs might represent a physiological disadvantage for instance under metalloid toxic conditions. In the genomes of species with F–A–A–R-type MIPs such as bacteria, charophytes (*K. nitens*), mosses (*P. patens*), ferns (*Ceratopteris richardii*, *Pteris vittata*), and lycophytes

(*S. moellendorffii*) ACR3 genes are present. ACR3 proteins are PM or tonoplast-localised trivalent As antiporters which belong to the BART (bile/arsenite/riboflavin transporter) superfamily and are crucial for As detoxification. No ACR3 homologs have been identified in angiosperms yet (Mansour *et al.*, 2007; Indriolo *et al.*, 2010). Therefore, the occurrence of F–A–R-type NIP channels correlating with that of ACR3 transporters in plant taxa, suggests that F–A–R-type NIPs did not replace ACR3 efflux transporters in As detoxification processes in plants, while F–A–R-type AqpNs have replaced active As efflux transporters in bacterial *ars* operons. It is tempting to hypothesise that the presence of F–A–R type NIPs in plants necessitate additional efficient As detoxification mechanisms such as ACR3 transporters. The physiological function of individual F–A–R-type NIPs in old terrestrial plant lineages remains to be elucidated.

In summary, the transport functions and the phylogenetic and genetic context analyses are consistent and suggest that today's seed plant NIP function has evolved from bacterial As efflux channels and that amino acid changes in the SF were necessary but not sufficient for that change. In charophytic algae, the original role of ancestral NIPs may remain conserved. While the role of NIPs in mosses remains enigmatic, in seed plants SF alterations together with currently uncharacterised molecular channel changes converted bacterial As efflux proteins into essential plant nutrient transporters. The functional analyses strongly suggest that nutritional demands of terrestrial plants were a strong driver for functional divergence of NIPs. Plant NIP paralogues of AqpNs underwent subfunctionalisation by specialising on B transport regulation and specificity. The ancestral, however, physiologically insignificant B transport ability of NIPs of early land plants gained high importance when stable and flexible B-dependent cell wall properties became crucial for upright plant growth. In addition, neofunctionalisation of NIPs occurred when a few NIP-II (in horsetails) but in particular NIP-III group isoforms developed their ability to transport Si, which was not generally intrinsic to ancestral F–A–R-type NIPs.

Acknowledgments

We thank Dr Tandron Moja and Dr Eggert (IPK-Gatersleben, Germany) for ICP-MS analyses, and Dr Takano (University of Hokkaido, Japan) for the provision of the plasmid pP46 containing the *AtNIP5;1* promoter. We are thankful to Dr Finn for providing processable MIP sequences which have been used for the phylogenetic analyses (Finn & Cerdà, 2015). The authors declare no competing interest. This work was supported by the Deutsche Forschungsgemeinschaft (1668/1-1 to GPB), Grant-in-Aid for Specially Promoted Research (16H06296 to JFM), Belgian National Fund for Scientific Research to FC, the Belgian French community (ARC16/21-075 to FC) and the Bauchau Award to FC.

Author contributions

GPB conceived the project; BP, TAD and GPB designed research; BP, TAD, NB, JF, ABräutigam, MDB, CS, ABieber,

NM-U, JFM and GPB performed research; BP, TAD, NB, MDB, CS, ABräutigam, NM-U, JFM, FC and GPB analysed data; MDB, BP and GPB wrote the paper with the input of all authors.

ORCID

Nadine Bernhardt  <https://orcid.org/0000-0003-0716-6284>
Annett Bieber  <https://orcid.org/0000-0002-8173-7084>
Gerd P. Bienert  <https://orcid.org/0000-0001-9345-4666>
Manuela D. Bienert  <https://orcid.org/0000-0001-5948-3789>
Andrea Bräutigam  <https://orcid.org/0000-0002-5309-0527>
François Chaumont  <https://orcid.org/0000-0003-0155-7778>
Till A. Diehn  <https://orcid.org/0000-0001-6381-5256>
Jacqueline Fuge  <https://orcid.org/0000-0003-0312-1468>
Jian Feng Ma  <https://orcid.org/0000-0003-3411-827X>
Namiki Mitani-Ueno  <https://orcid.org/0000-0002-6840-2407>
Benjamin Pommerrenig  <https://orcid.org/0000-0002-7522-7942>
Christoph Spitzer  <https://orcid.org/0000-0002-6356-0393>

References

- Abascal F, Irisarri I, Zardoya R. 2014. Diversity and evolution of membrane intrinsic proteins. *Biochimica et Biophysica Acta* **1840**: 1468–1481.
- Anderberg HI, Danielson JA, Johanson U. 2011. Algal MIPs, high diversity and conserved motifs. *BMC Evolutionary Biology* **11**: 110.
- Anderberg HI, Kjellbom P, Johanson U. 2012. Annotation of *Selaginella moellendorffii* major intrinsic proteins and the evolution of the protein family in terrestrial plants. *Frontiers in Plant Science* **3**: 33.
- Bienert GP, Chaumont F. 2011. Plant Aquaporins: roles in water homeostasis, nutrition, and signaling processes. In: Geisler M, Venema K, eds. *Transporters and Pumps in Plant Signalling*. Berlin/Heidelberg, Germany: Springer, 3–38.
- Bienert GP, Desguin B, Chaumont F, Hols P. 2013. Channel-mediated lactic acid transport: a novel function for aquaglyceroporins in bacteria. *Biochemical Journal* **454**: 559–570.
- Bienert MD, Bienert GP. 2017. Plant aquaporins and metalloids. In: Chaumont F, Tyerman SD eds. *Plant aquaporins. From transport to signaling*. Berlin-Heidelberg, Germany: Springer, 297–332.
- Bonilla I, García-González M, Mateo P. 1990. Boron requirement in cyanobacteria: its possible role in the early evolution of photosynthetic organisms. *Plant Physiology* **94**: 1554–1560.
- Carrano CJ, Schellenberg S, Amin SA, Green DH, Küpper FC. 2009. Boron and marine life: a new look at an enigmatic bioelement. *Marine Biotechnology* (NY) **11**: 431–440.
- Chen YR, Su YS, Tu SL. 2012. Distinct phytochrome actions in nonvascular plants revealed by targeted inactivation of phytylbin biosynthesis. *Proceedings of the National Academy of Sciences, USA* **109**: 8310–8315.
- Chen J, Bhattacharjee H, Rosen BP. 2015. ArsH is an organoarsenical oxidase that confers resistance to trivalent forms of the herbicide monosodium methylarsenate and the poultry growth promoter roxarsone. *Molecular Microbiology* **96**: 1042–1052.
- Danielson JA, Johanson U. 2008. Unexpected complexity of the aquaporin gene family in the moss *Physcomitrella patens*. *BMC Plant Biology* **8**: 45.
- Danielson JA, Johanson U. 2010. Phylogeny of major intrinsic proteins. In: Jahn TP, Bienert GP, eds. *MIPs and their role in the exchange of metalloids*. Springer Landes Biosciences: Austin, USA, 19–31.
- Deshmukh R, Bélanger RR. 2016. Molecular evolution of aquaporins and silicon influx in plants. *Functional Ecology* **30**: 1277–1285.

- Diehn TA, Bienert MD, Pommerrenig B, Liu Z, Spitzer C, Bernhardt N, Fuge J, Bieber A, Richet N, Chaumont F *et al.* 2019. Boron demanding tissues of *Brassica napus* express specific sets of functional Nodulin26-like Intrinsic Proteins and BOR1 transporters. *The Plant Journal*. doi: 10.1111/tjp.14428.
- Domozych DS, Ciana M, Fangel JU, Mikkelsen MD, Ulvskov P, Willats WG. 2012. The cell walls of green algae: a journey through evolution and diversity. *Frontiers in Plant Science* 3: 82.
- Fetter K, Van Wilder V, Moshelion M, Chaumont F. 2004. Interactions between plasma membrane aquaporins modulate their water channel activity. *Plant Cell* 16: 215–228.
- Finn RN, Chauvigné F, Hlidberg JB, Cutler CP, Cerdà J. 2014. The lineage-specific evolution of aquaporin gene clusters facilitated tetrapod terrestrial adaptation. *PLoS ONE* 9: e113686.
- Finn RN, Cerdà J. 2015. Evolution and functional diversity of aquaporins. *Biological Bulletin* 229: 6–23.
- Fu D, Libson A, Miercke LJ, Weitzman C, Nollert P, Krucinski J, Stroud RM. 2000. Structure of a glycerol-conducting channel and the basis for its selectivity. *Science* 290: 481–486.
- Hori K, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, Seo M, Sato S, Yamada T, Mori H, Tajima N *et al.* 2014. *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. *Nature Communications* 5: 3978.
- Indriolo E, Na G, Ellis D, Salt DE, Banks JA. 2010. A vacuolar arsenite transporter necessary for arsenic tolerance in the arsenic hyperaccumulating fern *Pteris vittata* is missing in flowering plants. *Plant Cell* 22: 2045–2057.
- Jensen MO, Tajkhorshid E, Schulten K. 2001. The mechanism of glycerol conduction in aquaglyceroporins. *Structure* 9: 1083–1093.
- Kozono D, Ding X, Iwasaki I, Meng X, Kamagata Y, Agre P, Kitagawa Y. 2003. Functional expression and characterization of an archaeal aquaporin. AqpM from *Methanothermobacter marburgensis*. *Journal of Biological Chemistry* 278: 10649–10656.
- Lee JK, Kozono D, Remis J, Kitagawa Y, Agre P, Stroud RM. 2005. Structural basis for conductance by the archaeal aquaporin AqpM at 1.68 Å. *Proceedings of the National Academy of Sciences, USA* 102: 18932–18937.
- Li RY, Ago Y, Liu WJ, Mitani N, Feldmann J, McGrath SP, Ma JF, Zhao FJ. 2009. The rice aquaporin Lsi1 mediates uptake of methylated arsenic species. *Plant Physiology* 150: 2071–2080.
- Ma JF, Tamai K, Yamaji N, Mitani N, Konishi S, Katsuhara M, Ishiguro M, Murata Y, Yano M. 2006. A silicon transporter in rice. *Nature* 440: 688–691.
- Ma JF, Yamaji N. 2015. A cooperative system of silicon transport in plants. *Trends in Plant Science* 20: 435–442.
- Mansour NM, Sawhney M, Tamang DG, Vogl C, Saier MH. 2007. The bile/arsenite/riboflavin transporter (BART) superfamily. *FEBS Journal* 274: 612–629.
- Matsunaga T, Ishii T, Matsumoto S, Higuchi M, Darvill A, Albersheim P, O'Neill MA. 2004. Occurrence of the primary cell wall polysaccharide rhamnogalacturonan II in peridophytes, lycophytes, and bryophytes. Implications for the evolution of vascular plants. *Plant Physiology* 134: 339–351.
- Mitani N, Yamaji N, Ma JF. 2008. Characterization of substrate specificity of a rice silicon transporter, Lsi1. *Plant Archives* 456: 679–686.
- 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.
- Murata K, Mitsuoka K, Hirai T, Walz T, Agre P, Heymann JB, Engel A, Fujiyoshi Y. 2000. Structural determinants of water permeation through aquaporin-1. *Nature* 407: 599–605.
- O'Neill MA, Ishii T, Albersheim P, Darvill AG. 2004. Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. *Annual Review of Plant Biology* 55: 109–139.
- Palmgren M, Engström K, Hallström BM, Wahlberg K, Søndergaard DA, Säll T, Vahter M, Broberg K. 2017. AS3MT-mediated tolerance to arsenic evolved by multiple independent horizontal gene transfers from bacteria to eukaryotes. *PLoS ONE* 12: e0175422.
- Pao GM, Wu LF, Johnson KD, Höfte H, Chrispeels MJ, Sweet G, Sandal NN, Saier MH Jr. 1991. Evolution of the MIP family of integral membrane transport proteins. *Molecular Microbiology* 5: 33–37.
- Park JH, Saier MH Jr. 1996. Phylogenetic characterization of the MIP family of transmembrane channel proteins. *The Journal of Membrane Biology* 153: 171–180.
- Pommerrenig B, Diehn TA, Bienert GP. 2015. Metalloido-porins: Essentiality of Nodulin 26-like intrinsic proteins in metalloid transport. *Plant Science* 238: 212–227.
- Popper ZA, Michel G, Hervé C, Domozych DS, Willats WG, Tuohy MG, Kloareg B, Stengel DB. 2011. Evolution and diversity of plant cell walls: from algae to flowering plants. *Annual Review of Plant Biology* 62: 567–590.
- Reski R, Abel WO. 1985. Induction of budding on chloronemata and autonomata of the moss, *Physcomitrella patens*, using isopentenyladenine. *Planta* 165: 354–358.
- Roberts DM, Routray P. 2017. The nodulin 26 intrinsic protein subfamily. In: Chaumont F, Tyerman SD eds. *Plant aquaporins. From transport to signaling*. Berlin-Heidelberg, Germany: Springer, 267–296.
- Sørensen I, Pettolino FA, Bacic A, Ralph J, Lu F, O'Neill MA, Fei Z, Rose JK, Domozych DS, Willats WG. 2011. The Charophycean green algae provide insights into the early origins of plant cell walls. *The Plant Journal* 68: 201–211.
- Tanaka M, Takano J, Chiba Y, Lombardo F, Ogasawara Y, Onouchi H, Naito S, Fujiwara T. 2011. Boron-dependent degradation of NIP5;1 mRNA for acclimation to excess boron conditions in Arabidopsis. *Plant Cell* 23: 3547–3559.
- Trembath-Reichert E, Wilson JP, McGlynn SE, Fischer WW. 2015. Four hundred million years of silica biomineralization in land plants. *Proceedings of the National Academy of Sciences, USA* 112: 5449–5454.
- Takano J, Wada M, Ludewig U, Schaaf G, von Wirén N, Fujiwara T. 2006. The Arabidopsis major intrinsic protein NIP5;1 is essential for efficient boron uptake and plant development under boron limitation. *Plant Cell* 18: 1498–1509.
- Wallace IS, Roberts DM. 2004. Homology modeling of representative subfamilies of Arabidopsis major intrinsic proteins. Classification based on the aromatic/arginine selectivity filter. *Plant Physiology* 135: 1059–1068.
- Xiao L, Wang H, Wan P, Kuang T, He Y. 2011. Genome-wide transcriptome analysis of gametophyte development in *Physcomitrella patens*. *BMC Plant Biology* 11: 177.
- Yang Y, Wu S, Lilley RM, Zhang R. 2015. The diversity of membrane transporters encoded in bacterial arsenic-resistance operons. *PeerJ* 3: e943.
- Zardoya R, Ding X, Kitagawa Y, Chrispeels MJ. 2002. Origin of plant glycerol transporters by horizontal gene transfer and functional recruitment. *Proceedings of the National Academy of Sciences, USA* 99: 14893–14896.
- Zardoya R, Irisarri I, Abascal F. 2016. Aquaporin discovery in the genomic era. In: Søgaard G, Nielsen S, Casini A, eds. *Aquaporins in health and disease: New molecular targets for drug discovery*. Boca Raton, FL, USA: CRC Press, 19–32.
- Zardoya R. 2005. Phylogeny and evolution of the major intrinsic protein family. *Biology of the Cell* 97: 397–414.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Phylogeny of bacterial major intrinsic proteins (MIPs) and plant nodulin 26-like intrinsic proteins of the AQP-N-III clade.

Fig. S2 Additional AqpN sequences being located in arsenic resistance (*ars*) operons.

Fig. S3 PpNIP5;3 is incorrectly spliced in *Atnip5;1* Arabidopsis knock-down plants.

Fig. S4 Responses of *Physcomitrella patens* to different metalloid conditions in the growth substrate.

Dataset S1 Information on major intrinsic protein (MIP) sequences which have been used for the phylogenetic analyses of this study.

Methods S1 Detailed information on methods.

Table S1 Nucleotide sequences of codon-optimised *MIP* genes.

Table S2 Primers used for cloning of constructs.

Table S3 Primers used for RT-qPCR.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.



About New Phytologist

- *New Phytologist* is an electronic (online-only) journal owned by the New Phytologist Trust, a **not-for-profit organization** dedicated to the promotion of plant science, facilitating projects from symposia to free access for our Tansley reviews and Tansley insights.
- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as ready' via *Early View* – our average time to decision is <26 days. There are **no page or colour charges** and a PDF version will be provided for each article.
- The journal is available online at Wiley Online Library. Visit **www.newphytologist.com** to search the articles and register for table of contents email alerts.
- If you have any questions, do get in touch with Central Office (np-centraloffice@lancaster.ac.uk) or, if it is more convenient, our USA Office (np-usaoffice@lancaster.ac.uk)
- For submission instructions, subscription and all the latest information visit **www.newphytologist.com**