Characterising Tobacco Aquaporins:

towards engineering improved photosynthesis and plant performance

Annamaria De Rosa

July 2020

Thesis submitted for the degree of Doctor of Phylosophy of

The Australian National University

Declaration

This thesis contains research undertaken at the Australian National University, Canberra, ACT. Except where otherwise indicated, this thesis is my own original work performed under the supervision of Professor John R Evans, Dr Michael Groszmann and Professor Ulrike Mathesius.

Annamaria De Rosa

February 2020

Acknowledgements

This research was carried out at the Australian National University, Canberra, and financially supported by the ANU Postgraduate Research Scholarship.

I would like to thank my supervisory panel, John Evans, Michael Groszmann and Ulrike Mathesius for their support, encouragement and guidance throughout my PhD. I am extremely grateful to have had the priviledge to work with them, share their passion for science and to have had the opportunity to learn such a variety scientific techniques under their supervision.

I would like to acknowledge the moral and technical support I received from Aleu Mani-George and Baxter Massey, it was a pleasure to work with them in the lab. I would also like to thank Soumi Bala for the wonderful encouragements during my PhD and for training me to conduct my gas exchange measurements and running the tunable diode laser. I would like to acknowledge Daryl Webb for his remarkable assistance for using the confocal microscope, Rosemary White for prviding GFP marker lines, and Alex Watson-lazowski for his help on processing RNA-seq data.

I would like to also thank Susanne von Caemmerer, Robert Furbank, Caitlin Byrt and all their lab members. It has been great sharing ideas with them all over the years through our lab meetings, creating a collaborative and supportive research environment. I would like to thank Alisha Duncan for involving me in outreach projects through the ARC Centre of Excellence for Translational Photosynthesis and for sharing her passion for science communication.

I am extremely thankful for the amazing friends I have made at ANU, in particular, I would like to thank Sally Buck and Lily Chen for their support, numerous coffee breaks and writing sessions together.

I would like to also thank my friends outside of ANU who are always there for me, and my wonderful family, always supporting and encouraging me to do my best. Finally, thank you to Tristan, I couldn't have asked for a more supportive and caring partner during my PhD journey.

Abstract

Future climatic conditions and an increasing global world population require the development of higher yielding and more resilient crops. Aquaporins are increasingly being studied as genetic engineering targets to tackle food security challenges. They constitute a major family of membrane spanning channel proteins, selectively facilitating the passive bidirectional passage of a range of solutes essential for numerous plant processes, including water relations, growth and development, stress responses, root nutrient uptake, and photosynthesis. In plants, aquaporins occur in five major subfamilies that differ in temporal and spatial gene expression, subcellular protein localisation, substrate specificity, and post-translational regulatory mechanisms, collectively providing a dynamic transportation network spanning the entire plant. Of particular interest are aquaporins in the Plasma membrane Intrinsic Proteins (PIP) subfamily, some of which have been shown to enhance membrane permeability to CO_2. This role could influence photosynthetic efficiency, which is limited by the ease with which CO_2 diffuses across cellular membranes from intercellular airspaces to the site of fixation within the chloroplast (i.e. mesophyll conductance). As such, PIP aquaporins are an attractive target for engineering enhanced photosynthesis. The ability to manipulate aquaporins towards improving plant productivity is reliant on expanding our insight into their diversity and functional roles.

My PhD project contributes towards a better understanding of aquaporin biology, providing a comprehensive overview of the *Nicotiana tabacum* (tobacco) aquaporin family. Tobacco is a popular model system capable of scaling from the laboratory to the field, it is closely related to several major economic crops (e.g. tomato, potato, eggplant and capsicum) and itself has new commercial applications. To date one PIP gene has been described in tobacco to be a CO_2 pore, NtAQP1. However, beyond this there is little characterization of tobacco aquaporins in the literature.

We identified that the tobacco genome encodes 76 aquaporins, making it the second largest characterised aquaporin family after that of *Brassica napus,* containing 121 genes*.* Tobacco aquaporins fall into five distinct subfamilies, for

which we characterised phylogenetic relationships, gene structures, protein sequences, selectivity filter compositions, sub-cellular localisation, and tissue-specific expression.

Once the tobacco aquaporin family was established, we functionally characterized nine tobacco AQPs occurring in the PIP, TIP and NIP subfamilies. We determined their sub-cellular localisation *in planta* and their substrate specificities through yeast-based functional assays developed within our laboratory. These yeast-based assays allowed us to test aquaporin permeability for a range of substrates important for plant function (water, hydrogen peroxide, urea and boron). 3D protein homology modelling was then used for an integrated aquaporin characterisation, linking substrate specificities and amino acid primary sequence to the pores' radius and physico-chemical properties.

Towards enhancing photosynthesis, we identified several PIP aquaporins that are likely candidates for transporting CO_2 and constitutively over-expressed these in tobacco. Gas exchange measurements showed gene-specific alterations to photosynthesis. Although enhanced photosynthetic rates and photosynthetic parameters (V_{cmax} and J) were observed with PIP1 over expression, no significant increase in mesophyll conductance was observed. Further investigations are needed to re-examine the consequence of overexpression of NtAQP1 on mesophyll conductance and clarify the possible pleiotropic effects that constitutive over-expression of tobacco PIPs has on leaf properties and plant growth.

Table of Contents

List of Figures

List of Tables

Abbreviations and symbols

A	Assimilation rate
AQP	Aquaporin
Ar/R	Aromatic/Arginine
AUC	Area under the curve
C_a	CO_2 concentration or partial pressure in ambient air
C_c	CO_2 concentration or partial pressure in the chloroplast
C_i	CO_2 concentration or partial pressure in intercellular air spaces
C_a-C_i	drawdown associated with stomatal conductance
C_i-C_c	drawdown associated with mesophyll conductance
CDS	Coding DNA Sequence
CFU	Colony forming units
ER	Endoplasmic Reticulum
GFP	Green Fluorescent Protein
g_m	mesophyll conductance
g_s	stomatal conductance
HKG	House Keeping Gene
H_2O_2	Hydrogen peroxide
J_{1500}	potential electron transport rate at 1500µmol photons $m^{-2}s^{-1}$
LMA	Leaf Mass per Area, g m^{-2}
MIP	Major Intrinsic Protein
NIP	Nodulin26-like Intrinsic Protein
NPA	Asparagine-Proline-Alanine

N. syl *Nicotiana sylvestris*

N. tom *Nicotiana tomentosiformis*

Nt *Nicotiana tabacum*

OD Optical density

OE Over Expression

PIP Plasma membrane Intrinsic Protein

PM Plasma membrane

ROS Reactive Oxygen Species

Sl *Solanum lycopersicum*

St *Solanum tuberosum*

TIP Tonoplast Intrinsic Protein

TM Transmembrane domain

Tono tonoplast

V_{cmax} maximum velocity of Rubisco carboxylation

XIP X Intrinsic Protein

$\delta^{13}C$ carbon isotope composition

Δ_i theoretical Delta, assuming $C_i = C_c$

Δ_o observed Delta

Chapter 1: General Introduction

1.1 Food security challenges

Unprecedented increases in crop yields are required in order to feed a projected world population of 9 billion people by the year 2050 (Ray et al. 2013). It is currently estimated that nearly ~12.5% of the global population are suffering from chronic hunger, with numbers on the rise (FAO 2009). In addition to population increases, global climate change will further challenge our efforts in ensuring sufficient food production, due to increasing temperatures and altered patterns of more erratic rainfall (Parry and Hawkesford 2010).

Trends in global crop yield (solid lines, Figure 1.1) are projected to increase at a rate of 0.9%-1.6% per year for key agricultural crops: maize, rice, wheat and soybean. In order to meet the requirement of doubling annual crop yields by the year 2050, a 2.4% annual yield increase is needed (indicated by dashed lines, Figure 1.1) (Ray et al. 2013). Previous major advances in agricultural production were achieved in the 1960's, with the 'green revolution' bringing about yield increases in major crops, predominantly through dwarfing which increased biomass partitioning to the grain (Khush 2001). Additionally, the expansion of irrigation structures, modernisation of farming techniques, utilisation of hybrid seed stocks, synthetic nitrogen fertilisers and pesticides, collectively facilitated a massive increase of crop yields globally. Since then, we have observed a plateauing of yield increases of food crops, with reports stating that crop yields are no longer improving on 24-39% of our most important crop land areas (Ray et al. 2012). In order to drastically increase crop yields and food production, innovations in genetic engineering are required, enabling a new 'green revolution' in agricultural productivity (Georges and Ray 2017).

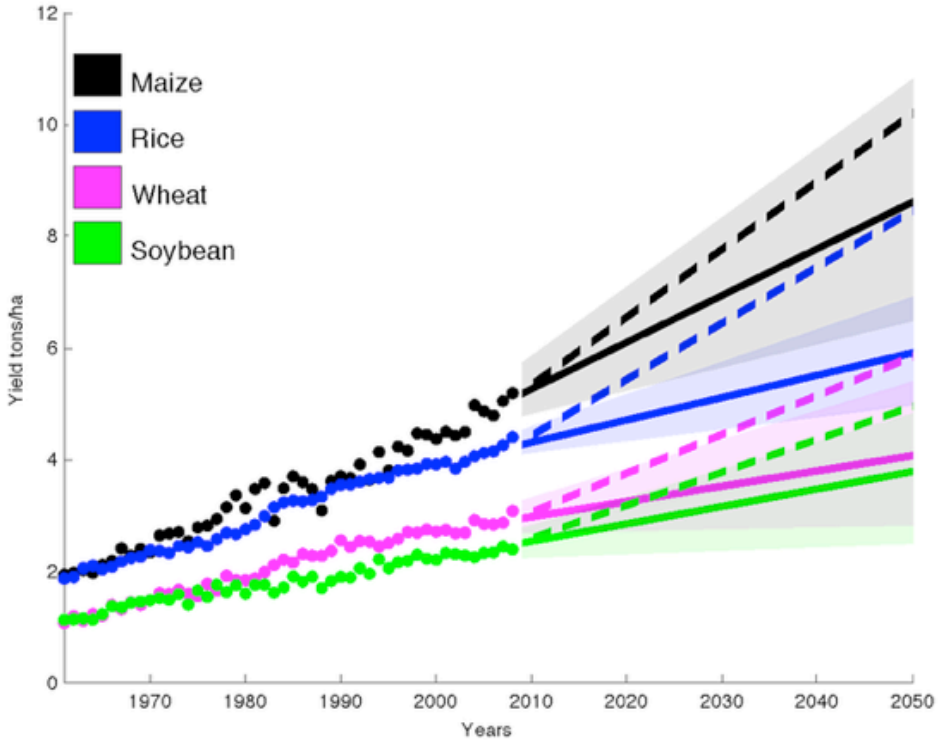


Figure 1. 1. Global projection of yield increases of major crop species to 2050. *Observed area-weighted global yield (closed circles) for maize (black), rice (blue), wheat (magenta) and soybean (green) between 1961-2008. Solid lines indicate yield projection to 2050. Dashed lines represent the annual ~2.4% yield improvement trend required to double crop production by 2050 (without an increase in cultivated agricultural land)(Ray et al. 2013) .*

1.2 Increasing biomass production through increased photosynthetic rates

Increasing agricultural productivity in order to address future food demands is reliant on the production of higher yielding crops. Crop yield can be described as a product of biomass and harvest index (ratio of grain yield to above ground biomass). Since the 'green revolution', the crop harvest index is thought to have reached a maximum biological limit, so researchers are now focusing on increasing crop biomass production in order to increase yield (Fischer et al. 2014).

Photosynthesis is thought to be one of the key remaining traits available to target for improvement in plants to increase crop yield, with an overwhelming weight of evidence from plants grown at elevated atmospheric carbon dioxide (CO_2) concentrations showing a link between increased photosynthesis and biomass production (Long et al. 2006). Photosynthesis is a vital process that sustains nearly

all life on earth and its efficiency can be limited by the diffusion of atmospheric CO_2 into the chloroplast (Evans et al. 2009). Stomatal conductance (g_s) to CO_2 enables the diffusion of ambient CO_2 (C_a) into the intercellular air spaces. Intercellular CO_2 (C_i) then diffuses into mesophyll cells to reach the chloroplast (C_c). The ease with which CO_2 is able to diffuse from the intercellular airspaces to the chloroplast can be referred to as the mesophyll conductance (g_m). Mesophyll conductance is one of the three main physiological processes limiting the uptake and fixation of CO_2 for photosynthesis, along with stomatal conductance and biochemical capacity (von Caemmerer and Evans 1991; Flexas et al. 2012). Mesophyll conductance can be affected by the surface area of chloroplasts exposed to intercellular airspaces, cell wall thickness, the permeability of the plasma membrane to CO_2 and cytosolic and stromal resistances (Evans et al. 1994; Terashima et al. 2011; Tholen and Zhu 2011).

A strategy to increase photosynthetic efficiency could be to facilitate diffusion of CO_2 into the chloroplasts by enhancing mesophyll conductance, as the increased partial pressure of CO_2 in the chloroplast would increase CO_2 fixation rates (Evans et al. 1994; von Caemmerer and Evans 2010). The cell wall and membranes restrict the ease with which CO_2 is able to diffuse into the chloroplast. It was traditionally thought that because CO_2 is lipophilic, it could readily pass through the lipid bilayer to meet CO_2 fixation rates in the chloroplast. However, biological membranes have high protein and sterol content, reducing their permeability for gas diffusion (Engelman 2005; Evans et al. 2009). In recent years, plant AQPs have been implicated with increasing membrane permeability to CO_2, suggesting that CO_2 diffusion across the plasma membrane of mesophyll cells could be possible through two parallel pathways, bulk membrane flow as well as through CO_2-permeable AQPs (Evans et al. 2009).

1.3 Facilitated diffusion through Aquaporins

Cellular membranes are dynamic structures, continuously adjusting their composition in order to allow plants to respond to developmental signals, stresses, and changing environments (Marschner 2011). The biological function of cell membranes is conferred by its protein composition, with the lipid bilayer providing

a basic structure and permeability barrier, and integral transmembrane proteins facilitating diffusion of selected substrates (Marschner 2011). Cell membrane diffusion is a fundamental process of biology and one of the oldest subjects studied in plant physiology (Hedrich and Marten 2006). Diffusional events at the cellular level eventuate in the coordinated transport of substrates throughout the plant to support development and growth.

Plant membranes contain three major classes of transport proteins known as ATP-powered pumps, transporters, and channel proteins (Chrispeels et al. 1999). Pumps, are active transporters that use the energy of ATP hydrolysis to move substrates across the membrane against a concentration gradient or electrical potential. Transporters move a variety of molecules across a membrane along or against a gradient at rates of 10^2 to 10^4 molecules per second. Unlike the first two classes, channel proteins are bidirectional and increase membrane permeability to a particular molecule. Channel proteins are permeable to a wide range of substrates and up to 10^8 molecules per second can pass through them.

In plants, aquaporins (AQPs) constitute a major family of such channel proteins that facilitate selective transport of substrates for numerous biological processes including, water relations, plant development, stress responses, and photosynthesis (Hachez et al. 2006; Groszmann et al. 2017). The AQP monomer forms a characteristic hour-glass membrane-spanning pore that assembles as tetrameric complexes in cell membranes. The union of the four monomers, creates a fifth pore at the centre of the tetramer which may provide an additional diffusional path (Frick et al. 2013).

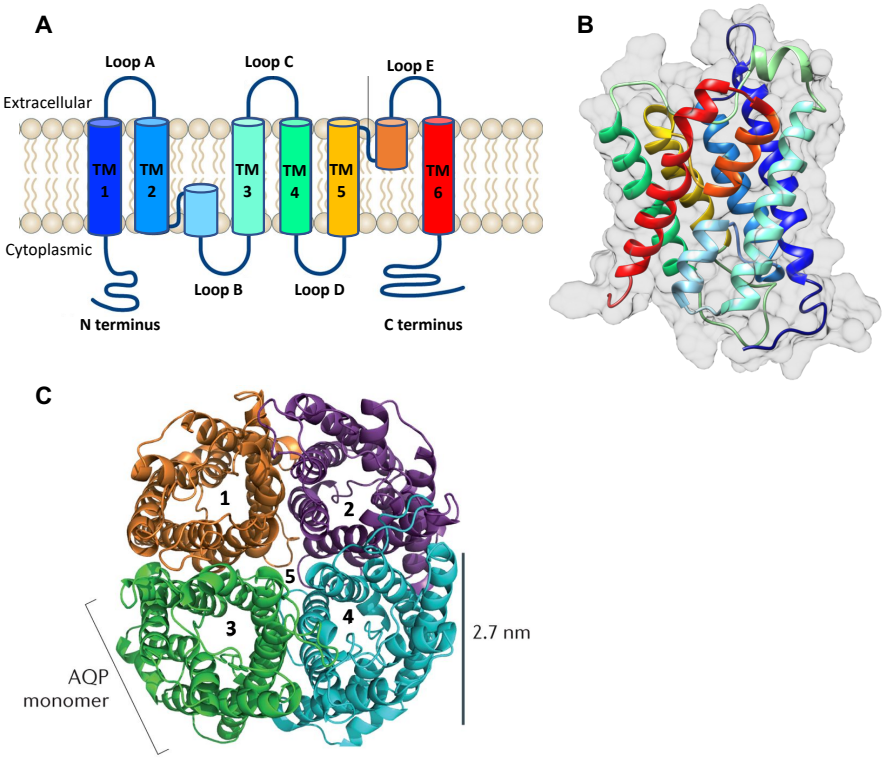


Figure 1. 2. **Schematic and 3D representation of an AQP monomer, and representation of an AQP tetramer. A.** AQP monomers have a highly conserved structure, comprising of six transmembrane helices (TM 1-6, Blue to Red), 5 connecting Loops (Loop A-Loop E) and cytoplasmic N- and C terminus. Loops B and Loop E form half helices folding into the membrane. **B.** 3D structure of SpinachPIP2;1 (SoPIP2;1, PDB:2b5f1.A) monomer, showing arrangement of transmembrane helical domains and connecting loops. **C.** Top view of the extracellular face of an AQP tetramer. Individual monomer pores are labelled 1-4. An additional central pore, pore 5, occurs upon tetramer assembly, providing an extra cavity for permeation of substrates. Panels A and C were adapted from (Verkman et al. 2014).

The substrate specificity of a given AQP is conferred by the complement of pore lining residues which achieve specificity through a combination of size exclusion and biochemical interactions with substrates (Hove and Bhave 2011). Key residues that have been identified to confer specificity include the dual Asn-Pro-Ala (NPA) motifs, the aromatic/Arginine filter (ar/R filter) and Froger's positions (P1-P5) (Froger et al. 1998; Mitani-Ueno et al. 2011; Murata et al. 2000). However, other pore-lining residues and lengths of the various transmembrane and loop domains of the AQP monomer are also known to influence substrate specificity through conformational changes of the pore size and accessibility (Wu and Beitz 2007; Hove

and Bhave 2011). It is likely that other residues that determine specificity and transport efficiency remain to be elucidated.

1.3.1 Diversification of Aquaporins in land plants

Aquaporins, which are members of the major intrinsic proteins (MIP) superfamily, are found across all taxonomic kingdoms (Abascal et al. 2014). In land plants, AQPs are by far the most extensively diversified. AQP diversification in land plants is facilitated by their propensity for gene duplication events (Groszmann et al. 2017). Plant AQPs are capable of transporting a wide variety of substrates including water, ammonia, urea, carbon dioxide, hydrogen peroxide, boron, silicon and other metalloids (Gomes et al. 2009; Pommerrenig et al. 2015; Hove and Bhave 2011). More recently, lactic acid, oxygen, and cations have been identified as permeating substrates (Choi and Roberts 2007; Zwiazek et al. 2017; Byrt et al. 2017; Bienert et al. 2013), with RNA molecules also implicated as a possible transported substrate (Reichel et al. 2016).

Plant AQPs are divided into five phylogenetically distinct sub-families and further into sub-groups; Plasma membrane Intrinsic Proteins (PIPs), Tonoplast Intrinsic Proteins (TIPs), Small basic Intrinsic Proteins (SIPs), Nodulin 26-like Intrinsic Proteins (NIPs), and X Intrinsic Proteins (XIPs) (Danielson and Johanson 2008; Johanson and Gustavsson 2002; Kaldenhoff and Fischer 2006). The sub-families differ to some degree in substrate specificity and integrate into different cellular membranes, providing plants with a versatile system for both sub-cellular compartmentalisation and intercellular transport, as illustrated in Figure 1.3 (Maurel et al. 2008). Further versatility in AQP function is achieved through tightly regulated spatial and temporal tissue-specific expression of different *AQP* genes, as well as post-translational modification of AQP proteins (e.g. phosphorylation) that controls membrane trafficking and channel activity (Santoni 2017; Luu and Maurel 2013).

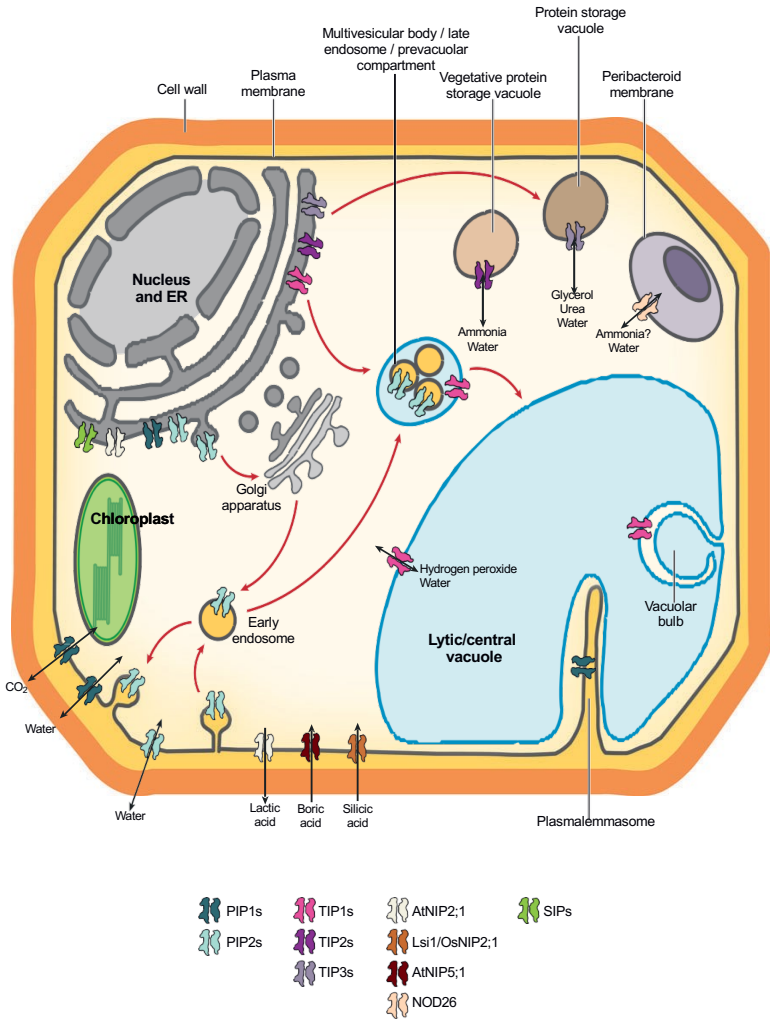


Figure 1. 3 **Multiple subcellular localisations and substrate specificities of plant AQPs.** Different AQP subclasses (illustrated by distinct colours) integrate in various membranes and subcellular compartments such as the plasma membrane, central vacuole membrane (tonoplast), endoplasmic reticulum (ER), protein storage vacuoles and endosomes. Permeating substrates include water, CO_2, hydrogen peroxide, lactic acid, glycerol, boric acid, silicic acid, ammonia and urea. Shuttling of substrates throughout the cell is coordinated via selective AQP-facilitated diffusion. Red arrows indicate the movement of AQPs through sub-cellular compartments through the secretory pathway (from the ER to the plasma membrane, via the golgi apparatus). Red arrows also illustrate the recycling of PIPs through repeated cycles of endocytosis, before being eventually targeted to the lytic vacuole**.** Figure from Maurel *et. al* (2008).

1.4 Aquaporins' potential to enhance photosynthetic efficiency

Aquaporins in the PIP subfamily have been associated with enhancing photosynthetic efficiency in C_3 plants due to their ability to permeate CO_2 across biological membranes; as illustrated in Figure 1.4 (von Caemmerer and Evans 2010). Tobacco AQP1 (NtAQP1), is part of the PIP1 subgroup which has been studied in heterologous expression systems and found to be CO_2 permeable (Uehlein et al. 2003; Otto et al. 2010). NtAQP1 has also been studied *in planta*, where RNAi knockdown and over expression experiments in tobacco plants altered net photosynthetic rate, consistent with NtAQP1 facilitating CO_2 diffusion across membranes of mesophyll cells (Flexas et al. 2006; Uehlein et al. 2008). Other PIP AQPs which have been implicated in CO_2 diffusion include barley PIP2;1 (Hanba et al. 2004), arabidopsis PIP1;2 (Uehlein et al. 2012; Heckwolf et al. 2011), ice plant PIP1 (McMIPB) (Kawase et al. 2013) and rice PIP1;2 (Xu et al. 2018). For a list of characterised CO_2-permeable AQPs, see Table 1 in Groszman et al. 2017. As such, AQPs could be ideal targets in attempts to enhance photosynthetic efficiency through increased CO_2 diffusion into the chloroplast (Flexas et al. 2006; Hanba et al. 2004; Uehlein et al. 2003; Uehlein et al. 2008; Uehlein et al. 2012).

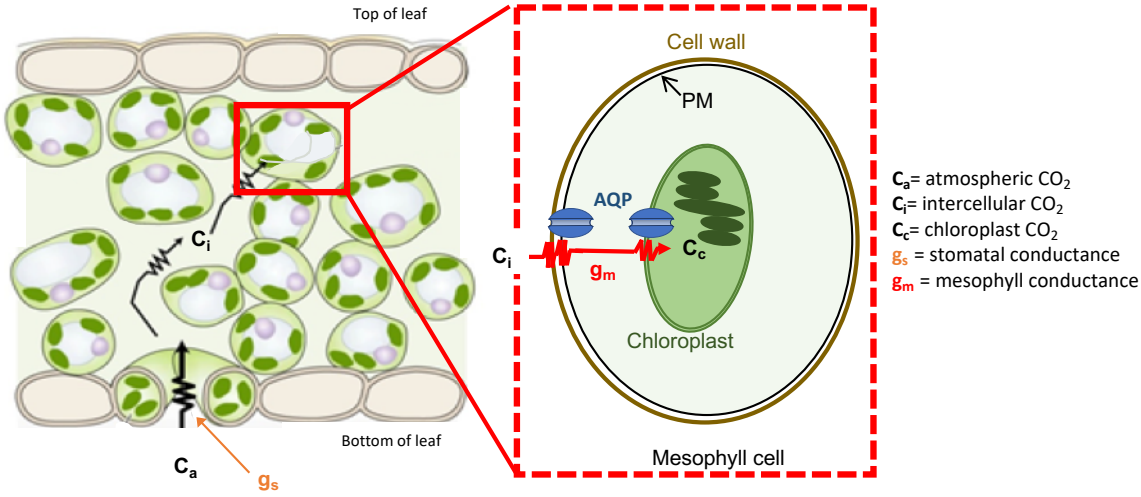


Figure 1. 4 Schematic representation of CO_2 diffusion into the chloroplast. *Atmospheric CO_2 (C_a) enters the leaf through the stomata. Stomatal conductance (g_s) is the ease with which CO_2 can diffuse through stomata. Intercellular CO_2 (C_i) diffuses through the intercellular airspaces within leaves. Mesophyll conductance (g_m) is the ease with which CO_2 is able to diffuse into the chloroplast (C_c); crossing the mesophyll cell wall, plasma membrane and chloroplast envelope to reach the site of its*

fixation by Rubisco. Aquaporins (AQPs) integrate into the plasma membrane (PM) and chloroplast envelope and facilitate diffusion of CO_2 into the chloroplast.

1.5 Aquaporin applications for improving other aspects of plant performance

Given their diverse complement of transported substrates and growing involvement in many developmental and stress responsive physiological roles, AQPs are increasingly becoming targets for engineering more resilient and productive plants (Li et al. 2014; Groszmann et al. 2017).

1.5.1 AQP-enhanced resilience to abiotic stresses

Water availability has been classed as the major constraint on world crop productivity (Baldocchi and Valentini 2004). As such, an immediate challenge for agriculture is increasing crop yields in water limited conditions. Engineering drought-tolerance in crops is quite complex due to the wide range of specific mechanisms regulating water balance. Potential processes could include improved hydraulic conductance, altered root architecture, increased sensitivity to water-stress and closure of stomata. AQPs are central players in plant water relations, with PIPs and TIPs largely mediating water uptake and transcellular water flow in the roots (Afzal et al. 2016). The over expression of AQP isoforms in plants has been shown to enhance plant resistance to drought through improved water use efficiency (WUE), hydraulic conductivity and retaining a better water status (Lian et al. 2004; Sade et al. 2010; Cui et al. 2008; Li et al. 2008b). In addition to having improved water status, over expressing a wheat PIP2 (*TaAQP7*) in tobacco plants resulted in the reduction of ROS accumulation and membrane damage via enhancing activities and expression of antioxidant enzymes, during exposure to drought conditions (Zhou et al. 2012). As well as drought, soil salinity can have negative effects on plant growth and development, severely disturbing the cell's osmotic balance. AQP gene expression has been shown to alter in response to salt stress, implicating them in salinity stress responses (Afzal et al. 2016). Constitutive over expression of particular AQP isoforms has provided improved tolerance to plants grown in salt stressed conditions. For example, over expression of rice PIP1;1

(*OsPIP1;1*) gene in rice increased germination of seeds under salt stress and improved tolerance of roots and leaf growth to salt stress (Liu et al. 2013).

1.5.2 AQP roles in micronutrient homeostasis

AQPs have also been implicated in nutrient uptake and redistribution within the plant. PIP, TIP and NIP isoforms have been reported to transport nitrogen compounds (ammonia and urea), and essential micronutrients such as boron and silicon (Gao et al. 2018). Studies have shown AQPs play an important role in reducing accumulation of micronutrients to toxic levels within plant tissues (Wang et al. 2016). Boron is an essential plant micronutrient and its homeostatic regulation is crucial. Exposure of plants to either excessive or insufficient boron results in significant reduction and quality of crop yields. AQP isoforms permeable to boron have been shown to be essential for plant growth under B limiting conditions (Takano et al. 2006; Tanaka et al. 2008). Conversely, increased tolerance to boron toxicity was observed when expression of boron-permeable Barley NIP2;1 (HvNIP2;1) was reduced in barley plants (Schnurbusch et al. 2010).

1.5.3 AQP involvement in pathogen interactions

In addition to the above-listed physiological roles, AQPs have been linked with mediating plants' defence responses to pathogens. Plant responses upon attack by pathogens include dehydration to regulate plant water relations and nutrient homeostasis, and the production of Reactive Oxygen Species (ROS, e.g. H_2O_2) (Afzal et al. 2016). Current literature implicates AQPs in plant immunity, due to their altered expression upon pathogen infection (Tian et al. 2016; Martins et al. 2015). Although further studies are required to obtain a full understanding of these AQP defence response pathways (Hooijmaijers et al. 2012), modified AQPs could be engineered into crop species in order to improve their defence response to pathogens, thereby reducing potential yield losses.

1.6 Deciphering Aquaporin function using heterologous expression systems

Aquaporins are involved in numerous processes within plants. Their physiological role can be controlled through spatial and temporal regulation of expression (Bots et al. 2005) and through interactions with other monomers forming the AQP tetramer (Otto et al. 2010). Furthermore, AQPs can be regulated post-translationally through gating mechanisms and interactions with other proteins (Tournaire-Roux et al. 2003; Törnroth-Horsefield et al. 2006; Roche and Törnroth-Horsefield 2017). Although several regulatory factors (e.g. localisation of gene expression, protein interactions and gating) can impact AQP functioning, identifying permeating substrates is pivotal to deciphering an AQP's biological role within the plant.

Teasing apart the complexities of aquaporin biology *in planta* can be difficult. Plants have many AQP isoforms, with some having redundancy of function under certain environmental conditions (Abascal et al. 2014; Fox et al. 2017). As such, modification of aquaporin expression within the plant, via over expression or down regulation of a specific AQP, may also affect the expression of closely related isoforms (Bi et al. 2015; Kaldenhoff et al. 2007). Plant protoplasts (cells lacking a cell wall) offer a robust *in planta* expression system for functional studies, being a single-cell alternative, removed from the complications arising from multicellularity in plants. (Yoo et al. 2007). Limitations around using plant protoplasts for functional studies include the laborious protoplast isolation procedure, with low protoplast yields for screening. Additionally, functional redundancy through native AQPs present within the protoplast membrane could impair our ability to accurately decipher permeating substrates.

Functional characterisation of plant AQPs is often accomplished through heterologous expression systems (Kaldenhoff et al. 2007), allowing for assignment of substrate specificity to an AQP monomer in isolation of other isoforms and external regulatory mechanisms. This can further our understanding of AQP pore biology through elucidation of which functional sites might act as determinants for pore selectivity.

Xenopus laevis oocytes can be used as a heterologous expression system for the study of AQPs. Although the use of *Xenopus* oocytes requires stringent ethics

approvals, and highly demanding technical skills, they are often considered a valuable system for the study of membrane membrane proteins due to the relatively large size of the egg and very low intrinsic membrane permeability through endogenous transporters. This results in limited background effect and less ambiguity in interpretation of permeating solutes (Deshmukh et al. 2016; Maurel et al. 1993). When water permeable plant AQPs are injected and expressed in the oocytes, their membrane permeabilities have been reported to increase 10- to 50-fold (Kaldenhoff et al. 2007).

Similarly to *Xenopus* oocytes, yeast (*Saccharomyces cerevisiae*) also has a low intrinsic membrane water permeability, making it another commonly used heterologous expression system for the functional characterisation of membrane transport proteins (Kaldenhoff et al. 2007). Their ease of growth, short generation time and well established transformation system, makes yeast a convenient and affordable option for studying AQPs (Deshmukh et al. 2016; Dreyer et al. 1999). A variety of experimental approaches have been used for the functional characterisation of transmembrane proteins in yeast, such as electrophysiology, tracking uptake of radiolabelled substrates, and the use of growth and survival assays (Bertl et al. 1995; Frommer et al. 1993; Riesmeier et al. 1992). Yeast can be a favourable system to use for AQP functional characterisation due to its ease of accessibility and/or availability compared to *Xenopus* oocytes, which can be technically difficult to work with and a more limited resource. Additionally, scale of replication is another advantage for using yeast as an heterologous expression system for functional characterisation of AQPs, with each yeast culture containing a starting cell number in the millions, as opposed to the 4-30 replicates often used in Xenopus oocyte experiments (Wang et al. 2018; Le Cahérec et al. 1996).

1.7 Tobacco as a model species for studying Aquaporins

The genomic era of plant biology has provided unprecedented opportunity to query AQP biology by exploring sequence diversity between isoforms in many species. Analysis between closely related species is especially advantageous as it allows close orthologous AQPs to be compared and contrasted for protein structural

features, tissue-specific expression patterns and sub-cellular localisation, all of which help towards elucidating physiological roles and regulation towards potential use in engineering efforts.

Nicotiana tabacum (tobacco) is a fitting candidate species to conduct such comparisons. Tobacco is a popular model system due to its suitability for experiments studying fundamental physiological processes (scaling from the cell level to the field). As both the nuclear and chloroplast genome can be genetically engineered, it is used as a model plant in biotechnology and disease susceptibility research and has renewed commercial applications in the biofuel and plant-made pharmaceutical sectors (Vanhercke et al. 2014; Tusé et al. 2014; Ma et al. 2015). Tobacco is part of the large family of Solanaceae, harbouring species of major economic importance such as tomato, potato, eggplant and capsicum (Gebhardt 2016). The AQP protein families of tomato (*Solanum lycopersicum*) and potato (*Solanum tuberosum*) have been established, each containing 45 and 47 AQP genes respectively (Reuscher et al. 2013; Venkatesh et al. 2013). The AQP families in tomato and potato that have already been characterised could facilitate cross-species comparison of sequences, expression profiles and function of various AQP isoforms in tobacco.

Also of interest is tobacco's recent evolutionary history. Tobacco is a complex allotetraploid with a large genome (4.5Gb), having evolved from an interspecific hybridisation of ancestors of *Nicotiana sylvestris* and *Nicotiana tomentosiformis* from distant clades within Nicotiana genus about 0.2M years ago (Sierro et al. 2014; Edwards et al. 2017; Leitch et al. 2008). The short time frame for redundant gene loss would result in tobacco still retaining two copies of AQP gene orthologs from each parent. Furthermore, the combination of parental genomes from phylogenetically separated sections within the Nicotiana genus, would enable us to trace the heritage of tobacco AQPs to their parental genomes and infer on gene family evolution processes (Leitch et al. 2008).

1.8 Project overview

Increases in crop yields are essential in order to address global food security challenges. Aquaporins are a protein family regulating numerous plant

physiological processes, making them useful genetic engineering targets to enhance plant photosynthesis and plant performance. Our ability to effectively engineer plant improvement relies on advancing our understanding of AQP biology and their regulation within plants. This requires studies which accurately elucidate AQP substrate specificities, and also investigates their sub-cellular and gene expression localisations, potential post-translational regulations (e.g. gating) and AQP tetramer interactions.

My PhD project aims to:

- Expand current knowledge of AQP biology, firstly by characterising the AQP family in tobacco, a popular model species in plant biology. The establishment of the tobacco AQPs family is a valuable resource, providing phylogenetic relationships, gene structures, protein sequences, selectivity filter compositions, identification of potential post-translational regulation sites and tissue-specific expression for all 76 tobacco AQP genes. The study of tobacco AQPs also allows us to draw comparisons with gene orthologs found in tomato and potato, closely related species within Solanaceae (Chapter 2).
- Develop high throughput and replicable functional assays in yeast in order to functionally test substrate specificities of candidate AQP genes. This will enable us to examine a range of permeating substrates (water, H_2O_2, boron and urea) to further our knowledge of AQP biology and potentially to engineer more resilient and higher yielding crops. I was involved with the development of the H_2O_2 and urea yeast-based functional assays (Chapter 3).
- Functionally characterise nine AQP isoforms occurring in the PIP, NIP and TIP subfamilies. The functional characterisation includes elucidating their substrate specificities through yeast-based functional assays (developed in Chapter 3) characterising their sub-cellular localisation *in planta* and conducting 3D homology modelling to link their substrate specificities and amino acid primary sequence to their pore's radius and

physico-chemical properties (Chapter 4). *Arabidopsis thaliana* was utilised as an *in planta* expression system for the characterisation of PIP, NIP and TIP sub-cellular localisations due to its ease of transformation and short generation time. This enabled us to compare membrane integration between these AQP sub-families with publicly available Arabidopsis GFP marker lines (for plasma membrane, tonoplast and ER integration). Constitutively over express several PIP1 isoforms with high homology to a characterised CO_2 pore, NtAQP1. These were identified to be likely candidates for transporting CO_2 and hypothesised to enhance photosynthetic efficiency through increases in mesophyll conductance. ., Gas exchange coupled to carbon isotope discrimination were measured on these overexpressing plants and used to describe changes in photosynthetic properties of their leaves and to calculate mesophyll conductance (Chapter 5).

Chapter 2: Characterisation of the tobacco AQP family

This chapter was submitted as a manuscript to BMC Plant Biology journal (published 9 June 2020). As such, there is some repetition of content in the 'Background' section and the introductory chapter (Chapter 1) of this thesis.

Genome-wide identification and expression analysis of Aquaporins in *Nicotiana tabacum* and their relationships with other Solanaceae species

2.1 Abstract

Background

Cellular membranes are dynamic structures, continuously adjusting their composition in order to allow plants to respond to developmental signals, stresses, and changing environments. To facilitate transmembrane transport of substrates, plant membranes are embedded with both active and passive transporters. Aquaporins (AQPs) constitute a major family of membrane spanning channel proteins that selectively facilitate the passive bidirectional passage of substrates across biological membranes at an astonishing 10^8 molecules per second. AQPs are the most diversified in the plant kingdom, comprising of five major subfamilies that differ in temporal and spatial gene expression, subcellular protein localisation, substrate specificity, and post-translational regulatory mechanisms; collectively providing a dynamic transportation network spanning the entire plant. Plant AQPs can transport a range of solutes essential for numerous plant processes including, water relations, growth and development, stress responses, root nutrient uptake, and photosynthesis. The ability to manipulate AQPs towards improving plant productivity, is reliant on expanding our insight into the diversity and functional roles of AQPs.

Results

We have characterised the AQP family from *Nicotiana tabacum* (NtAQPs; tobacco), a popular model system capable of scaling from the laboratory to the field. Tobacco is closely related to major economic crops (e.g. tomato, potato, eggplant and peppers) and itself has new commercial applications. Tobacco harbours 76 AQPs making it the second largest characterised AQP family. These fall into five distinct subfamilies, for which we characterised phylogenetic relationships, gene structures, protein sequences, selectivity filter compositions, sub-cellular localisation, and tissue-specific expression. We also identified the AQPs from Tobacco's parental genomes (*Nicotiana sylvestris* and *Nicotiana tomentosiformis*), allowing us to characterise the evolutionary history of the NtAQP family. Assigning orthology to tomato and potato AQPs allowed for cross-species comparisons of protein structures and gene expression profiles.

Conclusions

This study provides a comprehensive overview of the tobacco AQP family, and strengthens the value of existing AQP resources by providing additional comparative information. The gene expression analysis together with the cross-species comparisons, provide insight into conservation and diversification of AQP function and likely physiological roles of subfamilies and individual isoforms within the Solanaceae. Collectively, these results will support and help direct future functional studies.

Key words: Aquaporins, *Nicotiana tabacum*, phylogenetics, gene evolution, gene expression, *Nicotiana sylvestris*, *Nicotiana tomentosiformis*, Solanaceae

2.2 Background

Cellular membranes are dynamic structures, continuously adjusting their composition in order to allow plants to respond to developmental signals, stresses, and changing environments (Marschner 2011). The biological function of cell membranes is conferred by its protein composition, with the lipid bilayer providing a basic structure and permeability barrier, and integral transmembrane proteins

facilitating diffusion of selected substrates (Marschner 2011). Cell membrane diffusion is a fundamental process of plant biology and one of the oldest subjects studied in plant physiology (Hedrich and Marten 2006). Diffusional events at the cellular level eventuate in the coordinated transport of substrates throughout the plant to support development and growth.

Plant membranes contain three major classes of transport proteins known as ATP-powered pumps, Transporters, and Channel proteins (Chrispeels et al. 1999). Pumps, are active transporters that use the energy of ATP hydrolysis to move substrates across the membrane against a concentration gradient or electrical potential. Transporters move a variety of molecules across a membrane along or against a gradient at rates of 10^2 to 10^4 molecules per second. Unlike the first two classes, channel proteins are bidirectional and increase membrane permeability to a particular molecule. Channel proteins are permeable to a wide range of substrates and up to 10^8 molecules per second can pass through them (Chrispeels et al. 1999) In plants, aquaporins (AQPs) constitute a major family of such channel proteins that facilitate selective transport of substrates for numerous biological processes including, water relations, plant development, stress responses, and photosynthesis (Hachez et al. 2006; Groszmann et al. 2017).

The AQP monomer forms a characteristic hour-glass membrane-spanning pore that assembles as tetrameric complexes in cell membranes. The union of the four monomers, creates a fifth pore at the centre of the tetramer which may provide an additional diffusional path (Frick et al. 2013). The substrate specificity of a given AQP is conferred by the complement of pore lining residues which achieve specificity through a combination of size exclusion and biochemical interactions with substrates (Hove and Bhave 2011). Key identified specificity residues include the dual Asn-Pro-Ala (NPA) motifs, the aromatic/Arginine filter (ar/R filter) and Froger's positions (P1-P5) (Froger et al. 1998; Mitani-Ueno et al. 2011; Murata et al. 2000). However, other pore-lining residues and lengths of the various transmembrane and loop domains of the AQP monomer are also known to influence substrate specificity through conformational changes of the pore size and accessibility (Wu and Beitz 2007; Hove and Bhave 2011). It is likely that other residues that determine specificity and transport efficiency remain to be elucidated.

Aquaporins, which are members of the major intrinsic proteins (MIP) superfamily, are found across all taxonomic kingdoms (Abascal et al. 2014). In land plants, AQPs are by far the most extensively diversified, being capable of transporting a wide variety of substrates including water, ammonia, urea, carbon dioxide, hydrogen peroxide, boron, silicon and other metalloids (Gomes et al. 2009; Pommerrenig et al. 2015; Hove and Bhave 2011). More recently, lactic acid, oxygen, and cations have been identified as permeating substrates (Choi and Roberts 2007; Zwiazek et al. 2017; Byrt et al. 2017; Bienert et al. 2013), with RNA molecules also implicated as a possible transported substrate (Reichel et al. 2016). Plant AQPs are divided into five phylogenetically distinct sub-families and further into sub-groups; Plasma membrane Intrinsic Proteins (PIPs), Tonoplast Intrinsic Proteins (TIPs), Small basic Intrinsic Proteins (SIPs), Nodulin 26-like Intrinsic Proteins (NIPs), and X Intrinsic Proteins (XIPs) (Danielson and Johanson 2008; Johanson and Gustavsson 2002; Kaldenhoff and Fischer 2006). The sub-families differ to some degree in substrate specificity and integrate into different cellular membranes, providing plants with a versatile system for both sub-cellular compartmentalisation and intercellular transport. Further versatility is achieved through tightly regulated spatial and temporal tissue-specific expression of different *AQP* genes, as well as post-translational modification of AQP proteins (e.g. phosphorylation) that controls membrane trafficking and channel activity (Santoni 2017; Luu and Maurel 2013).

Given their diverse complement of transported substrate and growing involvement in many developmental and stress responsive physiological roles, AQPs are targets for engineering more resilient and productive plants (Li et al. 2014; Groszmann et al. 2017). The genomic era of plant biology has provided unprecedented opportunity to query AQP biology by exploring sequence diversity between isoforms in many species. Analysis between closely related species is especially advantageous as it allows close orthologous AQPs to be compared and contrasted for protein structural features, tissue-specific expression patterns and sub-cellular localisation, all of which help towards elucidating physiological roles and regulation towards potential use in engineering efforts. In order to conduct such comparisons, we have characterised the AQP family from *Nicotiana tabacum*

(NtAQPs; tobacco). Tobacco is a fitting candidate species to explore unknowns of AQP biology as it is a popular model system for studying fundamental physiological processes that is capable of scaling from the laboratory to the field. Tobacco is part of the large Solanaceae family, which includes species of major economic importance such as tomato, potato, eggplant and peppers (Gebhardt 2016), and itself has renewed commercial applications in the biofuel and plant-made pharmaceutical sectors (Vanhercke et al. 2014; Tusé et al. 2014; Ma et al. 2015). We found that Tobacco harbours 76 AQPs, making it the second largest family characterised to date. Tobacco is a recent allotetraploid, which accounts for its large AQP family size. Phylogenetic relationships, gene structures, protein sequences, selectivity filter compositions, sub-cellular localisation, and tissue-specific expression profiles were analysed. We also identified the AQPs of the Tobacco parental genomes (*Nicotiana sylvestris* and *Nicotiana tomentosiformis*), allowing us to characterise the evolutionary history of the NtAQP family. Furthermore, using the already defined AQP families of tomato (*Solanum lycopersicum*) and potato (*Solanum tuberosum*) (Reuscher et al. 2013; Venkatesh et al. 2013), we made cross-species comparisons of gene structures, protein sequences and expression profiles, which contributes insight into conservation and diversification of protein function and physiological role for future studies.

2.3 Methods

2.3.1 Identification of tobacco, N.sylvestris and N.tomentosiformis AQPs

The tobacco genome and the protein sequences for TN90 (Sierro et al. 2014) and K326-Nitab4.5v (Edwards et al. 2017) cultivars were obtained from the Solanaceae Genomics Network (Fernandez-Pozo et al. 2014) and imported into the Geneious (V9.1.5) software (Drummond et al. 2011). To comprehensively identify putative aquaporin genes in Tobacco, multiple BLASTP searches were performed against the TN90 tobacco predicted proteome, using each of the potato (*Solanum tuberosum*) and tomato (*Solanum lycopersicum*) aquaporin proteins sequences as queries. From each individual homology search, the top 3-5 matches were compiled as putative NtAQPs; with the list being consolidated at the end of the search routine.

A similar process was used to identify AQPs in *N. sylvestris* and *N. tomentosiformis* (tobacco parental genomes), however tobacco aquaporin coding sequences were used in BLASTN queries. Sequence alignments were conducted using MUSCLE (Edgar 2004). Whole family and sub-family sequence alignments were used to flag aberrant AQP protein sequences for closer inspection.

2.3.2 Phylogenetic analysis and classification of tobacco, N. sylvestris and N. tomentosiformis AQPs

MUSCLE aligned nucleotide or protein sequences were used to construct phylogenetic trees using neighbour-joining (NJ) method (pair-wise deletion; bootstrap = 1000) in MEGA7 software (Kumar et al. 2016). Tobacco AQP naming convention was based on homology to that of the tomato AQPs. *N. sylvestris* and *N.tomentosiformis* AQP gene names were assigned based on homology to tobacco AQPs.

2.3.3 Structural features of tobacco AQPs

The tobacco aquaporin intron/exon structures were identified by aligning CDS and genomic sequences. Comparisons of gene sequences (computed and our curations) and RNA-seq data were visualised through JBrowse. The topologies of the curated NtAQPs were defined using TOPCONS (Tsirigos et al. 2015). The complement of known functionally relevant residues were collected from MUSCLE aligned NtAQP protein sequences. Alignment statistics (e.g. % sequence identity and similarity using BLSM62 matrix) were collected from MUSCLE aligned sequences of individual sub-families. Prediction of phosphorylation sites were performed using NetPhos 3.1 prediction score ≥ 0.8 (Blom et al. 1999).
Subcellular localisation predictions were achieved using; YLoc (Briesemeister et al. 2010), Wolf PSort (Horton et al. 2007) and Plant-mPloc (Chou and Shen 2010)

2.3.4 Subcellular localisation in planta (Arabidopsis)

Tobacco AQP GFP fusion constructs were generated via Gateway cloning of commercially synthesised TIP (*NtTIP1,1s*), PIP (*NtPIP2;5t*) and NIP (*NtNIP5;1t*)

coding sequences in pZeo entry vectors into the pMDC43 destination vector (Curtis and Grossniklaus 2003); which produced N-terminal GFP:NtAQP fusion proteins driven by the constitutive 2x35S CaMV promoter. Arabidopsis transgenic lines were generated via agrobacterium (GV3101) floral dipping plant transformation method (Clough and Bent 1998). The GFP marker line (MG0100.15) used as a cytosolic localisation marker was generated in our lab via the Gateway cloning of the mGFP6 variant of GFP contained as a pZeo entry clone into the pMDC32 destination vector (Curtis and Grossniklaus 2003); which drives constitutive expression of the mGFP6 transgene via the 2x35S CaMV promoter. The PM:GFP line was also generated in our lab, built in the pMDC83 Gateway destination vector and consisting of the Arabidopsis PIP2;1 (an already established PM marker (Nelson et al. 2007)) with a mGFP6 C-terminal fusion, all driven by the 2x35S CaMV promoter.

Arabidopsis seeds were liquid sterilised using hypochlorite, washed several times and sown on Gamorg's B5 medium containing 0.8% Agar and the antibiotic hygromycin for selection of transformants. After 8 days of growth, arabidopsis seedlings were gently removed from the agar, mounted in Phosphate Buffer (100mM $NaPO_4$ buffer, *pH* 7.2) on a standard slide and covered with coverslip, and visualised with a Zeiss LSM 780 Confocal microscope using a 40x water immersion objective (1.2 NA). Light micrographs of cortical cells in the root elongation zone were visualised using Differential Interference Contrast (DIC), with GFP fluorescence captured using excitation at 488 nm and emission detection across the 490-526 nm range. Autofluorescence was detected in the 570-674 nm range and excluded from GFP detection channel. Images were processed using Fiji (ImageJ) program (Schindelin et al. 2012).

2.3.5 AQP gene expression analysis

Transcript expression of the identified aquaporins was extracted from published, publicly available datasets, via two avenues; (1) mining of processed transcript expression matrices and (2) analyses of raw RNA-Seq reads uploaded to GenBank Sequence Read Archive (SRA). Processed transcript expression of *N. tabacum* K326 (Edwards et al. 2017) was extracted from The Sol Genomics Network (Fernandez-Pozo et al. 2014). Data was extracted as transcripts per million (TPM)

and so was mined without further processing. This data set contained tissue specific expression of the leaf and root. Raw RNA-Seq reads from both *N. tabacum* K326 and TN90 (Sierro et al. 2014) were downloaded from the GenBank SRA via command line into paired end fastq files. Reads were tissue specific from either the leaf, root, young leaf, young flower, mature leaf, mature flower, senescent leaf, senescent flower or dry capsule. Raw reads were first processed using Trimmomatic (Bolger et al. 2014) to remove any adapter sequences. Reads were then aligned to the *N. tabacum* transcriptome, either the K326 (Edwards et al. 2017) or TN90 (Sierro et al. 2014), depending on the variety of the sample using the Quasi align mode within Salmon (Patro et al. 2017) using a k-mer length of 31. This gave a relative abundance measure for each transcript (TPM). Mapping rates to the K326 transcriptome were between 73-78%, while mapping rates to the TN90 transcriptome were between 89-94%.

Tomato and potato root, leaf and flower expression data was retrieved through the EMBL-EBI Expression Atlas, and originally published by (Chattopadhyay et al. 2012) and (Consortium 2011).

2.4 Results

2.4.1 Identification and classification of NtAQP genes

A homology search, using tomato and potato AQP protein sequences as queries, identified 85 loci putatively encoding AQP-like genes in the genome of the TN90 tobacco cultivar (Sierro et al. 2014). Nine of these genes encode for severely truncated proteins and were classified as pseudogenes (Table S2. 1). The remaining 76 genes had a level of homology to tomato and potato AQPs to be considered 'bona fide' tobacco AQPs (NtAQPs; Table 2. 1). 73 of these 76 tobacco AQP genes were also identified in the genome of the more recently sequenced K326 cultivar (Nitab4.5v) (Edwards et al. 2017) (Table 2. 1). To determine the precise protein sequences and gene structures of the Tobacco AQPs, the surrounding genomic region of the identified coding sequences were examined in all forward translated frames. The likely protein products and associated intron/exon structures were curated through alignments with respective Solanaceae homologues. Our gene models were then

independently validated and supported by alignments against Tobacco whole transcriptome mRNA-seq data (obtained from Edwards et al., 2017), which also aided in defining the 5' and 3' UTRs. A comparison between our manually curated AQP protein and gene models against the computational predictions for the TN90 and K326 cultivars (Sierro et al. 2014; Edwards et al. 2017) revealed that 15% of TN90 and 50% of K326 computed AQP models were incorrectly annotated (Table 2. 1). Errors in the computed gene models were encountered across all NtAQP sub-families and consisted of either missing or truncated 5' and 3'UTRs, absent exons, truncated exons (ranging from 4-87 amino acids), and exon insertions (16-57 amino acids) due to inclusion of adjacent intron sequence (Figure 2. 1, Figure S2.1). A summary of our NtAQP gene models, identifiers and genomic locations for the TN90 and K326 cultivars are available in Table S2.2. FASTA sequencing files of coding DNA sequence (CDS), protein, and genomic sequence can be found in Additional file 3. Sequences of these high confidence NtAQP protein and gene models have been submitted to NCBI (Table 2. 1).

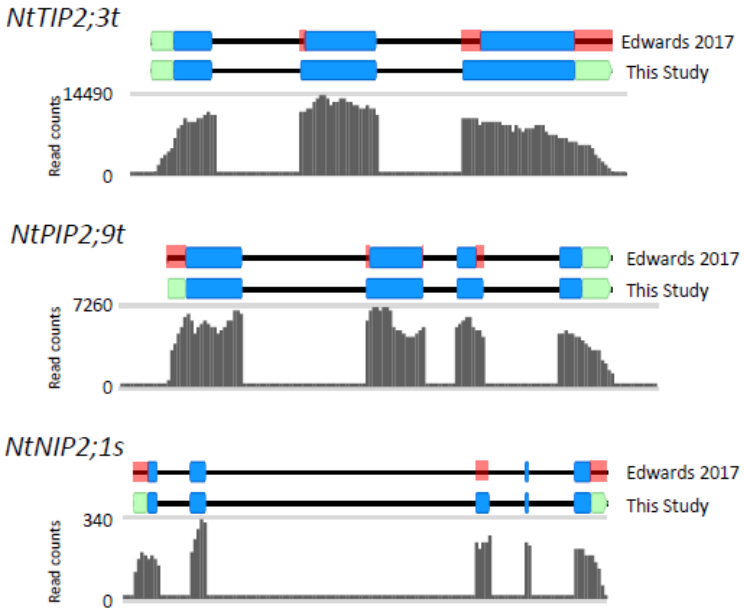


Figure 2. 1 Representative examples of our curated gene models validated with RNA-seq data. *Our curated models were aligned to those computed in Edwards et al. (2017). The examples depicted in the figure have high (NtTIP2;3t), medium (NtPIP2;9t) and low expression levels (NtNIP2;1s). Mapped genomic reads locate to mRNA encoding regions and as such denote exon boundaries and UTRs. Red*

boxes in the Edwards predicted gene models denote missing coding regions as indicated by deviations from the RNA-seq localisation.

Through the process of curating the tobacco AQP gene and protein sequences, we have made corrections to several previously mis-annotated AQP genes of tomato and potato namely, *StXIP3;1, StXIP4;1*, *SlXIP1;6*, *SlPIP2;1*, and *SlTIP2;2* (Table S2.3). We also identified through our Tobacco genome sequence analysis an erroneous non-synonymous single nucleotide mutation (C > T, CDS position 619) in the reported mRNA sequence of the frequently studied Tobacco AQP1 gene (NtAQP1; assigned as NtPIP1;5s in this study). The mutation results in a Histidine (H) to Tyrosine (Y) substitution at amino acid position 207 being incorrectly reported in the initial cloning of this gene and subsequent use ((Biela et al. 1999); NCBI AF024511 and AJ001416). This substitution is notable since His207,which corresponds to the His193 position of the well-studied crystal structures of Spinach PIP2;1 (Törnroth-Horsefield et al. 2006; Nyblom et al. 2009; Frick et al. 2013), is highly conserved across all angiosperm PIP AQPs and is a key regulator in the gating, and therefore transport capacity, of the AQP channel (Frick et al. 2013; Törnroth-Horsefield et al. 2006; Tournaire-Roux et al. 2003). The inadvertent use of this H207Y NtAQP1 mutant in functional characterisation studies may have implication on the conclusions drawn for this frequently studied plant AQP. In support of His207 being the correct residue in NtAQP1, we found that several closely related NtAQP1 orthologues across several Solanaceae species, including 3 additional *Nicotiana* species, all had the His207 residue (Figure S2.2).

2.4.2 Gene structures and phylogenetic analysis of Tobacco AQPs

A phylogeny of the 76 curated NtAQP protein sequences with the tomato, potato and Arabidopsis AQP genes (Figure S2.3), segregated the NtAQPs into five distinct sub-families, namely the NIPs (16), SIPs (5), XIPs (4), PIPs (29) and TIPs (22) (Figure 2.2). An emerging problem among the increasing number of studies characterising plant AQP families across species is the confusion in nomenclature that either misses or incorrectly assigns orthology between AQP genes. Such confusion is seen in the nomenclature between tomato and potato AQPs. At least in

this case, the naming inconsistency is predominantly a result of the two family characterisations being published concurrently by different groups (Reuscher et al. 2013; Venkatesh et al. 2013). Towards contributing to a more congruent naming structure of AQPs between species, especially within a single family of angiosperms, we aligned our NtAQP naming convention with that of tomato AQPs, given their more consistent nomenclature to likely Arabidopsis AQP orthologues. Table S2.2 lists the tobacco AQPs with their corresponding tomato and potato orthologous genes.

65 of the 76 NtAQP genes had clear orthologs in tomato which directed their naming (Figure S2.3 and Table S2.2). The 11 tobacco *AQPs* with no apparent tomato or potato ortholog were allocated designations unique to tobacco (denoted by black stars in Figure S2.3). Gene lengths varied between NtAQPs from 1,091bp to 6,627bp, with a single extreme instance of 17,278bp (*NtPIP2;11s*) due to a large intron insertion (Figure 2.2). The exon-intron patterning of NtAQP genes were highly conserved with that of their tomato and potato orthologs (Table S2.2) (Reuscher et al. 2013; Venkatesh et al. 2013). Individual AQPs within the PIP, TIP, NIP and SIP sub-families were well conserved across the three Solanaceae species (Figure S2.3). The XIPs were an exception as they predominantly phylogenetically clustered within each separate species, pointing to a high degree of intra-species XIP diversification within the Solanaceae (Figure S2.3).

A distinctive feature in the phylogeny was that most NtAQPs reside as pairs, supported by high bootstrap values (Figure 2.2). The high homology in protein sequences between members of these phylogenetic pairs also extended to highly similar nucleotide sequences and gene structures (Figure 2. 2).

Table 2. 1. List of the 76 tobacco aquaporin genes identified in this study. *Provided are protein lengths, gene identifiers in the TN90[(1)] (Sierro et al. 2014) and K326[(2)] (Edwards et al. 2017) cultivar genomes, comparison of whether the computational gene models derived from each study matched the curated gene structures (Y-yes or N-no) and NCBI accession identifiers. NtTIP2;5s, NtNIP4;2s and NtNIP4;3t genes were not identified in the K326[(2)] cultivar's genome. For further details, including gene identifiers, previous NCBI accession identifiers etc.., see the expanded version of this table in Table S2.2.*

This Study			TN90 Sierro *et al.* 2014			K326 Edwads *et al.* 2017	
Name	**Protein (aa)**	**NCBI accession-This study**	**Gene name [(1)]**	**mRNA ID [(1)]**	**Accurate gene model?[(1)]**	**Gene name [(2)]**	**Accurate gene model? [(2)]**
NtPIP1;1s	289	BK011392	gene_35182	mRNA_75678_cds	Y	Nitab4.5_0004836g0030.1	**N**
NtPIP1;1t	289	BK011393	gene_27714	mRNA_59319_cds	Y	Nitab4.5_0006090g0020.1	**N**
NtPIP1;2s	288	BK011394	gene_58674	mRNA_125284_cds	Y	Nitab4.5_0011459g0010.1	Y
NtPIP1;2t	286	BK011395	gene_10991	mRNA_23602_cds	**N**	Nitab4.5_0000583g0150.1	Y
NtPIP1;3s	288	BK011396	gene_79275	mRNA_170144_cds	Y	Nitab4.5_0007597g0010.1	Y
NtPIP1;3t	288	BK011397	gene_84661	mRNA_181592_cds	Y	Nitab4.5_0003043g0010.1	Y
NtPIP1;5s	288	BK011398	gene_40739	mRNA_87599_cds	Y	Nitab4.5_0010813g0020.1	**N**
NtPIP1;5t	288	BK011399	gene_80239	mRNA_172222_cds	Y	Nitab4.5_0001615g0140.1	y
NtPIP1;7s	287	BK011400	gene_59749	mRNA_127708_cds	Y	Nitab4.5_0006718g0030.1	**N**
NtPIP1;8s	286	BK011401	gene_86041	mRNA_184690_cds	Y	Nitab4.5_0000737g0120.1	Y
NtPIP2;1s	284	BK011402	gene_9798	mRNA_21207_cds	Y	Nitab4.5_0009795g0020.1	Y
NtPIP2;1x	284	BK011403	gene_9795	mRNA_21200_cds	**N**	Nitab4.5_0009795g0010.1	**N**
NtPIP2;2t	284	BK011404	gene_87071	mRNA_186851_cds	Y	Nitab4.5_0000101g0110.1	Y
NtPIP2;3t	284	BK011405	gene_8898	mRNA_19079_cds	**N**	Nitab4.5_0000101g0120.1	**N**
NtPIP2;4s	288	BK011406	gene_84258	mRNA_180721_cds	Y	Nitab4.5_0004314g0010.1	Y
NtPIP2;4t	288	BK011407	gene_71307	mRNA_152443_cds	Y	Nitab4.5_0000181g0120.1	**N**
NtPIP2;5s	286	BK011408	gene_31592	mRNA_67547_cds	Y	Nitab4.5_0001192g0080.1	**N**
NtPIP2;5t	286	BK011409	gene_32945	mRNA_70539_cds	Y	Nitab4.5_0001297g0050.1	**N**
NtPIP2;6s	288	BK011410	gene_22735	mRNA_48588_cds	Y	Nitab4.5_0004108g0020.1	Y
NtPIP2;6t	288	BK011411	gene_34319	mRNA_73634_cds	Y	Nitab4.5_0000650g0260.1	Y
NtPIP2;7t	284	BK011412	gene_84225	mRNA_180649_cds	Y	Nitab4.5_0000106g0170.1	Y
NtPIP2;8s	285	BK011413	gene_75147	mRNA_160866_cds	Y	Nitab4.5_0003914g0040.1	Y
NtPIP2;8t	285	BK011414	gene_53392	mRNA_114047_cds	Y	Nitab4.5_0000283g0420.1	Y

Table 2.1 continued.

NtPIP2;9s	284	BK011415	gene_84936	mRNA_182184_cds	Y	Nitab4.5_0005236g0020.1	**N**
NtPIP2;9t	284	BK011416	gene_9787	mRNA_21186_cds	**N**	Nitab4.5_0002763g0030.1	**N**
NtPIP2;11s	269	BK011417	gene_40272	mRNA_86606_cds	Y	Nitab4.5_0008552g0040.1	**N**
NtPIP2;11t	269	BK011418	gene_62966	mRNA_134569_cds	Y	Nitab4.5_0001789g0070.1	**N**
NtPIP2;13s	284	BK011419	gene_55607	mRNA_118532_cds	Y	Nitab4.5_0014443g0010.1	Y
NtPIP2;13t	284	BK011420	gene_81728	mRNA_175293_cds	Y	Nitab4.5_0000575g0130.1	Y
NtTIP1;1s	252	BK011426	gene_4702	mRNA_9987_cds	Y	Nitab4.5_0003155g0010.1	Y
NtTIP1;1t	252	BK011427	gene_17915	mRNA_38368_cds	Y	Nitab4.5_0001163g0070.1	**N**
NtTIP1;2s	253	BK011428	gene_62289	mRNA_133117_cds	Y	Nitab4.5_0001068g0010.1	Y
NtTIP1;2t	253	BK011429	gene_18091	mRNA_38668_cds	Y	Nitab4.5_0000766g0050.1	Y
NtTIP1;3s	249	BK011430	gene_81216	mRNA_174227_cds	Y	Nitab4.5_0011193g0010.1	Y
NtTIP1;3t	249	BK011431	gene_34364	mRNA_73720_cds	Y	Nitab4.5_0022765g0010.1	Y
NtTIP1;4t	252	BK011432	gene_44062	mRNA_94602_cds	Y	Nitab4.5_0000173g0030.1	**N**
NtTIP2;1s	249	BK011433	gene_13886	mRNA_29709_cds	**N**	Nitab4.5_0009267g0020.1	**N**
NtTIP2;1t	249	BK011434	gene_84779	mRNA_181854_cds	Y	Nitab4.5_0003039g0050.1	**N**
NtTIP2;2s	251	BK011435	gene_65205	mRNA_139414_cds	Y	Nitab4.5_0001381g0190.1	**N**
NtTIP2;3s	251	BK011436	gene_8782	mRNA_18868_cds	Y	Nitab4.5_0001076g0030.1	**N**
NtTIP2;3t	251	BK011437	gene_77281	mRNA_165620_cds	Y	Nitab4.5_0000618g0070.1	**N**
NtTIP2;4s	249	BK011438	gene_44575	mRNA_95620_cds	Y	Nitab4.5_0007573g0030.1	Y
NtTIP2;5s	249	BK011439	gene_55803	mRNA_118941_cds	Y	*Not identified*	-
NtTIP2;5t	249	BK011440	gene_36783	mRNA_79146_cds	Y	Nitab4.5_0011578g0040.1	N
NtTIP3;1s	260	BK011441	gene_7183	mRNA_15505_cds	Y	Nitab4.5_0005315g0010.1	Y
NtTIP3;1t	260	BK011442	gene_54243	mRNA_115698_cds	Y	Nitab4.5_0000477g0090.1	Y
NtTIP3;2t	259	BK011443	gene_79868	mRNA_171429_cds	**N**	Nitab4.5_0009307g0020.1	Y
NtTIP4;1s	248	BK011444	gene_76645	mRNA_164188_cds	Y	Nitab4.5_0000837g0080.1	**N**
NtTIP4;1t	248	BK011445	gene_2305	mRNA_4886_cds	Y	Nitab4.5_0000151g0360.1	Y
NtTIP5;1s	251	BK011446	gene_8008	mRNA_17273_cds	Y	Nitab4.5_0010023g0020.1	Y
NtTIP5;1t	251	BK011447	gene_33209	mRNA_71126_cds	Y	Nitab4.5_0002816g0050.1	**N**
NtNIP1;1s	275	BK011376	gene_27146	mRNA_58000_cds	**N**	Nitab4.5_0005428g0060.1	**N**
NtNIP1;2s	288	BK011377	gene_42864	mRNA_92251_cds	Y	Nitab4.5_0008572g0060.1	**N**
NtNIP1;2t	282	BK011378	gene_42851	mRNA_92212_cds	**N**	Nitab4.5_0001778g0110.1	Y
NtNIP2;1s	287	BK011379	gene_24518	mRNA_52562_cds	Y	Nitab4.5_0001638g0020.1	**N**

Table 2.1 continued.

NtNIP3;1s	348	BK011380	gene_85282	mRNA_182892_cds	Y	Nitab4.5_0013395g0010.1	**N**
NtNIP4;1s	271	BK011381	gene_11802	mRNA_25329_cds	Y	Nitab4.5_0003360g0080.1	**N**
NtNIP4;1t	272	BK011382	gene_33173	mRNA_71046_cds	Y	Nitab4.5_0004399g0020.1	**N**
NtNIP4;2s	273	BK011383	gene_47152	mRNA_100755_cds	Y	Not identified	-
NtNIP4;2t	273	BK011384	gene_36231	mRNA_78016_cds	Y	Nitab4.5_0000742g0130.1	**N**
NtNIP4;3s	282	BK011385	gene_55126	mRNA_117533_cds	Y	Not identified	-
NtNIP5;1s	298	BK011386	gene_36225	mRNA_77996_cds	Y	Nitab4.5_0005519g0010.1	**N**
NtNIP5;1t	298	BK011387	gene_38118	mRNA_82133_cds	Y	Nitab4.5_0000799g0080.1	Y
NtNIP6;1s	304	BK011388	gene_39457	mRNA_84834_cds	**N**	Nitab4.5_0012943g0030.1	Y
NtNIP6;1t	304	BK011389	gene_8958	mRNA_19210_cds	**N**	Nitab4.5_0001454g0120.1	**N**
NtNIP7;1s	294	BK011390	gene_69139	mRNA_147657_cds	Y	Nitab4.5_0007039g0010.1	**N**
NtNIP7;1t	281	BK011391	gene_41519	mRNA_89450_cds	Y	Nitab4.5_0002600g0020.1	**N**
NtSIP1;1t	238	BK011421	gene_54009	mRNA_115212_cds	**N**	Nitab4.5_0000001g0350.1	Y
NtSIP1;2s	244	BK011422	gene_73217	mRNA_156641_cds	Y	Nitab4.5_0007223g0030.1	Y
NtSIP1;2t	243	BK011423	gene_74850	mRNA_160182_cds	Y	Nitab4.5_0000812g0160.1	**N**
NtSIP2;1s	241	BK011424	gene_42066	mRNA_90601_cds	Y	Nitab4.5_0001918g0070.1	**N**
NtSIP2;1t	241	BK011425	gene_29131	mRNA_62430_cds	**N**	Nitab4.5_0000721g0170.1	Y
NtXIP1;6s	327	BK011448	gene_13292	mRNA_28463_cds	Y	Nitab4.5_0007293g0050.1	**N**
NtXIP1;6t	327	BK011449	gene_52652	mRNA_112498_cds	Y	Nitab4.5_0000956g0150.1	**N**
NtXIP1;7s	314	BK011450	gene_34706	mRNA_74489_cds	Y	Nitab4.5_0007733g0020.1	**N**
NtXIP1;7t	314	BK011451	gene_50247	mRNA_107135_cds	Y	Nitab4.5_0006828g0010.1	**N**

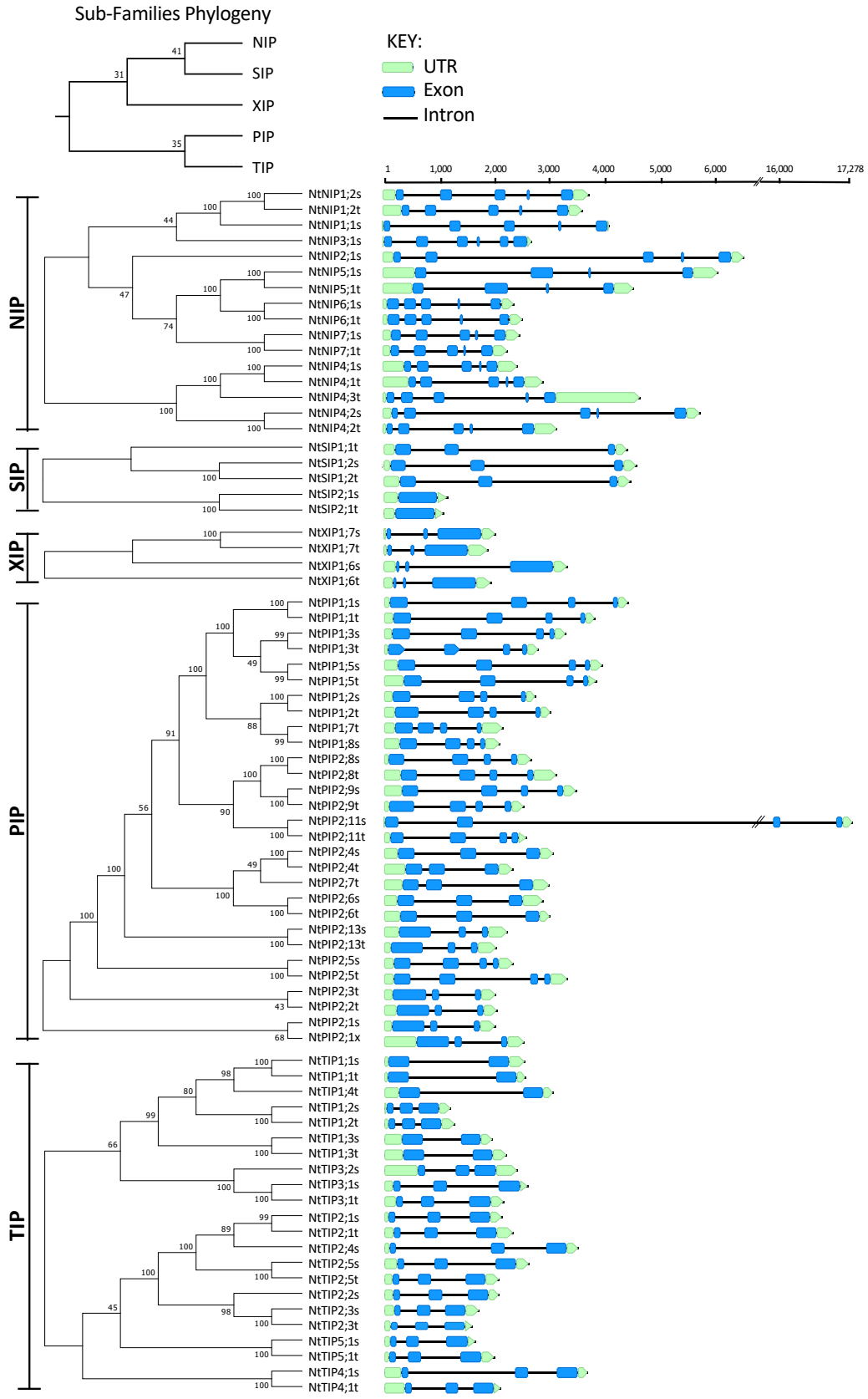


Figure 2. 2 Phylogeny and gene structures of 76 tobacco aquaporins. *Phylogenetic tree was generated using the neighbour-joining method (via MEGA7) from MUSCLE aligned protein sequences.*

Confidence levels (%) of branch points generated through bootstrapping analysis (n=1000). Gene structures are located adjacent to their respective location on the phylogenetic tree; blue rectangles correspond to the exons; green rectangles and arrows to the 5' and 3'UTRs, respectively. Scale bar at the top of gene structures indicates nucleotid length. The last letter in the NtAQP names denote the likely origin of the gene (s = N. sylvestris, t = N. tomentosiformis, x = unknown).

2.4.3 Tobacco AQP protein sequence comparisons

2.4.3.1 General structural features of NtAQP proteins

Topological analysis using TOPCON (see materials and methods), predicted that all NtAQP proteins consist of six transmembrane helical domains, five intervening loop regions and cytoplasmic localised N- and C- terminal tails, which is consistent with the typical structure of AQPs (Figure 2. 3). The size of the transmembrane helical domains appear to be an integral property of the AQP structure given their remarkably conserved lengths across the sub-families (Figure 2. 3A). Conversely, the length of the loop regions showed substantial variability between sub-families (Figure 2.3A). The most pronounced was Loop A, which is prominently longer and apoplastically exposed in the PIP2s (18aa) and shorter in the NIPs (8aa) compared to the average length of TIPs, SIPs, and XIPs (14aa). The cytoplasmic Loop B, is shorter in XIPs (20aa vs. 24aa). Loop C is nearly double the length in the XIPs (38aa) compared to the other sub-families (20aa). Loop D is slightly longer in the PIPs (12aa) and shorter in the SIPs (7aa), while Loop E is substantially longer in the XIPs (32aa) and shorter in the NIPs (20aa) (Figure 2. 3A). The cytoplasmically localised N- and C-terminal tails are the most varied in size of any of the AQP domains (Figure 2. 3A). The N-terminal tail ranges from 59aa in the NIPs to just 7aa in the SIPs and the C-terminal tail from 30aa in the NIPs to 14aa in the PIPs.

Examining sequence conservation of the different protein domains across the sub-families, revealed that the transmembrane helices are generally the more highly conserved feature of the AQP (Figure 2. 3B). Loop B and E are also highly conserved relative to the other domains, which is likely owing to their direct role in forming the transmembrane pore. Conversely, Loops A and C, along with the two

terminal tails were found to be the least conserved domains within each NtAQP sub-family (Figure 2. 3B).

To learn more about the putative functional characteristics of the different NtAQPs, we used multiple protein sequence alignments to report residue compositions at key positions in the protein known to regulate AQP function (Table 2. 2). Included are the dual Asn-Pro-Ala (NPA) motifs, the five Froger's position residues (P1-P5), and the residues of the aromatic/Arginine filter (ar/R filter), all of which are specific pore lining residues that contribute to determining which substrates permeate though the AQP pore. We also reported on several other sites known to be post-translationally modified, which influence channel activity and membrane localisation (Table 2. 2).

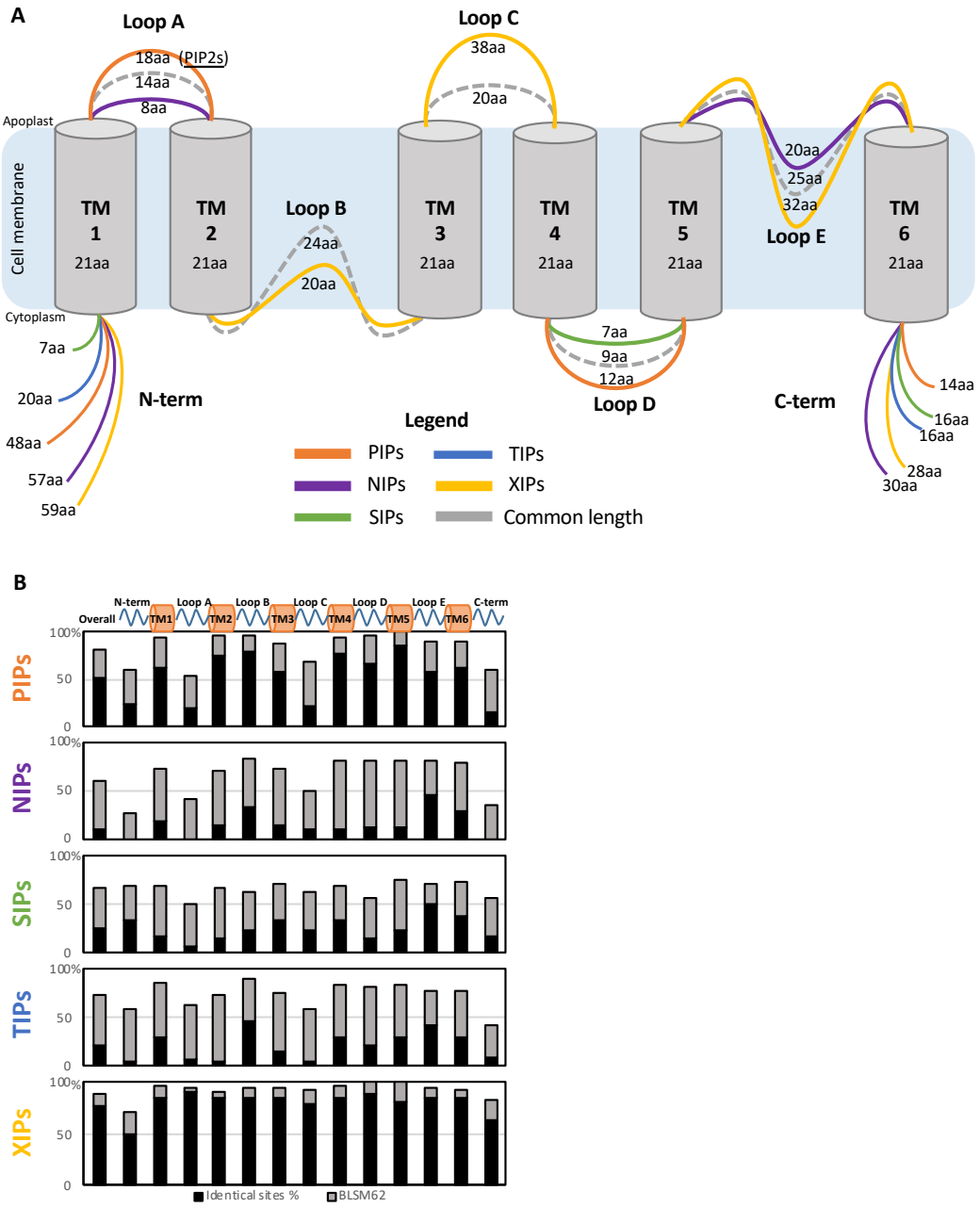


Figure 2. 3 Protein sequence comparisons of NtAQP sub-families. (A) *Diagrammatic illustration of an AQP depicting protein topology and lengths of the; N-terminal tail (N-term), TransMembrane domains (TM) 1-6, Loops A-E, and C-terminal tail (C-term). The average amino acid (aa) lengths of each structural feature are listed for the different NtAQP sub-families. Common length of a domain is represented in grey, while deviations from the common length are in colour; PIPs (orange), NIPs (purple), SIPs (green), TIPs (blue) and XIPs (yellow).* ***(B)*** *Overall and intra-domain sequence similarities for each NtAQP sub-family. Schematic representation of the AQP domains is illustrated at the top, with aligned columns showing protein sequence identical sites (black) and the BLSM62 similarity score (grey) between members of the given NtAQP subfamily.*

2.4.3.2 NtPIP subfamily

The NtPIPs represent the largest NtAQP subfamily with 29 members that are phylogenetically divided into PIP1 and PIP2 subgroups. Despite being the largest subfamily, the NtPIPs were among the most conserved in protein sequence (>50%; Figure 2. 3B). The apoplastic exposed Loops A and Loop C were the exceptions having only ~20% sequence identity and varying in size between PIP1 and PIP2 proteins (Figure 2. 3). This sequence diversification could be of functional importance given Loop A is involved in PIP-PIP dimerization mediated primarily through a conserved cysteine residue, which is present in all NtPIPs (Roche and Törnroth-Horsefield 2017; Bienert et al. 2012). The generally high sequence similarity across most of the PIP protein domains was also reflected in both PIP1s and PIP2s having identical configuration of residues across the NPA and ar/R motifs; which were predominantly hydrophilic residues (Table 2. 2). Only Froger's position 2 showed variation with amino acids of different properties (G, M or Q) occupying this position (Table 2. 2). The NtPIP1s are predominantly distinguished from NtPIP2s by having longer N-terminal and shorter C-terminal tail sequences. The N-terminal tail is involved in calcium-dependent gating of the pore which occurs through interactions involving two acidic residues (Asp28 and Glu31, Table 2. 2) (Törnroth-Horsefield et al. 2006). Pore gating is also triggered by pH involving protonation of a Loop D histidine (His-193, Table 2. 2) and phosphorylation of a Loop B serine (Ser115, Table 2. 2) (Tournaire-Roux et al. 2003; Törnroth-Horsefield et al. 2006). These four residues were identified in each NtPIPs indicating the entire subfamily retains these modes of regulation (Table 2. 2). The Loop B serine (Ser115), or phosphorylatable threonine, was also conserved in members of XIPs, TIPs and SIPs (but not NIPs), suggesting a shared mechanism of gating regulation between different NtAQPs (Table 2. 2). Two commonly phosphorylated serine sites were found conserved in the longer C-terminal tail of NtPIP2s (Ser274 and Ser277; Table 2. 2, Figure S2. 4). The phosphorylation status of these serine residues are known to facilitate protein-protein interactions, influence trafficking to and from the plasma membrane, and alter the transport capacity of the pore (Chevalier and Chaumont 2014; Groszmann et al. 2017). NtPIP1 proteins have the second of these serine residues (Ser277), but are not predicted to be phosphorylated (Table 2. 2;

Figure S2. 4). A strongly conserved positively charged lysine or arginine directly preceding the second phosphorylated serine is found across all NtPIPs, and also more broadly across PIPs from other plant species (data not shown), with the exception of NtPIP1;5 and PIP2;11 which have a histidine (Figure S2. 4). Histidine can achieve a positive charge through protonation, indicating a possible pH regulated functional state of the C-terminal tail in these NtPIPs.

2.4.3.3 NtNIP subfamily

NIPs were found to have the lowest overall sequence identical sites (~10%), suggesting a highly divergent subfamily at the sequence level (Figure 2. 3B). The sequence variation was evenly distributed across all AQP domains, with only Loop B and Loop E retaining modest conservation with >30% identical residues per site. This comparatively higher conservation likely reflects these two loops being directly involved in forming the main pore structure and controlling substrate selectivity. Loops B and E each contain a NPA motif, and Loop E also contains ar/R and Froger's residues (Table 2. 2). Across the NtNIPs, there was substantial variation in the residues constituting the dual NPA motifs (NP<u>A</u>/<u>S</u>/<u>V</u>) and across all 5 Froger's positions (Table 2. 2). And all but LE2 of the ar/R residues were variable, although the residue that were present tended to be more hydrophobic (Table 2. 2). Also notable in the NtNIPs, were their distinctively longer N and C terminals (~57-30aa) compared to those in other sub-families (Figure 2. 3A). The extended C-terminal tail contains numerous serine residues, many of which were predicted to be phosphorylated (Figure S2. 4). Included were serine residues at homologous positions to the confirmed phosphorylated sites of Ser262 in GmNOD26 (a soybean NIP) and Ser277 in PIPs (Table 2. 2). The Ser115 phosphorylation site that controls aspects of pore gating in PIPs was conserved and predicted to be phosphorylated in only NtNIP4;3s, with all other NtNIPs having a structurally rigid proline residue at this position (Table 2. 2).

2.4.3.4 NtTIP subfamily

Conservation among the NtTIPs was ~22% sequence identity (Figure 2. 3B). Similar to the NIPs, the highest sequence conservation occurred in Loops B and E (>40%). The dual NPA motif, ar/R H2 and Froger's P3 to P5 are well conserved among the different TIP subgroups. The exceptions being NtTIP2;1s with a NPD configuration of the first NPA motif, and the NtTIP5;1 proteins which have a H>N substitution at ar/R H2 (Table 2. 2). The other ar/R and Froger's sites are rather variable among the NtTIPs, especially ar/R LE2 which varies between amino acids of quite differing properties (V, R or Y; Table 2. 2). A histidine opposed to phenylalanine located at ar/R LC of NtTIP2s, TIP4s and TIP5s (Table 2. 2), suggests an enhanced capacity to transport ammonia (Kirscht et al. 2016). The Ser115 phosphorylation site that controls pore gating in PIPs was identified in 5 of the 22 NtTIPs, with the remaining NtTIPs possessing a threonine which is also a potentially phosphorylatable residue. NtTIP2 and NtTIP5 proteins have a conserved histidine (His131) in Loop C that is involved in a similar pH regulated gating of the pore to that of His193 in Loop D of PIPs and NIPs (Leitão et al. 2012; Soto et al. 2010). The C-terminal tail of NtTIPs contained on average less than 2 serine residues, none of which were predicted to be phosphorylation targets (data not shown).

Table 2. 2 Amino acid composition of NtAQPs at known functionally important positions. *Listed are the two NPA motifs (NPA I in Loop B, NPA II in Loop E), ar/R residues (in Helix 2, 5, Loop C and Loop E) and the Froger's residues (Positions 1-5). Other known regulatory residues involved with post-translational regulation of AQPs are listed: Asp28, Glu31, Ser115, His131, His193, Ser274, Ser277. Amino acid positions of these sites are relative to Spinach PIP2;1 or Arabidposis TIP2;1 (the His131 residue), as their crystal structures have been defined and well-studied. Serine (S) residues with a superscript P (i.e. S^P) are predicted to be phosphorylated according to NetPhos3.1. Also included are the predicted subcellular localisations from Plant-mPloc, WolfPsort and YLoc; outputs include plasma membrane (plas), cytosol (cyto), tonoplast (tono), chloroplast (chlo), endoplasmic reticulum (ER), peroxisome (pero) and extracellular (extra) localisation. The last letter in the NtAQP names denote the likely origin of the gene (s = N. sylvestris, t = N. tomentosiformis, x = unknown).*

		NPA I	NPA II	ar/R					Froger's positions					Other regulatory residues							Pred. subcellular loc.		
	Gene Name	LB	LE	H2	LC	H5	LE1	LE2	P1	P2	P3	P4	P5	Asp28	Glu31	Ser115	Ser188	His193	Ser274	Ser277	mPLoc	WolfPsort	Yloc
PIP	NtPIP1;1s	NPA	NPA	F	G	H	T	R	M	S	A	F	W	E	E	S^P	N	H	-	S	plas	plas	cyto
	NtPIP1;1t	NPA	NPA	F	G	H	T	R	M	S	A	F	W	E	E	S^P	N	H	-	S	plas	plas	cyto
	NtPIP1;2s	NPA	NPA	F	G	H	T	R	Q	S	A	F	W	E	E	S^P	N	H	-	S	plas	plas	cyto
	NtPIP1;2t	NPA	NPA	F	G	H	T	R	Q	S	A	F	W	E	E	S^P	N	H	-	S	plas	plas	cyto
	NtPIP1;3s	NPA	NPA	F	G	H	T	R	M	S	A	F	W	E	E	S^P	N	H	-	S	plas	plas	cyto
	NtPIP1;3t	NPA	NPA	F	G	H	T	R	M	S	A	F	W	E	E	S^P	N	H	-	S	plas	plas	cyto
	NtPIP1;5s	NPA	NPA	F	G	H	T	R	M	S	A	F	W	E	E	S^P	N	H	-	S	plas	plas	cyto
	NtPIP1;5t	NPA	NPA	F	G	H	T	R	M	S	A	F	W	E	E	S^P	N	H	-	S	plas	plas	cyto
	NtPIP1;7t	NPA	NPA	F	G	H	T	R	G	S	A	F	W	E	E	S^P	S^P	H	-	S	plas	plas	cyto
	NtPIP1;8s	NPA	NPA	F	G	H	T	R	G	S	A	F	W	E	E	S^P	S^P	H	-	S	plas	plas	plas
	NtPIP2;1s	NPA	NPA	F	G	H	T	R	Q	S	A	F	W	D	E	S^P	N	H	S^P	S	plas	plas	plas
	NtPIP2;1x	NPA	NPA	F	G	H	T	R	Q	S	A	F	W	D	E	S^P	N	H	S^P	S	plas	plas	plas
	NtPIP2;2t	NPA	NPA	F	G	H	T	R	Q	S	A	F	W	D	E	S^P	N	H	S^P	S	plas	plas	plas
	NtPIP2;3t	NPA	NPA	F	G	H	T	R	Q	S	A	F	W	D	E	S^P	N	H	S^P	S	plas	plas	plas
	NtPIP2;4s	NPA	NPA	F	G	H	T	R	Q	S	A	F	W	D	E	S^P	N	H	S^P	S	plas	plas	plas
	NtPIP2;4t	NPA	NPA	F	G	H	T	R	Q	S	A	F	W	D	E	S^P	N	H	S^P	S	plas	plas	plas
	NtPIP2;5s	NPA	NPA	F	G	H	T	R	Q	S	A	F	W	D	E	S^P	N	H	S^P	S	plas	plas	plas
	NtPIP2;5t	NPA	NPA	F	G	H	T	R	Q	S	A	F	W	D	E	S^P	N	H	S^P	S	plas	plas	plas
	NtPIP2;6s	NPA	NPA	F	G	H	T	R	Q	S	A	F	W	D	E	S^P	S^P	H	S^P	S	plas	plas	plas
	NtPIP2;6t	NPA	NPA	F	G	H	T	R	Q	S	A	F	W	D	E	S^P	S^P	H	S^P	S	plas	plas	plas
	NtPIP2;7t	NPA	NPA	F	G	H	T	R	Q	S	A	F	W	D	E	S^P	S^P	H	S^P	S	plas	plas	plas
	NtPIP2;8s	NPA	NPA	F	G	H	T	R	M	S	A	F	W	D	E	S	S^P	H	SP	S	plas	plas	plas
	NtPIP2;8t	NPA	NPA	F	G	H	T	R	M	S	A	F	W	D	E	S	S^P	H	SP	S	plas	plas	plas
	NtPIP2;9s	NPA	NPA	F	G	H	T	R	M	S	A	F	W	D	E	S^P	S^P	H	S^P	S	plas	plas	plas
	NtPIP2;9t	NPA	NPA	F	G	H	T	R	M	S	A	F	W	D	E	S^P	S^P	H	S^P	S	plas	plas	plas
	NtPIP2;11s	NPA	NPA	F	G	H	T	R	M	S	A	F	W	D	E	S^P	N	H	S^P	S	plas	plas	plas
	NtPIP2;11t	NPA	NPA	F	G	H	T	R	M	S	A	F	W	D	E	S^P	N	H	S^P	S	plas	plas	plas
	NtPIP2;13s	NPA	NPA	F	G	H	T	R	Q	S	A	F	W	D	E	S^P	N	H	S^P	S^P	plas	plas	plas
	NtPIP2;13t	NPA	NPA	F	G	H	T	R	Q	S	A	F	W	D	E	S^P	N	H	S^P	S^P	plas	plas	plas
NIP	NtNIP1;1s	NPS	NPA	W	A	V	A	R	F	S	A	Y	L	-	-	P	-	-	S^P	-	plas	plas	cyto
	NtNIP1;2s	NPA	NPA	W	L	V	A	R	F	S	A	Y	M	-	-	P	-	-	S	-	plas	plas	plas
	NtNIP1;2t	NPA	NPA	W	L	V	A	R	F	S	A	Y	M	-	-	P	-	-	S	-	plas	plas	plas
	NtNIP2;1s	NPA	NPA	G	T	S	G	R	L	T	A	Y	I	-	-	P	-	-	S^P	-	plas	plas	plas
	NtNIP3;1s	NPA	NPA	W	V	I	A	R	F	S	A	Y	I	-	-	P	-	-	S	-	plas	plas	plas
	NtNIP4;1s	NPA	NPA	W	F	V	A	R	F	S	A	Y	I	-	-	P	-	-	S	-	plas	plas	plas
	NtNIP4;1t	NPA	NPA	W	F	V	A	R	F	S	A	Y	I	-	-	P	-	-	S	-	plas	plas	plas
	NtNIP4;2s	NPA	NPA	W	F	V	A	R	L	S	A	Y	I	-	-	P	-	-	T	-	plas	plas	plas
	NtNIP4;2t	NPA	NPA	W	F	V	A	R	L	S	A	Y	I	-	-	P	-	-	T	-	plas	plas	plas
	NtNIP4;3s	NPA	NPA	W	F	V	A	R	L	S	A	Y	I	-	-	S^P	-	-	S	-	plas	tono	plas
	NtNIP5;1s	NPA	NPA	S	V	I	A	R	F	T	A	Y	I	-	-	P	-	-	S	-	plas	tono	cyto
	NtNIP5;1t	NPS	NPV	A	V	I	A	R	F	T	A	Y	L	-	-	P	-	-	S	-	plas	tono	cyto
	NtNIP6;1s	NPA	NPV	S	V	I	A	R	L	T	A	Y	L	-	-	P	-	-	S	-	plas	plas	plas
	NtNIP6;1t	NPA	NPV	T	V	I	A	R	L	T	A	Y	L	-	-	P	-	-	S	-	plas	plas	plas
	NtNIP7;1s	NPA	NPA	A	I	V	G	R	Y	S	A	Y	V	-	-	P	-	-	T	-	plas	plas	plas
	NtNIP7;1t	NPA	NPA	A	I	V	G	R	Y	S	A	Y	V	-	-	P	-	-	T	-	plas	plas	plas

Table 2.2 continued.

		NPA I	NPA II	ar/R					Froger's positions					Other regulatory residues							Pred. subcellular loc.		
	Gene Name	LB	LE	H2	LC	H5	LE1	LE2	P1	P2	P3	P4	P5	Asp28	Glu31	Ser115	Ser188	His193	Ser274	Ser277	mPLoc	WolfPsort	Yloc
TIP	NtTIP1;1s	NPA	NPA	H	F	I	A	V	A	S	A	Y	W	-	-	T	-	-	-	-	tono	plas	plas
	NtTIP1;1t	NPA	NPA	H	F	I	A	V	A	S	A	Y	W	-	-	T	-	-	-	-	tono	plas	plas
	NtTIP1;2s	NPA	NPA	H	F	I	A	V	A	S	A	Y	W	-	-	T	-	-	-	-	tono	plas	pero
	NtTIP1;2t	NPA	NPA	H	F	I	A	V	A	S	A	Y	W	-	-	T	-	-	-	-	tono	plas	pero
	NtTIP1;3s	NPA	NPA	H	F	I	A	V	T	S	A	Y	W	-	-	T	-	-	-	-	tono	plas	plas
	NtTIP1;3t	NPA	NPA	H	F	I	A	V	T	S	A	Y	W	-	-	T	-	-	-	-	tono	plas	plas
	NtTIP1;4t	NPA	NPA	H	F	I	A	V	A	S	A	Y	W	-	-	T	-	-	-	-	tono	plas	cyto
	NtTIP2;1s	NPD	NPA	H	H	I	G	R	V	S	A	Y	W	-	-	T	-	-	-	-	tono	tono	plas
	NtTIP2;1t	NPA	NPA	H	H	I	G	R	V	S	A	Y	W	-	-	T	-	-	-	-	tono	plas	plas
	NtTIP2;2s	NPA	NPA	H	H	I	G	R	V	S	A	Y	W	-	-	T	-	-	-	-	tono	tono	plas
	NtTIP2;3s	NPA	NPA	H	H	I	G	R	V	S	A	Y	W	-	-	T	-	-	-	-	tono	tono	Extra
	NtTIP2;3t	NPA	NPA	H	H	I	G	R	V	S	A	Y	W	-	-	T	-	-	-	-	tono	tono	plas
	NtTIP2;4s	NPA	NPA	H	H	I	G	R	V	S	A	Y	W	-	-	T	-	-	-	-	tono	plas	plas
	NtTIP2;5s	NPA	NPA	H	H	I	G	R	V	S	A	Y	W	-	-	T	-	-	-	-	tono	plas	plas
	NtTIP2;5t	NPA	NPA	H	H	I	G	R	V	S	A	Y	W	-	-	T	-	-	-	-	tono	plas	plas
	NtTIP3;1s	NPA	NPA	H	F	I	A	R	A	A	A	Y	W	-	-	S^P	-	-	-	-	tono	plas	plas
	NtTIP3;1t	NPA	NPA	H	F	I	A	R	A	A	A	Y	W	-	-	S^P	-	-	-	-	tono	plas	plas
	NtTIP3;2t	NPA	NPA	H	F	V	G	R	A	A	A	Y	W	-	-	S^P	-	-	-	-	tono	plas	plas
	NtTIP4;1s	NPA	NPA	H	H	I	A	R	V	S	A	Y	W	-	-	T	-	-	-	-	tono	tono	plas
	NtTIP4;1t	NPA	NPA	H	H	I	A	R	L	S	A	Y	W	-	-	T	-	-	-	-	tono	tono	plas
	NtTIP5;1s	NPA	NPA	N	H	V	G	Y	T	S	A	Y	W	-	-	S	-	-	-	-	tono	plas	plas
	NtTIP5;1t	NPA	NPA	N	H	V	G	Y	T	S	A	Y	W	-	-	S	-	-	-	-	plas, vac	plas	plas
SIP	NtSIP1;1t	NPT	NPA	T	E	V	P	N	M	A	A	Y	W	-	-	H	-	-	-	-	plas	plas	Extra
	NtSIP1;2s	NPA	NPA	T	G	V	P	N	I	A	A	Y	W	-	-	T	-	-	-	-	plas	tono	Extra
	NtSIP1;2t	NPA	NPA	T	G	V	P	N	I	A	A	Y	W	-	-	T	-	-	-	-	plas	tono	ER
	NtSIP2;1s	NPL	NPA	H	R	H	G	S	F	V	A	Y	W	-	-	S	-	-	-	-	plas	chlo	Extra
	NtSIP2;1t	NPL	NPA	H	R	H	G	S	F	V	A	Y	W	-	-	S	-	-	-	-	plas	chlo	Extra
XIP	NtXIP1;6t	NPV	NPA	A	-	T	A	R	V	C	A	F	W	-	-	S	-	-	-	-	plas	plas	plas
	NtXIP1;6s	NPV	NPA	A	-	T	A	R	V	C	A	F	W	-	-	S	-	-	-	-	plas	plas	plas
	NtXIP1;7s	NPV	NPA	I	-	T	A	R	V	C	A	F	W	-	-	S	-	-	-	-	plas	plas	pero
	NtXIP1;7t	NPV	NPA	I	-	T	A	R	V	C	A	F	W	-	-	S	-	-	-	-	plas	plas	pero

2.4.3.5 NtSIP subfamily

While only comprising of 5 genes, the NtSIP subfamily had low sequence conservation, with Loop A the least conserved (Figure 2. 3B). The first NPA motif varied with NP<u>A</u>/<u>T</u>/<u>L</u> combinations (Table 2. 2). Substantial variation was also was found in other key residues with completely different configuration of residues in the ar/R and Froger's P1-P2 between NtSIP1 and NtSIP2 proteins (Table 2. 2). The N- terminal tail of NtSIPs were distinctly shorter than other sub-families (~7aa) (Figure 2. 3A).

2.4.3.6 NtXIP subfamily

The XIPs are a small sub-family with high sequence identity (~75%). The first NPA motif is replaced by a NPV motif in all four NtXIP proteins (Table 2. 2).

There is a strong consensus in the residues residing in the Froger's and dual NPA motifs, with the only variation being I/A at ar/R H2 (Table 2. 2). Concordant with other studies of XIPs, the loop C of NtXIP is substantially longer (~38aa) compared to that of other subfamilies (Gupta and Sankararamakrishnan 2009). NtXIPs have the conserved phosphorylated Ser115, although it was not a predicted phosphorylation target (Table 2. 2). The C-terminal tail of NtXIPs contained a single serine residue which was not predicted to be phosphorylated (data not shown).

2.4.4 Subcellular localisation of tobacco AQPs in planta

AQPs can facilitate diffusion of a range of substrates across various plant membranes and the specific membrane localisation can vary between the different subfamilies. Computational prediction programs can be used as an initial inference of subcellular localisation to further help elucidate putative biological activities and physiological functions of candidate proteins (Briesemeister et al. 2010). We conducted subcellular prediction analyses using three commonly used software programs, Plant-mPLoc, Wolf Psort and YLoc (see materials and methods). Consistency in prediction across the three programs was found for 35 (46%) of NtAQPs (Table 2. 2). Consensus in predicted localisation was mainly observed for the PIP2s and the NIPs, which generally predicted to be plasma membrane (PM) localised. The TIPs and SIPs appeared to have the most contrasting predictions in subcellular localisation results, with TIP localisations ranging between tonoplast, PM, peroxisome, cytoplasmic and extra cellular localisation; and SIPs having PM, tonoplast, chloroplast, ER and extra cellular localisations across the 3 prediction tools (Table 2. 2).

As an initial exploration into the actual membrane targeting of the different NtAQP subfamilies and to validate the software-predictions, the subcellular localisations of a representative tobacco PIP, TIP and NIP was visualised *in plant*a (Arabidopsis) using NtAQP-GFP fusions and confocal microscopy. SIPs are a smaller AQP subfamily and were not included in this analysis. The NtXIPs have already been established as localising to the PM (Bienert et al. 2011).

In plant cells, the vacuole occupies much of the internal volume pushing the cytoplasm and its contents to the periphery. This can give the illusion of PM localisation even for cytosolic proteins such as 'free' GFP, especially if only examined as a 2D-optical slice at the whole cell level (Figure 2. 4Ai). To correctly assess the subcellular localisation of the NtAQPs in plant cells, we used established GFP marker lines to define the characteristic features of proteins localised to the cytoplasm, plasma membrane (PM), ER, and tonoplast (tono) (Figure 2. 4; (Nelson et al. 2007)). Key defining features between the four sub-cellular compartments were identified by examining localisation in the vicinity of the nucleus, the topology of the signal, and 3D renders of series Z-stack images of the cells (Figure 2. 4B-G). Consistent with the subcellular structures being labelled, the PM:GFP marker localised exclusively to the periphery of the cell when adjacent to the nucleus (Figure 2. 4Bii), ER:GFP wrapped around the nucleus (Figure 2. 4Dii), and Tono:GFP localised internally to the nucleus leaving a signal void on the cell surface (Figure 2. 4Fii). PM:GFP produced a sharp defined integration with the cell margin (Figure 2. 4Biii), featuring as an outer shell in the 3D render (Figure 2. 4Biv). The ER:GFP peripheral signal was mottled in appearance (Figure 2. 4Di), consisting of localised bright specks with distinct regions of no signal (Figure 2. 4Diii), that appeared as a 'web' in the 3D render (Figure 2. 4Div). Tono:GFP was present as large undulating 'sheets' of signal associated with the trans-vacuolar strands (tonoplast-delimited cytoplasmic tunnels) and folds of vacuole membrane (tonoplast) (Figure 2. 4Fi-iv), which had a distinct 'wavy' topology (Figure 2. 4Fiii).

Distinct *in planta* subcellular localisation patterns were observed for the PIP, TIP and NIP NtAQPs, consistent with the known membrane targeting properties of these different AQP subfamilies (Figure 2. 4C,E,G). The GFP signal of the representative PIP (NtPIP2;5t) appeared sharp and uniformed around the cell periphery, with signal running external to the nucleus and forming a smooth outer shell in the 3D render with no discernible signal in any internal structures (Figure 2. 4C). This pattern was concordant with a PM:GFP marker (Figure 2. 4B), indicating a strong integration of NtPIP2;5t into the PM.

The representative NtNIP (NtNIP2;1s), had features indicating it co-localises to the PM and ER. The peripheral localised NIP:GFP signal was mottled in

appearance with distinct specks of intense bright signal similar to the ER marker. However, unlike the ER marker, these specks were dispersed along a consistent basal signal continuous around the cell periphery, which indicates PM localisation (Figure 2. 4Ei-iii). The 3D render of NIP:GFP further demonstrated the shared shell-like PM signal overlapping the mottled web-like ER patterned signal (Figure 2. 4Eiv). The localisation of the representative NtTIP (NtTIP 1;1s) is consistent with integration into the tonoplast. NtTIP:GFP showed a uniform yet diffuse localisation within the cell, with signal circling the nucleus on the cytosolic but not plasma membrane side (Figure 2. 4Gi-ii). The labelled membrane had a wavy topology consistent with the tonoplast, with occurrence of internal membranes resembling transvacuolar strands (Figure 2. 4Giii-iv).

The PM integration of *NtPIP2;5* was predicted by all 3 software programs, whereas the tonoplast localisation of *NtTIP1;1s* was only predicted by Plant-mPLoc. Lastly, the PM localisation of *NtNIP2;5s* was anticipated in all 3 programs, but none predicted its co-localisation with the ER (Table 2. 2).

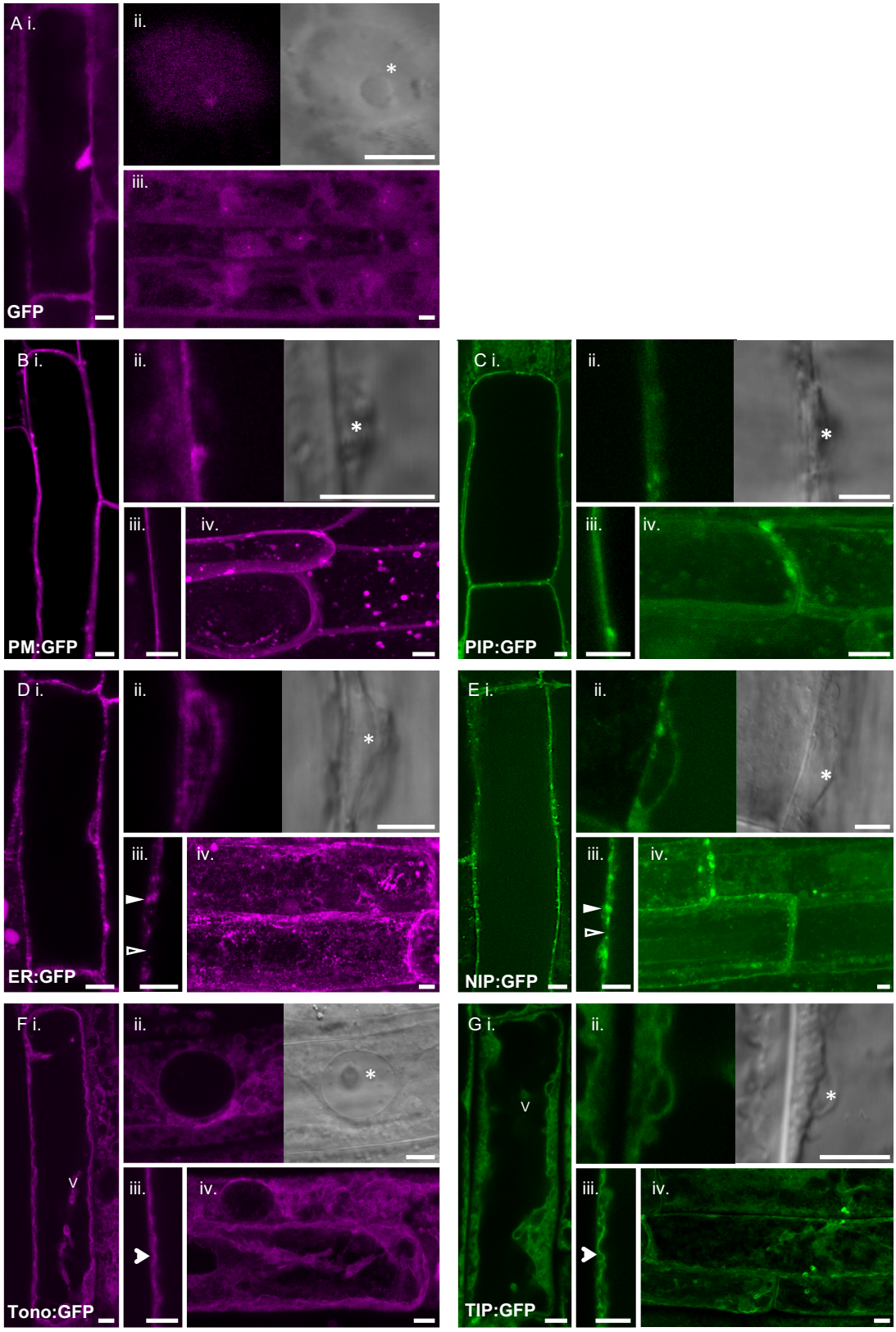


Figure 2. 4 In planta sub-cellular localisation of PIP, TIP and NIP aquaporins*. Confocal images of root cortical cells of transgenic 8-day-old Arabidopsis seedlings.* ***(A, B, D, F)*** *GFP marker lines; false coloured purple.* ***(C, E, G)*** *NtAQP:GFP lines; false coloured green.* ***Subpanels (i-iv)*** *are;* ***(i.)*** *Optical cross-*

section midway through a root cortical cell. ***(ii)*** *GFP signal associated with nucleus; confocal image (left) DIC image (right).* ***(iii.)*** *Close-up of cell peripheral margin.* ***(iv.)*** *Maximum intensity projections compiled from z-stacks.* ***(A)*** *GFP-only localisation.* ***(B)*** *Plasma membrane (PM:GFP) marker.* ***(C)*** *NtPIP2;5t (PIP:GFP).* ***(D)*** *Endoplasmic reticulum (ER:GFP) marker. The ER is known not uniformly be present around the cell periphery which is reflected by regions of bright GFP signal (solid arrowhead) interspersed with regions of no GFP signal (open arrowhead).* ***(E)*** *NtNIP2;1s (NIP:GFP).* ***(F)*** *Tonoplast (Tono:GFP) marker showing characteristic features of the tonoplast membrane including, transvacuolar strands (v) and general undulating appearance (arrow).* ***(G)*** *NtTIP1;1s (TIP:GFP). Notable sub-cellular features are marked by a; asterisks for the nucleus, 'V' for transvacuolar strands, arrowheads indicate instances of varied brightness (solid = high signal, empty = no signal) in GFP fluorescence in* ***D (iii)*** *and* ***E (iii)****, or undulations of the tonoplast in* ***F (iv)*** *and* ***G (iv)****. Scale bar 5µm.*

2.4.5 Parental association and recent evolutionary history of Tobacco AQPs

The distinctive phylogenetic pairing of most NtAQPs in our initial phylogenetic characterisation, is likely characteristic of the recent evolutionary origin of tobacco, which arose from an allotetraploid hybridisation event between *N. sylvestris* and *N. tomentosiformis* only ~0.2M years ago (Edwards et al. 2017; Sierro et al. 2014). To explore the evolution of the tobacco AQP family, we identified the AQP gene families in the two parental lines using NtAQP nucleotide coding sequences as queries in BLAST searches. Initially, 40 and 41 AQPs were identified in both N. s*ylvestris* and *N. tomentosiformis* respectively, which is comparable to the number of AQP genes found in the related diploid species of tomato and potato (Table 2. 3). Tobacco having approximately double the complement of AQP genes to its parents, is consistent with a recent allotetraploid hybridisation event where insufficient time has passed for a reduction in genome size through redundant gene-loss. The introduction of the parental *N. sylvestris* and *N. tomentosiformis* AQPs into the NtAQP phylogeny, transformed the majority of the distinct NtAQP phylogenetic pairs into small clades of four genes where each of the paired NtAQPs was now clearly associated with an AQP from one of the two parents (e.g. NtPIP1;1 sub-clade, Figure 2. 5). This phylogenetic relationship confirmed that the distinctive phylogenetic pairing of NtAQPs corresponds to orthologous 'sister' genes arising from hybridisation, with both parental genomes having contributed one AQP gene

to each tobacco sister pair (Figure 2. 5). Initially 30 sister gene pairs were identified that had a clear match to an orthologous gene from both *N. sylvestris* and *N. tomentosiformis* (Figure 2. 5). The ancestral origin of the *NtAQP* genes were denoted in the nomenclature by the addition of a suffix 's' or 't' (e.g. *NtPIP1;1s* and *NtPIP1;1t*), to indicate a *N. sylvestris* or *N. tomentosiformis* lineage, respectively.

One *NtAQP* gene had no resolved match to a *N. sylvestris* or *N. tomentosiformis* parental AQP and was assigned a suffix 'x' (*NtPIP2;1x*). The lack of a clear parental match to *NtPIP2;1x* likely means that the orthologous gene has been lost in the parental genome post tobacco emergence, or the orthologous parental AQP was not identified due to incomplete coverage of sequencing data. Either way, the presence of this gene in the tobacco genome allows us to infer its presence in a parental genome at the time of hybridisation. We predict that *NtPIP2;1x* was inherited from *N. tomentosiformis*, as it occurs in a distinct clade with a tobacco sister gene (*NtPIP2;1s*) and an orthologous *N. sylvestris* AQP (*N.sylPIP2;1*), but lacks a *N. tomentosiformis* progenitor ortholog (orange box, Figure 2. 5). As such, assigning *NtPIP2;1x* as a *N. tomentosiformis* descendant, brings the total number of AQPs in the parental genomes to 40 in *N. sylvestris* and 42 in *N. tomentosiformis*, with the total number of genes within the PIP, NIP and TIP sub-families being very similar to those of tomato and potato (Table 2. 3).

The phylogenetic analysis also revealed recent evolutionary events in the tobacco*, N. sylvestris* and *N. tomentosiformis* AQP families. These events were recognised by deviations from the conventional four-gene small sub-clade groupings comprised of the tobacco sister genes and their respective parental orthologs. Seven *AQP* gene loss events were recognised in *N. sylvestris*, six of which occurred prior to the tobacco hybridisation event as the given *AQP* was absent in both *N. sylvestris* and tobacco (blue stars, Figure 2. 5). In several cases, the remnants of the eroding *N. sylvestris* pseudo gene were also inherited and identifiable in the tobacco genome (e.g. *SIP1;1* and *PIP2;7*; Figure 2. 5). Only a single gene loss event was recognised in *N. tomentosiformis*, with no representative *NIP1;1* orthologs identified in either *N. tomentosiformis* or tobacco (red star, Figure 2. 5). Five parental AQP genes have been lost in tobacco, three from *N. tomentosiformis* and two from *N. sylvestris* origins (green stars, Figure 2. 5). Instances of gene gains were

also evident in both parental species prior to the tobacco hybridisation event (purple and orange stars, Figure 2. 5). These gained genes were distinct in the phylogenies as they did not uniquely match a specific Solanaceae gene ortholog, appearing instead as a duplicate copy of an existing *AQP* gene within the tobacco parental species (Figure S2. 3). Four AQP gene gain events occurred in *N. sylvestris*, two of which (*N.sylNIP3;1* and *N.sylPIP1;2*), began redundant gene erosion prior to tobacco hybridization (purple stars, Figure 2. 5). The third, *N.sylPIP2;11b*, is retained as a functional unit in *N. sylvestris* but has eroded in tobacco; hence the designation 'b' as opposed to a unique numerical identifier. The fourth gene, *N.sylPIP1;8*, has been retained in both *N. sylvestris* and tobacco as a functional gene (purple star, Figure 2. 5). A single gene duplication event was recognized in *N. tomentosiformis*, giving rise to *PIP2;2* and *PIP2;3* orthologs which were both inherited and subsequently retained as functional genes in tobacco (orange star, Figure 2. 5).

Table 2. 3. Summary of total AQPs currently identified within Solanaceae. *Tomato and potato AQP families were characterised by Reuscher et al. (2013) and Venkatesh et al. (2013), respectively. N.sylvestris, N.tomentosiformis and tobacco gene families were established in this study. Numbers indicate genes occurring in each AQP sub-family (PIP, TIP, NIP, SIP and XIP). The number of identified N.tomentosiformis PIPs through the BLAST searches was 15 (noted by superscript value), however, we predict the total number of PIPs at the time of hybridisation to be 16 due to NtPIP2;1x likely being of N.tomentosiformis origin. This brings the overall number of N.tomentosiformis AQPs to 42.*

Solanaceae species	PIPs	TIPs	NIPs	SIPs	XIPs	Total AQPs
Tomato (*Reuscher et al. 2013)*	14	11	10	4	6	45
Potato (*Venkatesh et al. 2013)*	15	11	10	3	4	47
N. sylvestris	**15**	**11**	**10**	**2**	**2**	**40**
N. tomentosiformis	**16**$^{(15)}$	**13**	**8**	**3**	**2**	**42**$^{(41)}$
Tobacco (N. tabacum)	29	22	16	5	4	**76**

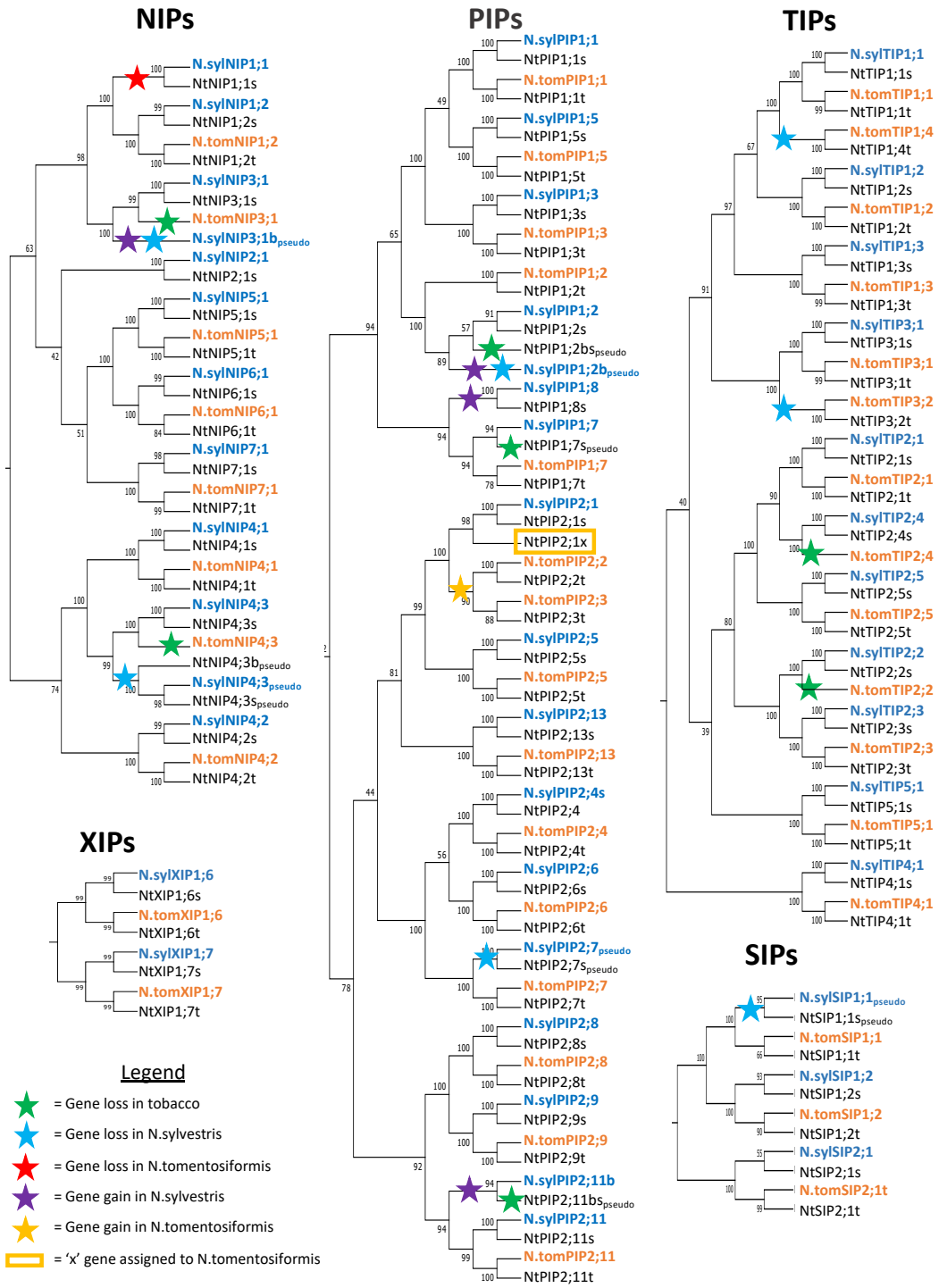


Figure 2. 5 Phylogenetic relationship of tobacco, N. sylvestris and N. tomentosiformis AQPs. *Phylogenetic trees for each AQP sub-family were generated using the neighbour-joining method from MUSCLE alignments of nucleotide coding sequences. Confidence levels (%) of branch points generated through bootstrapping analysis (n=1000). N. sylvestris (N. syl) and N. tomentosiformis (N. tom) AQPs are colour coded in blue and orange, respectively. Green stars indicate a loss of a parental gene in tobacco post-hybridisation; Blue and Red stars indicate gene loss events in N. sylvestris and N. tomentosiformis, respectively. Purple and Yellow stars indicate pre-hybridisation gene gain events in N. sylvestris and N .tomentosiformis, respectively.*

2.4.6 Tobacco AQP gene expression analysis

To provide insight into possible physiological roles of the various AQP isoforms, publicly available whole transcriptome RNA-seq datasets (Sierro et al. 2014; Edwards et al. 2017) were processed and analysed to compare organ-specific expression patterns of the 76 tobacco AQPs. The Sierro et al. (2014) transcriptome dataset from the TN90 cultivar was chosen for analyses as it provided the most extensive sampling of tissues at various developmental stages (young leaf, mature leaf, senescent leaf, stem, root, young flower, mature flower, senescent flower and dry capsules).

Although the *NtAQP* sister genes are highly homologous in their nucleotide coding sequences (~96.5%), the SNPs that are present occur at a frequency and distribution enabling unique mapping of reads to differentiate between sister genes. In the TN90 dataset, we detected expression from 75 out of 76 NtAQPs, with only *NtXIP1;4t* having no mapped mRNA reads. However, *NtXIP1;4t* is an expressed gene, as low transcript abundance was detected in the K326 cultivar (data not shown). Among the sub-families, gene expression of PIPs and TIPs was generally greater than for SIPs, XIPs and NIPs (Figure 2. 6A). Among the most highly expressed *NtAQPs*, *PIP1;5s*, *PIP1;5t, PIP1;3s* and *PIP1;3t* stood out as being constitutively expressed in all major plant organs, while *TIP1;1s* and *TIP1;1t*, were present in all tissues except for the dry capsule (Figure 2. 6A). Some highly expressed genes also showed a level of tissue specificity, with *NIP4;1s* and *NIP4;1t* expressed only in flowers, and *TIP3;1s, TIP3;1t* and *TIP3;2t* predominantly in the flower capsule (Figure 2. 6A).

To examine differential expression between plant organs, the expression levels of a given *AQP* were standardised relative to its highest expressing tissue (Figure 2. 6B). *AQP*s with a broad expression distribution throughout the plant could be readily identified (e.g. *SIP1;2* and *PIP1;5* sister pairs, Figure 2. 6B). Other AQPs show tissue specific expression: young flowers (*PIP2;11s* & *PIP2;11t*; *NIP2;1s*), leaves (*XIP1;6s*; *PIP2;5t*; *PIP2;1x*) or roots (*TIP1;*2, *TIP1;3* and *TIP2;3* genes). At the sub-family scale, *NtNIPs* and *NtTIPs* are found to be preferentially expressed in roots, stems and flowers, with a low tendency for expression in leaves (Figure 2. 6B).

NtPIPs and *NtSIPs* are more broadly expressed, while there is no expression of *NtXIPs* in either the stem or dry capsule (Figure 2. 6B). Within sub-families we see gene members with specialised or preferential tissue expression. For example, some NtPIPs preferentially expressed in the roots (*PIP1;1s & PIP1;1t*; *PIP2;4s* & *PIP2;4t*), others express preferentially in leaves (e.g. *PIP2;5t* & *PIP2;1x*), while *PIP2;11s* & *PIP2;11t* have become specialised in young flowers (Figure 2. 6B). Discrete tissue-specific specialisation was also observed for members of the other families, for instance, *TIP3;1* and *TIP3;2* genes express only in dry capsule (seeds), and expression of *NIP4;1* and *NIP4;2* was only detected in flowers (Figure 2. 6B).

Next we compared differences in expression between sister genes to explore possible functional divergence. In general, sister gene pairs showed matching patterns of tissue-specific expression (Figure 2. 6B). However, of the 31 proposed sister gene pairs, 26 showed differential expression levels in at least one tissue (Figure 2. 6C). In the majority of these instances a single sister gene of the pair was more highly expressed in several plant organs. Examples include, *NIP5;1s, SIP2;1t, SIP1;2t, PIP2;6t, PIP2;4s, PIP1;3t* and *PIP1;1s*. There were also instances of contrasting expression where sister genes show distinctions in preferential expression between plant organs. For example, *TIP3;1s* with 4-fold higher expression in the capsule compared to its sister pair *TIP3;1t*, which is expressed >10-fold higher in roots (Figure 2. 6C). Further examples of contrasting expression include, *NtPIP2;5t* (leaves) against *NtPIP2;5s* (roots) and *NtNIP6;1s* (leaves and dry capsule) against *NtNIP6;1t* (roots) (Figure 2. 6C).

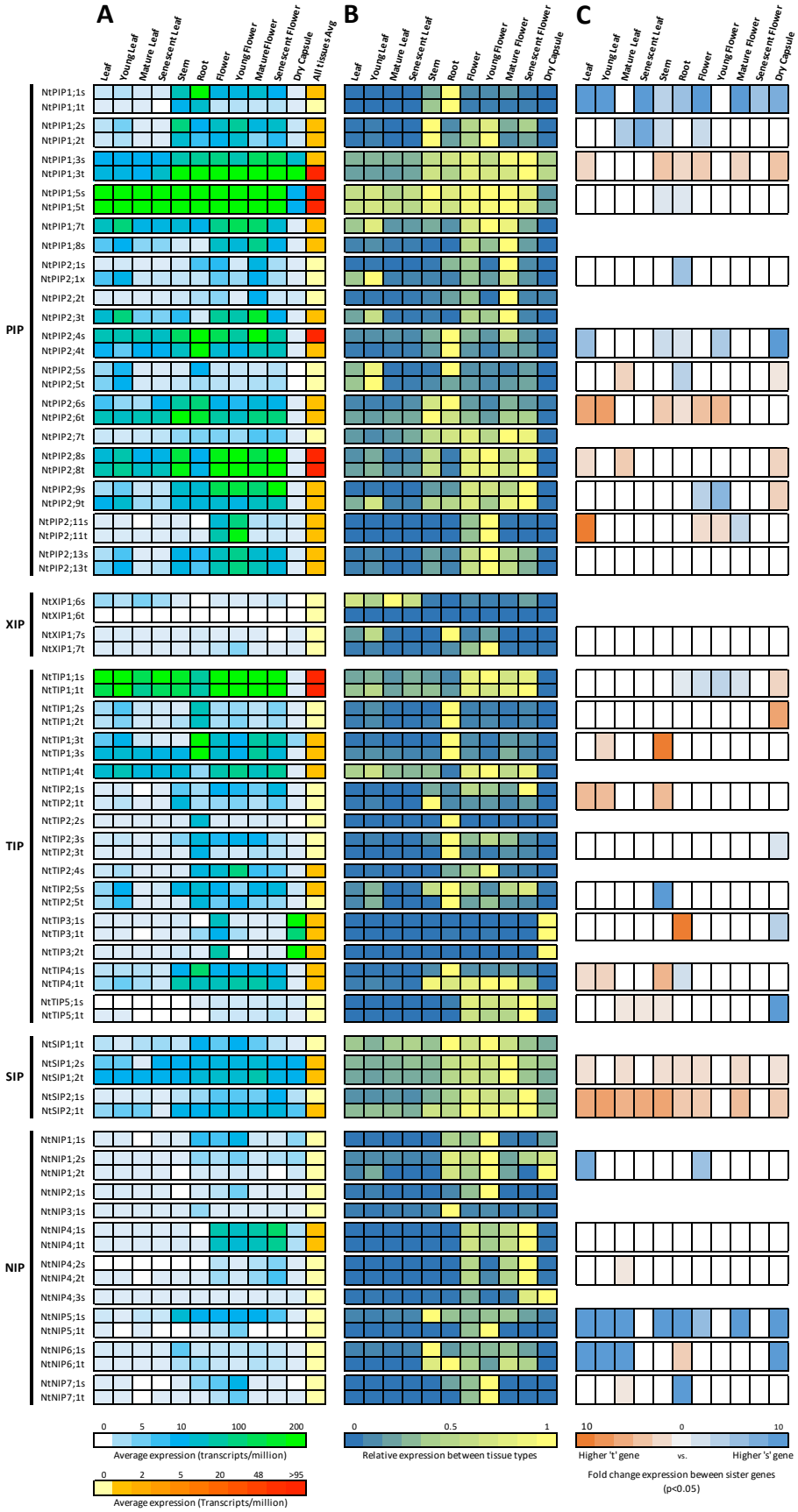

A
B
C
Leaf
Young Leaf
Mature Leaf
Senescent Leaf
Stem
Root
Flower
Young Flower
Mature Flower
Senescent Flower
Dry Capsule
All tissues Avg
NtPIP1;1s
NtPIP1;1t
NtPIP1;2s
NtPIP1;2t
NtPIP1;3s
NtPIP1;3t
NtPIP1;5s
NtPIP1;5t
NtPIP1;7t
NtPIP1;8s
NtPIP2;1s
NtPIP2;1x
NtPIP2;2t
NtPIP2;3t
NtPIP2;4s
NtPIP2;4t
NtPIP2;5s
NtPIP2;5t
NtPIP2;6s
NtPIP2;6t
NtPIP2;7t
NtPIP2;8s
NtPIP2;8t
NtPIP2;9s
NtPIP2;9t
NtPIP2;11s
NtPIP2;11t
NtPIP2;13s
NtPIP2;13t
PIP
NtXIP1;6s
NtXIP1;6t
NtXIP1;7s
NtXIP1;7t
XIP
NtTIP1;1s
NtTIP1;1t
NtTIP1;2s
NtTIP1;2t
NtTIP1;3t
NtTIP1;3s
NtTIP1;4t
NtTIP2;1s
NtTIP2;1t
NtTIP2;2s
NtTIP2;3s
NtTIP2;3t
NtTIP2;4s
NtTIP2;5s
NtTIP2;5t
NtTIP3;1s
NtTIP3;1t
NtTIP3;2t
NtTIP4;1s
NtTIP4;1t
NtTIP5;1s
NtTIP5;1t
TIP
NtSIP1;1t
NtSIP1;2s
NtSIP1;2t
NtSIP2;1s
NtSIP2;1t
SIP
NtNIP1;1s
NtNIP1;2s
NtNIP1;2t
NtNIP2;1s
NtNIP3;1s
NtNIP4;1s
NtNIP4;1t
NtNIP4;2s
NtNIP4;2t
NtNIP4;3s
NtNIP5;1s
NtNIP5;1t
NtNIP6;1s
NtNIP6;1t
NtNIP7;1s
NtNIP7;1t
NIP
0 5 10 100 200
Average expression (transcripts/million)
0 2 5 20 48 >95
Average expression (Transcripts/million)
0 0.5 1
Relative expression between tissue types
10 0 10
Higher 't' gene vs. Higher 's' gene
Fold change expression beween sister genes
(p<0.05)

Figure 2. 6 Expression patterns of NtAQP genes in different tissues. (A) Absolute NtAQP gene expression. *Heatmap of gene expression (transcripts per million) of NtAQPs across different tissues. Green shading represents higher expression, graduating to a light blue for lower expression, as per key. Included in the last column is the average gene expression across all tissues examined; red shading for high expression moving towards yellow for low expression, as per key.* ***(B) Relative expression compared to the highest expressing tissue for the given NtAQP.*** *Heatmap of tissue-specific gene expression with values standardised to the tissue showing the highest expression for that given NtAQP. Yellow indicates high expression graduating towards blue for low expressing tissue.* ***(C) Comparison of expression patterns between AQP sister genes.*** *Heatmap of significant fold change differences in expression ($p<0.05$) between sister genes across the different examined tissues. Blue indicates higher expression of the 's' gene and orange higher expression of the 't' gene.*

2.4.7 Gene expression conservation within Solanaceae

As an initial means of exploring conservation in biological activities and physiological functions between AQP orthologous of different species, we compared tissue-specific expression levels of NtAQPs with their orthologs from the closely related tomato and potato species. This was done by comparing the relative gene expression across root, leaf and floral tissues of *AQP* genes we have identified as being orthologs between the Solanaceae species (e.g. *NtPIP1;1s* & *NtPIP1;1t* in tobacco, *SlPIP1;1* in tomato and *StPIP1;2* in potato; listed Table S2. 2). We were able to perform this analysis on the PIPs, TIPs, NIPs and SIPs but not the XIPs given the previously mentioned difficulty of assigning orthology between the species. Even randomised pairwise comparisons of expression patterns between NtXIPs with those of tomato and potato, could not find consensus patterns, hinting further towards the unique intra-species diversification of XIPs within the Solanaceae (Figure S2. 5).

In the majority of instances (25 of 36 Solanaceae AQP ortholog sets), the tobacco sister genes had similar patterns of relative expression levels between the three organs to their orthologs from both tomato and potato, implying similar physiological roles for the orthologs across the Solanaceae family (e.g. *NIP1;1*, *NIP3;1*, *NIP4;*2, *PIP2;6*, *PIP2;9*, *PIP2;11*, *TIP5;1*, and *SIP1;1* ; Figure 2. 7). Some deviations in tissue-specific expression patterns were observed between orthologs, suggesting possible species-specific functional diversification. The predominant observed deviations were instances where either the tobacco, tomato or potato AQP differed in their tissue-specific expression pattern compared to the orthologs from

the other Solanaceae. Examples include; the tobacco *NtNIP5;1*, *NtPIP1;2, TIP1;1*; the tomato *SlPIP2;8*, *SlTIP2;*1, *SlTIP3;1,* and *SlTIP3;2* genes; and the potato *StPIP1;2* (*NtPIP1;1* ortholog), *StTIP1;2, StTIP1;1* (*NtTIP1;3* ortholog) and *StTIP2;4* (*NtTIP2;3* ortholog) genes (Figure 2.7). Additionally, we observed one case where a NtAQP sister gene (*NtPIP2;5s*), differed in expression from the tomato, potato and its NtAQP "t" sister gene; suggesting a potential diversification in gene function within tobacco. More complex deviations were also observed involving tobacco sister genes having contrasting expression to each other, that matched a similar contrast in expression between the tomato and potato orthologs (e.g. *NtPIP2;1* and *NtNIP6;1* sister genes; Figure 2.7).

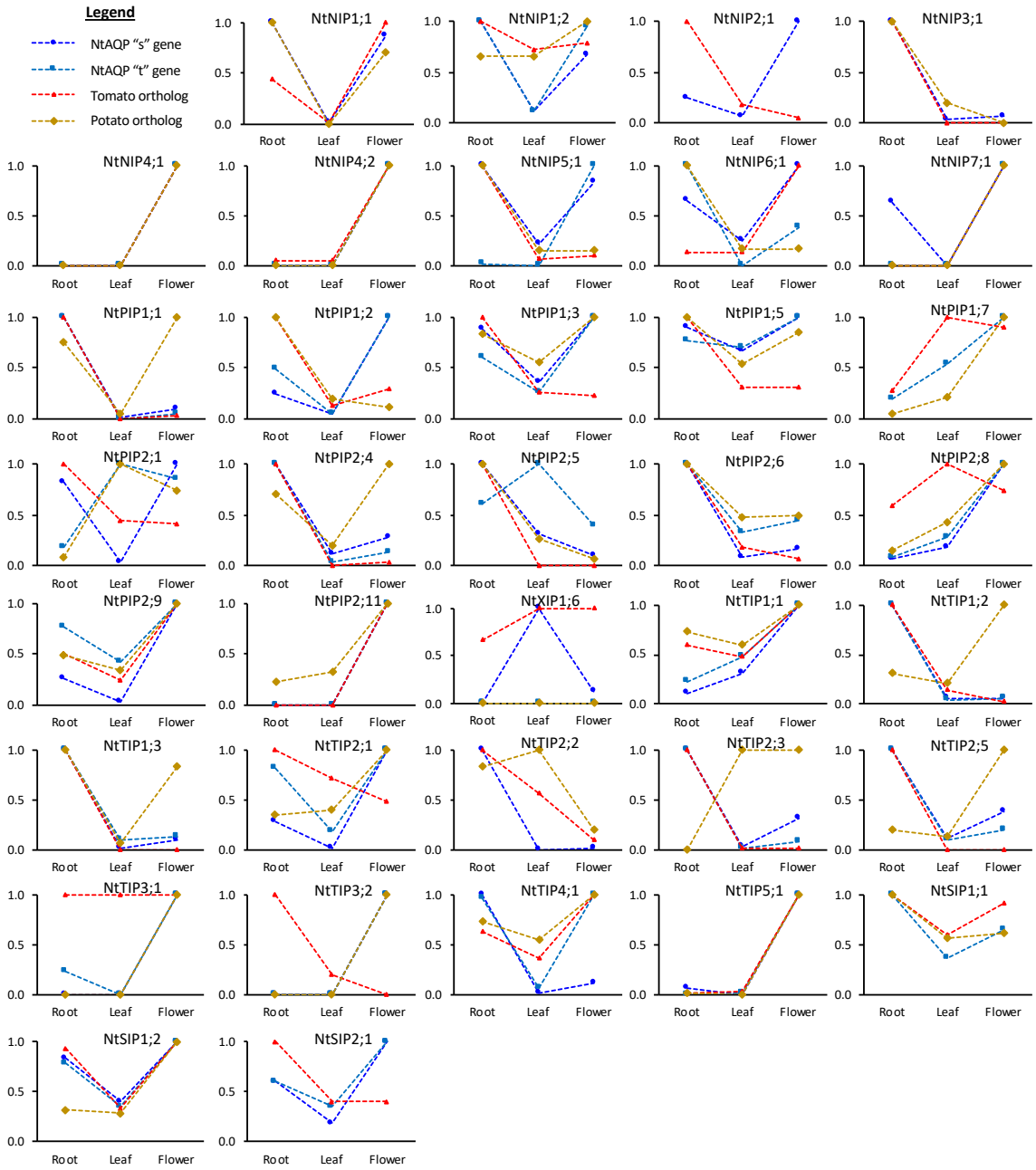


Figure 2. 7 Tissue-specific gene expression patterns of AQP isoforms in tobacco, tomato and potato. *Graphs contain relative gene expression (standardised to highest expressing tissue) across root, leaf and flower tissues for tobacco sister genes (light and dark blue) and their corresponding tomato (red) and potato (brown) orthologs as listed in Table S2.2.*

2.5 Discussion

The growing amount of research into AQPs is greatly advancing our understanding of their diversity and functional roles, towards manipulating them to potentially enhance plant performance and resilience to environmental stresses (Li et al. 2014; Uehlein et al. 2003; Maurel et al. 2008; Xu et al. 2018; Groszmann et al. 2017). The establishment of the tobacco AQP gene family allowed us to efficiently contribute to the current knowledge of AQP biology by, comparing regions of homology within and across closely related species, analysing pore-lining residues, identifying key structural characteristics, and providing necessary information and candidates for future functional screens. Furthermore, elucidating orthology between the already characterised tomato (Reuscher et al. 2013) and potato (Venkatesh et al. 2013) AQPs, enables comparisons between isoforms across these Solanaceae species, which will facilitate the translation of knowledge from tobacco into its closely related and horticulturally important crop species.

2.5.1 NtAQP protein sequence analysis and associations with AQP function

We found that the tobacco AQP family comprises of 76 members, making it one of the largest AQP families characterised to date; second only to the polyploid canola (*Brassica napus*) with 121 members (Sonah et al. 2017; Yuan et al. 2017). Correctly defining and analysing the NtAQP protein structures (Figure 2.3A), sequence homology (Figure 2. 3B), and comparison of functionally relevant residues (Table 2.2), helps towards predicting potential permeating substrates, post-translational regulation, and subcellular localisations. AQP monomers have a highly conserved structure, with transmembrane (TM) segments providing a structural scaffold and defining the channel environment, with the connecting loops also having significant roles in channel function (Törnroth-Horsefield et al. 2006). We found a high conservation in length and sequence identity of the NtAQP TM domains (Figure 2. 3); their variability likely constrained to maintain structural integrity of the AQP monomer (Berny et al. 2016). Additionally, conservation of critical residues in TM domains is essential for tetramer formation, with modifications leading to aberrant AQP oligomerisation (Yoo et al. 2016). NtAQP loops and termini had

notable differences in lengths (Figure 2.3A) and lower sequence conservation across subfamilies (Figure 2.3B); such variation has implications for AQP monomer interactions, pore accessibility and cellular membrane destinations (Gupta and Sankararamakrishnan 2009; Takano et al. 2017).

AQP solute selectivity are conferred through specific structural features of the AQP monomer's pore, and substrate interactions with pore-lining residues. We surveyed known specificity-determining residues across the NtAQPs, including the aromatic arginine (ar/R) filter, NPA domains, and Froger's positions (Wu and Beitz 2007; Sui et al. 2001; Froger et al. 1998; Hove and Bhave 2011). We observed an increased sub-family conservation in the loops harbouring these specificity-determining residues, in particular Loops B and E which have a direct role in forming the transmembrane pore (Figure 2. 3B). Each subfamily had their unique characteristic combination of amino acids at these locations concordant with known subfamily substrate specificities. For example, NtPIPs have more polar residues in their ar/R filter which is consistent with PIPs in general having the propensity to permeate water, whereas the NtNIPs have more hydrophobic amino acids in their ar/R filter, which predicts poorer water permeability and a preference for substrates such as ammonia, urea and metalloids instead (Wu and Beitz 2007; Hove and Bhave 2011).

Additional to the specificity-determining pore lining residues, post-translational modification of specific residues (e.g. through protonation or phosphorylation), also directly or indirectly determine the transport mechanics of the AQP monomer (Luang and Hrmova 2017). Plants rely on these secondary mechanisms to ensure tight regulation of AQPs, especially in response to stresses. Gating of the monomeric pore in response to external stimuli is a key control over AQP function. Among currently characterised residues involved in gating (listed in Table 2. 2), we found subfamily-specific conservation across the NtAQPs. For example, all NtPIPs had the Loop D Histidine (His193) which is highly conserved across all plant PIPs, and can be protonated in response to changes in cytosolic pH (e.g. flooding induce hypoxia), and leading to the closure of the PIP pores (Tournaire-Roux et al. 2003). pH regulated responses are important for AQP as is the C-terminal tail of the PIP proteins (Luang and Hrmova 2017). These facts drew

our attention to the identified Lysine/Arginine > Histidine substitution in the C-terminal tails of *NtPIP1;5* and *NtPIP2;11* (Addition file 2: Figure 2. S4). The normally positively charged Lysine/Arginine residue present in all other NtPIPs and highly conserved across plant PIPs in general, directly precedes a functionally important phosphorylated serine. Together this suggests a likely functional relevance of a positively charged residue at this position in PIP regulation. The Histidine present at the equivalent position in NtPIP1;5 and NtPIP2;11 can still obtain the conserved positive charge upon protonation, implying a possible novel pH control over the regulatory influences normally imposed by the PIP C-terminal tail.

Some sharing of gating mechanisms between NtAQPs from different sub-families can be inferred from our analysis. For example, the Loop B serine (Ser155) which in PIPs is involved in phosphorylation dependent disruption of N-terminal tail gating (Nyblom et al. 2009; Törnroth-Horsefield et al. 2006), is conserved in some members of the other NtAQP subfamilies. NtPIPs and NtNIPs both seem to be regulated by phosphorylation in their C-terminal tails given the abundance of serine residues. The phosphorylation state of the C-terminal tail is known to regulate channel activity and also control trafficking to the plasma membrane (Prak et al. 2008; Nyblom et al. 2009). Interestingly, the NtTIPs had a dearth of serine residues in the C-terminal tail, suggesting a lack of a C-terminal phosphorylation-dependent regulation mechanism. This perhaps is due to their integration into the vacuole membrane versus the plasma membrane integration of PIPs and NIPs. Consistent with differing regulatory requirements, we found that NtTIP2 and NtTIP5 proteins possessed a conserved histidine (His131) in loop C that is involved in a similar pH regulated gating of the pore to that of His193 in Loop D of PIPs and NIPs (Leitão et al. 2012; Soto et al. 2010) (erroneously reported as located to loop D of VvTnTIP2;1 in Leitão et al., 2012) . However, unlike the cytosolic PIP/NIP Loop D His193, the TIP Loop C His131 is likely orientated into the vacuole and thus responding to the vacuole contents and environment.

Other structural features NtAQP of note include: the longer Loop D of PIPs compared to the other sub-families which aids in its ability to cap the pore entrance (Törnroth-Horsefield et al. 2006); the substantially longer Loop A of PIPs compared

to the other NtAQPs, known to play a role in tetramer formation by mediating disulphide bonds between PIP1 and PIP2 isoforms (Roche and Törnroth-Horsefield 2017); the long N- and C-terminal tails of NtNIPs, important for protein regulation, trafficking, and protein-protein interactions (Diehn et al. 2015); the distinctly short N-terminal of SIPs associated with their intracellular destination into the ER (Maeshima and Ishikawa 2008); the long Loop C of NtXIPs, characteristically enriched with flexible glycine residues allowing it to tuck into the channel opening and interact with selectivity filter residues and permeating solutes (Newby et al. 2008; Gupta and Sankararamakrishnan 2009).

2.5.2 NtAQP subcellular localisation

Determining AQP subcellular localisations can help elucidate physiological roles within the plant. For instance, integration into plasma membrane indicates solute transport in and out of the cell; localisation to the tonoplast implies a role in vacuole storage; or retained in the ER membranes to coordinate shuttling of substrates and nutrients between plant membranes (Luu and Maurel 2013; Mizutani et al. 2006; Bienert et al. 2011; Ishikawa et al. 2005; Maeshima and Ishikawa 2008). We utilised sub-cellular localisation prediction software commonly used for fast *in silico* predictions of AQP isoform membrane integration. These software incorporate known sorting signals, amino acid composition and functional domains to generate results (Briesemeister et al. 2010). We found a 46% consensus of predicted AQP subcellular localisations using three software tools (Plant-mPLoc, WolfPsort and YLoc). The discrepancies most likely due to the complexity of AQP membrane integration processes (Takano et al. 2017). Overall, the PIP, NIPs and XIPs were predominantly localised to the PM; all of the Plant-mPloc and some of the WolfPsort outputs predicted tonoplast localisation for the TIPs; and the SIP localisations were quite varied. AQP:GFP fusions allowed us to visualise *in planta* sub-cellular localisation of representative AQPs and compare these to the software predictions. The representative PIP (*NtPIP2;5t*), NIP (*NtNIP2*;1s) and TIP (*NtTIP1;1s*) AQPs had distinct sub-cellular localisations, consistent with the known diversity of the AQP subfamilies across plants (Luu and Maurel 2013). Concordant with studies in other species (Takano et al. 2006; Choi

and Roberts 2007), we found that NtNIP co-localise to the PM and ER, which was not captured with the prediction software, instead reporting only PM integration.

2.5.3 Nicotiana AQP gene evolution

Tobacco recently descended from a allotetraploid hybridisation event between *N. sylvestris* and *N. tomentosiformis*, which are distantly related within the Nicotiana genus (Leitch et al. 2008). Genome downsizing is a widespread biological response to polyploidization, eventually leading to diploidization (Leitch and Bennett 2004). However, due to the short evolutionary time frame since its inception (0.2M years), tobacco has undergone a limited amount of genome downsizing. As a result, the NtAQP family is characteristically comprised of sister gene pairs, which we could assign to their given parental origins. Tobacco has lost only around 10% of it duplicated genes with no observed preferential gene loss from either parent (Edwards et al. 2017). Concordant with this estimation, 7 gene loss events (~8.6% of total inherited parental AQPs) were identified in tobacco, with 3 and 4 of these being redundant ortholog losses from the *N. sylvestris* and *N. tomentosiformis* genomes, respectively. According to our expression analysis, the NtAQP gene copies inherited from both *N. sylvestris* and *N. tomentosiformis* ('s' and 't' genes, respectively), were overall equally expressed, which agrees with broader genomic studies on tobacco (Edwards et al. 2017). The redundancy of the homeologs presumably would allow for one of the sister genes to accumulate mutations without immediate effect on fitness, most often leading to non-functionalisation (gene-loss), or in some instances sub-functionalisation or even neo-functionalisation. To this end, we observed instances where one AQP gene of a sister pair was consistently preferentially expressed throughout several plant organs (e.g. *PIP1;1s, PIP1;3t, SIP2;1t and NIP5;1s*); suggesting that the redundant lower-expressing sister gene could become non-functional over time. Alternatively, some sister genes showed distinct tissue-specific diversification, such as the *NtPIP2;5* gene pair, where the s- and t-genes were more highly expressed in the roots and leaves, respectively, and which maybe candidates for sub- or even neo-functionalisation.

We were able to identify several AQP gene gain and loss events between the parents since their divergence within the Nicotiana genus, ~15 Million years ago (Sierro et al. 2013). Both the *N. sylvestris* and *N.tomentosiformis* have a genome rich of repeat expansion (accumulation of transposable elements), making them nearly 3 times the size of that of tomato and potato (2.6 Gb vs. 0.9 Gb) (Sierro et al. 2013; Sato et al. 2012; Xu et al. 2011). Regardless of the discrepancy in genome size, there was close conservation of AQP ortholog numbers within these diploid Solanaceae species; with the PIPs and TIPs consistently the larger subfamilies. We saw a significant diversity in XIPs occurring in the Solanum (tomato and potato) and the Nicotiana species. This diversity manifested as discrepancies in isoform numbers between the species and as lower sequence identity; depicted in the phylogeny as a separation of tomato, potato and Nicotiana isoforms into distinct groups. XIPs are a more recently characterised AQP subfamily, with isoforms lacking in monocots and in Brassicaceae, and having a lower overall sequence identity compared to other AQP subfamilies (Danielson and Johanson 2008). The tomato and potato XIP are predominantly found clustered on a single chromosome, indicating that recent segmental gene duplications that occurred independently within tomato and potato likely explain the lack of direct gene orthology to tobacco XIPs (Venkatesh et al. 2015).

2.5.4 Gene expression analysis

Our NtAQP gene expression analysis revealed a wide range of patterns across tissue types, consistent with the known diversity of AQP functions (Hachez et al. 2006). It revealed that some AQPs had high levels across numerous tissues throughout the plant (e.g. *PIP1;3t* and *PIP1;5, TIP1;1* sister pairs), implicating involvement in broad spanning processes (e.g. substrate transport from roots to shoots to flowers), while others had highly organ specific expression (e.g. *TIP1;*3, *NIP4;1*, and *TIP3;1* sister genes, in roots, flowers and seed capsules respectively). In general, the XIPs and majority of NIPs had lower overall expression levels, although there is the possibility that their expression might change in response to a specific stimulus, or that they are expressed at similar levels, but in very specific cell types making up a small population of the total tissue sampled for RNA-seq.

Tissue specific expression patterns can help towards assigning physiological roles for the NtAQPs. We observed general trends between the AQP sub-families. The tobacco PIPs appeared to have more isoforms with expression in the leaf compared to the other sub-families. These are likely to be involved in roles typically reported for PIPs across plants species, including; leaf cell expansion, leaf movement, mediating water exiting the xylem, control of stomatal aperture and gas transport (e.g. CO_2) for photosynthesis (Wei et al. 2007; Hachez et al. 2008; Heinen et al. 2009). Several PIPs have targeted expression in flowers (*PIP1;7t*, *PIP1;8s*, *PIP2;2t*, *PIP2;3t*, and *PIP2;8*, *PIP2;9*, *PIP2;11, PIP2;13* sister pairs), some of which would be involved in mediating water supply during stigma, anther and petal development (Bots et al. 2005; Ma et al. 2008).

Much like the PIPs, several isoforms within the NIPs (*NIP4;3s* and *NIP4;1* and *NIP4;2* sister genes) and TIPs (*TIP5;1* sister genes) had targeted expression to the flower. The tissue-specificity of these *NtNIP*s and *NtTIP*s is consistent with the floral tissue localisation of Arabidopsis *NIP4;1*, *NIP4;2* and *TIP5;1*, which have known roles in pollen development and pollen germination (Soto et al. 2010; Di Giorgio et al. 2016). Additionally, we identified *NtTIP3;1* and *NtTIP3;2* as being exclusively expressed in the seed capsule. This is consistent with the seed-specific expression of their orthologs in other species (Utsugi et al. 2015; Li et al. 2008a; Footitt et al. 2019) where they accumulate in mature embryos and later function in water uptake during seed imbibition and germination (Mao and Sun 2015; Footitt et al. 2019). The consistent expression pattern between species implies functional conservation, meaning that *NtNIP4;1*, *NtNIP4;2* and *NtTIP5;1* likely fulfil roles in different aspects of tobacco pollen biology, and *NtTIP3;1* and *NtTIP3;2* are expected to aid tobacco seed germination.

Several PIP and TIP isoforms were found with exclusive or preferential expression in the roots (e.g. *PIP1;1*, *PIP2;4*, *PIP2;5s*, *PIP2;6*, *TIP1;2*, *TIP1;3*, *TIP2;5* and TIP2;2s), where they could be functioning in lateral root emergence (Péret et al. 2012; Reinhardt et al. 2016), regulation of cell water uptake and homeostasis (Gambetta et al. 2017), or nutrient absorption through ammonium loading in vacuoles (Lopez et al. 2004; Loqué et al. 2005). That latter possible role of

ammonium loading is especially pertinent to the two NtTIP2 proteins listed, which have a histidine residue in the ar/R LC position characteristic of ammonia transporting TIPs (Kirscht et al. 2016) .

The putative roles put forward for the various NtAQPs above, could equally apply to many of the tomato and potato AQPs given the general family-wide conservation in tissue-specific expression patterns between the three examined Solanaceae species. The generally high conservation in expression patterns between Solanaceae AQP orthologs supports the accuracy of our NtAQP orthology; assigned based on protein sequence homology. The similarity at both the protein and transcript levels strongly implies functional conservation for many of the AQP orthologs across these Solanaceae species. Knowledge of the extent of such conservation is valuable as it can help facilitate translation of findings across Solanaceae species for traits of agronomic importance. Deviations are also interesting (of which we observed several), as they indicate possible novel species-specific roles or functions.

2.6 Conclusions

We determined that the tobacco AQP family consists of 76 members divided into five subfamilies each with subtle characteristic variations in protein structures, pore lining residues, and putative post-translational regulatory mechanisms. Members of the different NtAQP subfamilies were found to localise to specific sub-cellular membranes, which contribute collectively to a dynamic and extensive transport system. Tobacco is a recent allotetraploid, which accounts for its large AQP family size and characteristic phylogenetic pairing of sister genes inherited and retained from its parents; *Nicotiana sylvestris* and *Nicotiana tomentosiformis*. Identifying the AQPs in the parental genomes allowed us to characterise the evolutionary history of the NtAQP family. Expression analysis of the NtAQPs revealed diverse tissue-specificities, consistent with the broad spanning physiological functions of AQP. Some NtAQPs were expressed widely, while other showed specialised or strong preferential expression within a single tissue. We found that the expression specificity for a number of NtAQPs resembled that of orthologous AQPs with established physiological roles in other species, allowing us

to assign putative functions in tobacco. The conservation in AQP protein structure and gene expression patterns were high with other Solanaceae species, which will facilitate the translation of knowledge from tobacco into closely related and horticulturally important crops.

2.7 Author contributions

The results presented in this chapter were generated by myself with the supervision of Prof John Evans and Dr Michael Groszmann. Mapping of RNA-seq reads for gene expression analyses of the tobacco AQPs was conducted by Dr Alex Watson-Lazowski.

2.8 Supplementary Tables

Table S2. 1. Tobacco AQP pseudo genes. *Table of sequences that encode for incomplete AQPs within the tobacco TN90 genome sequence (Sierro et al. 2014), that we have subsequently assigned as pseudo genes. Notes on trans-membrane domains were sourced from analysis using TOPCONs protein topology prediction software.*

Name	mRNA ID	Gene ID	Length (aa)	Notes
NtPIP1;2as$_{pseudo}$	mRNA_18180_cds	gene_8467	160	N- and C- terminals truncated
NtPIP1;7s$_{pseudo}$	mRNA_69157_cds	gene_32307	131	C-terminus truncation Only found in TN90 genome
NtPIP2;7s$_{pseudo}$	mRNA_133596_cds	gene_62530	214	C-terminus truncation
NtPIP2;11as$_{pseudo}$	mRNA_86603_cds	gene_40271	142	N-terminus truncation
NtNIP4;3as$_{pseudo}$	mRNA_37003_cds	gene_17291	120	C-terminus truncation Only found in TN90 genome
NtNIP4;3b$_{pseudo}$	mRNA_ 187913_cds	gene_87535	159	C-terminus truncation
NtSIP1;1s$_{pseudo}$	mRNA_113569_cds	gene_53164	156	C-terminus truncation
Undetermined	mRNA_124044_cds	gene_58110	134	No trans-membrane domains predicted
Undetermined	mRNA_185032_cds	gene_86211	137	No trans-membrane domains predicted

Table S2. 2. Extended information on the 76 tobacco aquaporins identified in this study.
Provided are protein lengths, gene identifiers, gene structures, chromosome and/or scaffold locations

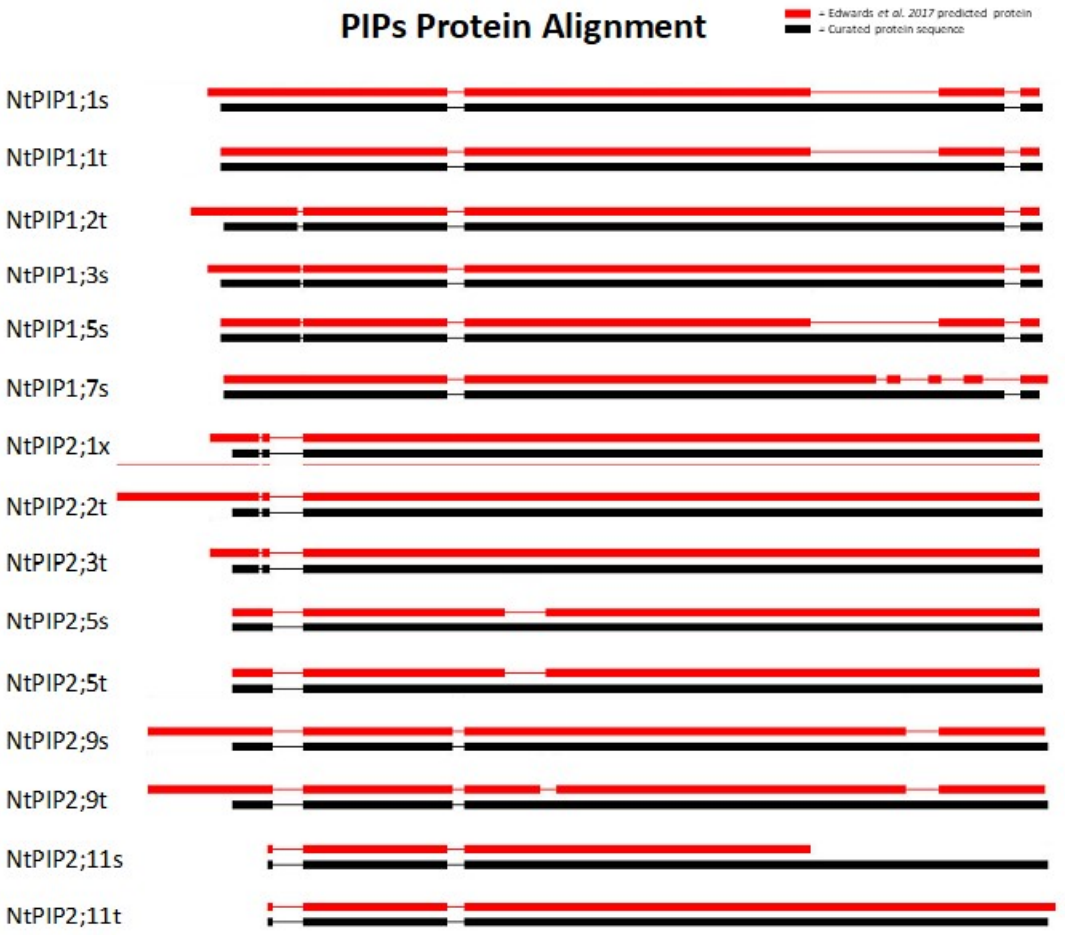
in the TN90[1] (Sierro et al. 2014) and K326[2] (Edwards et al. 2017) cultivar genomes, comparison of whether the computed gene models derived from each study matched gene structures curated in this study (Y-yes or N-no) and NCBI accessions. NtTIP2;5s, NtNIP4;2s and NtNIP4;3t genes were not identified in the K326[2] cultivar's genome. Also listed are the corresponding tomato and potato orthologs and their respective gene (inton/exon) structures.

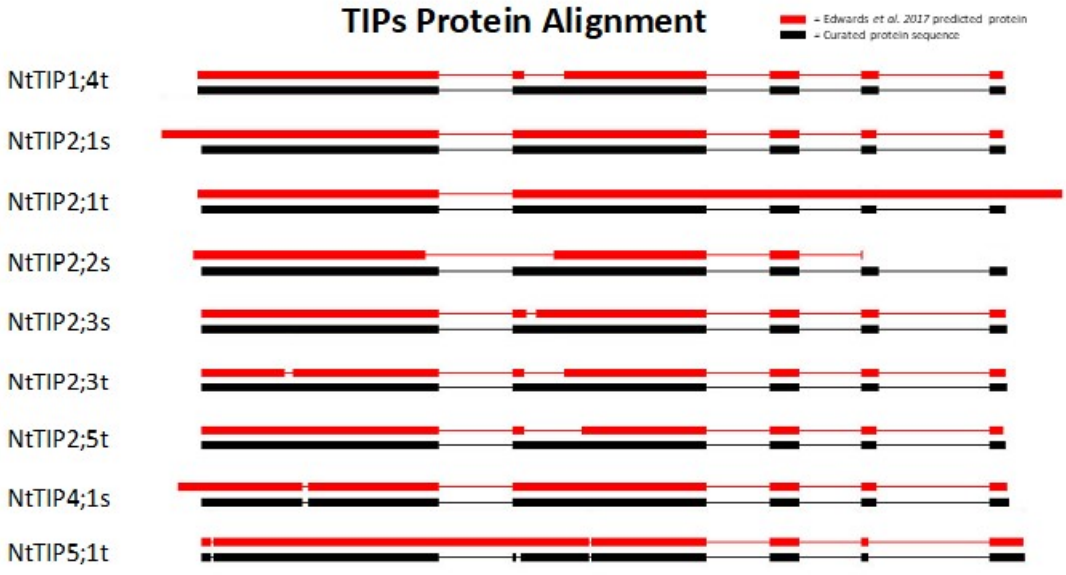
Table not included in thesis document due to size/formatting limitations.
Please refer to Table S2 in publication (De Rosa et al. 2020)

Table S2. 3 Amended annotations of previously reported tomato, potato and tobacco AQPs. *In analysing the NtAQP family, we identified misannotations in previously reported AQPs from Tomato (Solanum lycopersicum), Potato (Solanum tuberosum) and Tobacco (Nicotiana tabacum). Provided is a brief description of the error. Corrected sequences can be found in Additional file 3 (not included in Thesis).*

AQP	Original annotation ID	Reported by	mis-annotation	NCBI accession ID of corrected annotation
StXIP3;1	PGSC0003DMG400026706	Venkatesh et al., 2013	Missing 66aa from N-terminal end. The fully encoded StXIP3;1 occurs directly upstream of the reported gene.	PGSC0003DMG400026705
StXIP4;1	PGSC0003DMG400001640	Venkatesh et al., 2013	Missing 80aa from N-terminal end.	XP_006359635.1
SlXIP1;6	Solyc01g111010.2	Reusher et al., 2013	Extended N-terminus due to inclusion of upstream gene (NCBI XP_025884611.1) into the ORF of SlXIP1;6.	XP_004231447.1
SlPIP2;1	Solyc09g007770.2	Reusher et al., 2013	Truncated N- and C-termini due to SNP sequencing errors at positions 45 and 794.	NM_001302894.2
SlTIP2;2	Solyc03g120470.2	Reusher et al., 2013	Initial gene identifier incorporated a two gene fusion of a tRNase and SlTIP2;2.	NM_001287367.1 (Solyc03g120475.1.1)
NtAQP1	AF024511 & AJ001416	Biela et al., 1999	C > T SNP sequencing error at position 619 resulting in a Histidine (H) to Tyrosine (Y) substitution at amino acid position 207.	NtPIP1;5 (this study)

2.9 Supplementary figure





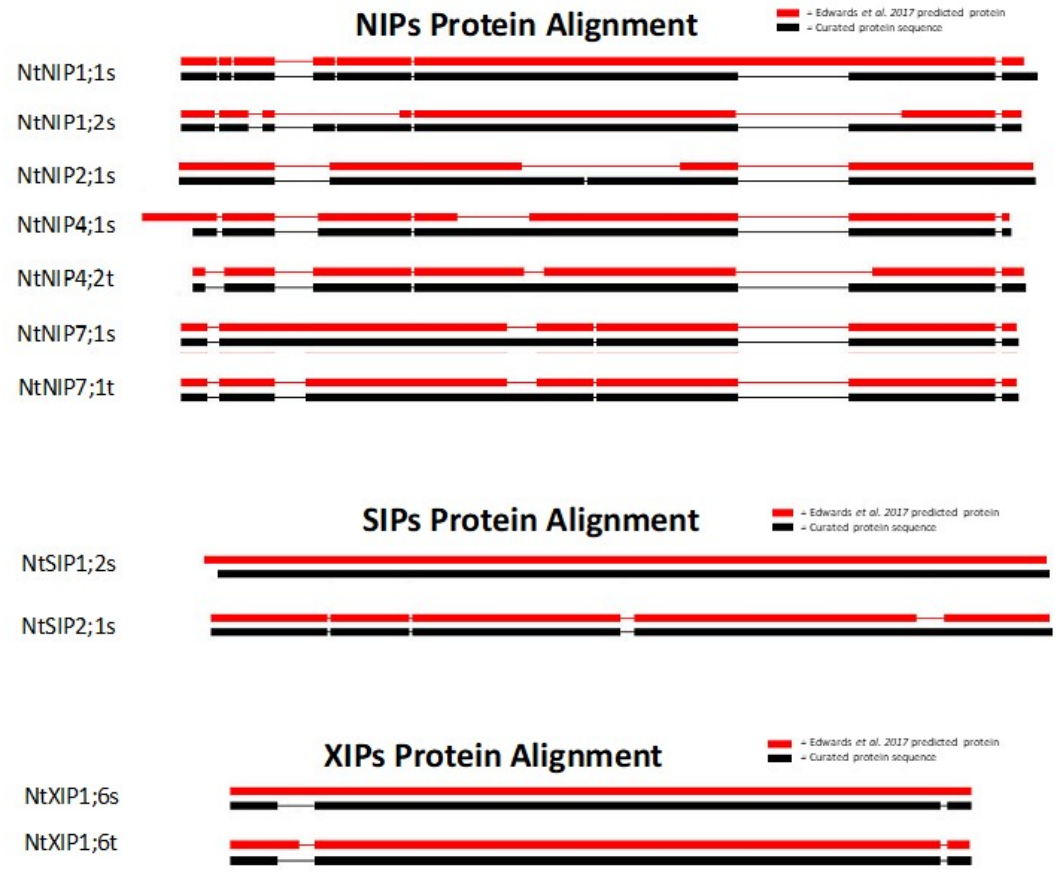


Figure S2. 1 AQP subfamily alignments for genes with incorrect protein sequences reported in Edwards et al. (2017). *In red is the Edwards et al. (2017) predicted protein sequence and in black is the curated protein sequence from this study.*

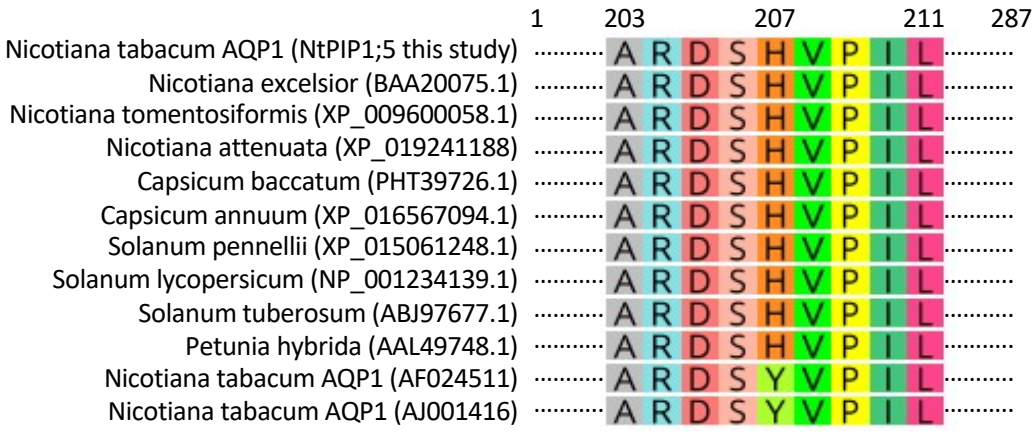
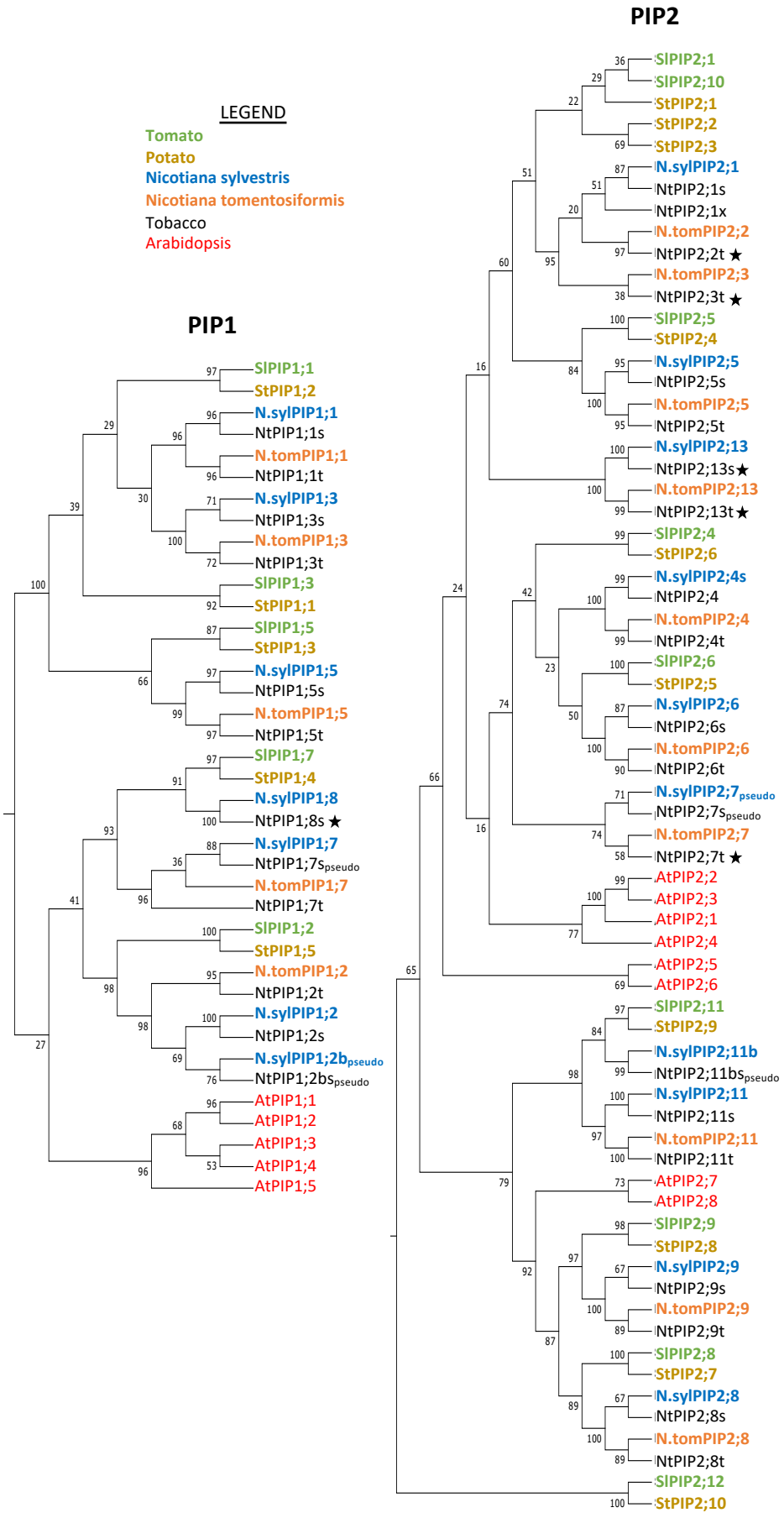
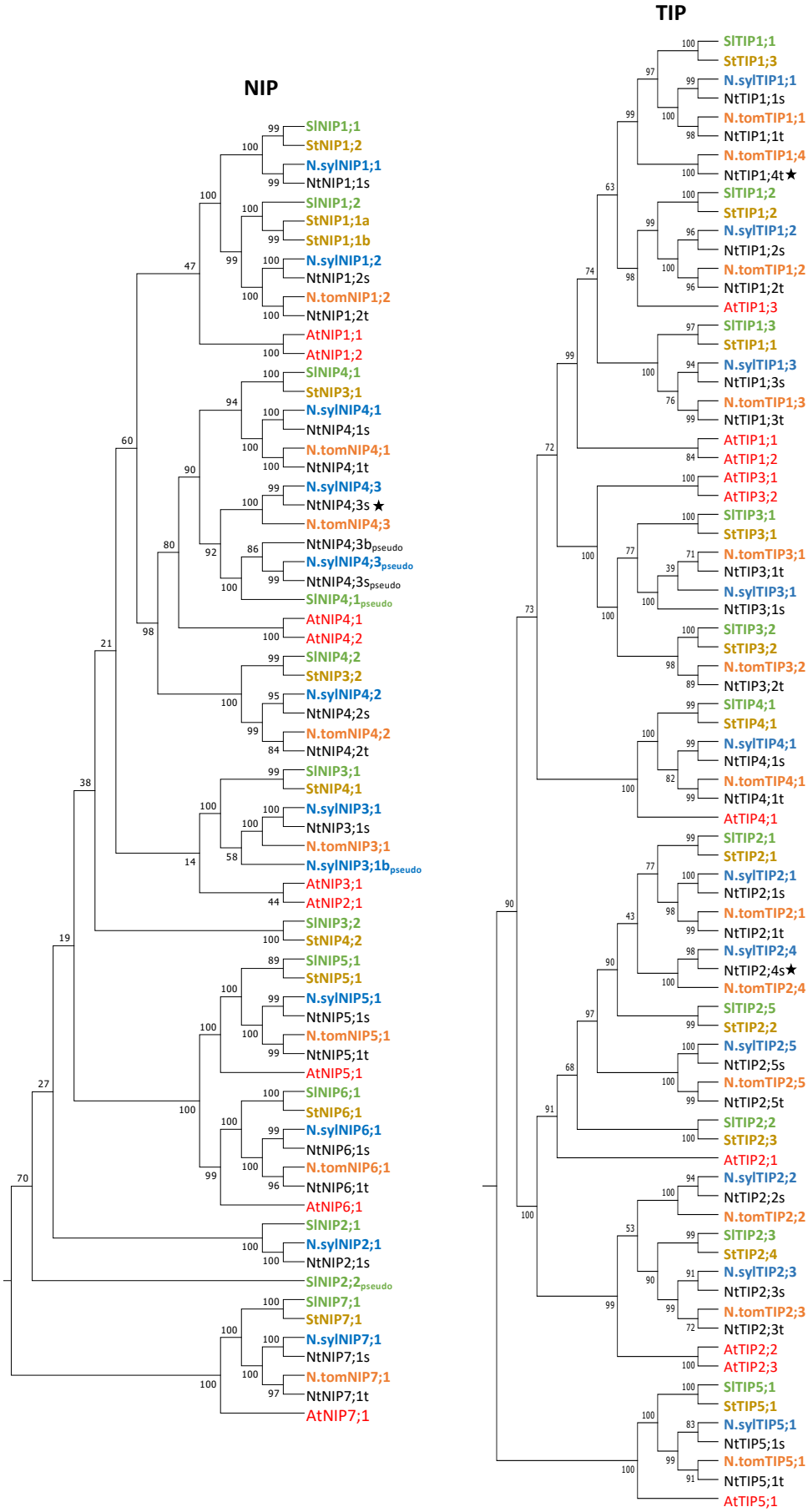


Figure S2. 2 Alignment of regions surrounding Histidine 207 in NtAQP1 (NtPIP1;5s). *Partial regions of a protein sequence alignment surrounding Histidine 207 of the NtAQP1 (NtPIP1;5) identified in this study, against the seemingly erroneous NtAQP1 sequence reported in (Biela et al., 1999; NCBI AF024511 and AJ001416) and closest BlastP matches from various other Solanaceae species.*


LEGEND
Tomato
Potato
Nicotiana sylvestris
Nicotiana tomentosiformis
Tobacco
Arabidopsis
PIP1
SlPIP1;1
StPIP1;2
N.sylPIP1;1
NtPIP1;1s
N.tomPIP1;1
NtPIP1;1t
N.sylPIP1;3
NtPIP1;3s
N.tomPIP1;3
NtPIP1;3t
SlPIP1;3
StPIP1;1
SlPIP1;5
StPIP1;3
N.sylPIP1;5
NtPIP1;5s
N.tomPIP1;5
NtPIP1;5t
SlPIP1;7
StPIP1;4
N.sylPIP1;8
NtPIP1;8s ★
N.sylPIP1;7
NtPIP1;7spseudo
N.tomPIP1;7
NtPIP1;7t
SlPIP1;2
StPIP1;5
N.tomPIP1;2
NtPIP1;2t
N.sylPIP1;2
NtPIP1;2s
N.sylPIP1;2bpseudo
NtPIP1;2bspseudo
AtPIP1;1
AtPIP1;2
AtPIP1;3
AtPIP1;4
AtPIP1;5
PIP2
SlPIP2;1
SlPIP2;10
StPIP2;1
StPIP2;2
StPIP2;3
N.sylPIP2;1
NtPIP2;1s
NtPIP2;1x
N.tomPIP2;2
NtPIP2;2t ★
N.tomPIP2;3
NtPIP2;3t ★
SlPIP2;5
StPIP2;4
N.sylPIP2;5
NtPIP2;5s
N.tomPIP2;5
NtPIP2;5t
N.sylPIP2;13
NtPIP2;13s★
N.tomPIP2;13
NtPIP2;13t★
SlPIP2;4
StPIP2;6
N.sylPIP2;4s
NtPIP2;4
N.tomPIP2;4
NtPIP2;4t
SlPIP2;6
StPIP2;5
N.sylPIP2;6
NtPIP2;6s
N.tomPIP2;6
NtPIP2;6t
N.sylPIP2;7pseudo
NtPIP2;7spseudo
N.tomPIP2;7
NtPIP2;7t ★
AtPIP2;2
AtPIP2;3
AtPIP2;1
AtPIP2;4
AtPIP2;5
AtPIP2;6
SlPIP2;11
StPIP2;9
N.sylPIP2;11b
NtPIP2;11bspseudo
N.sylPIP2;11
NtPIP2;11s
N.tomPIP2;11
NtPIP2;11t
AtPIP2;7
AtPIP2;8
SlPIP2;9
StPIP2;8
N.sylPIP2;9
NtPIP2;9s
N.tomPIP2;9
NtPIP2;9t
SlPIP2;8
StPIP2;7
N.sylPIP2;8
NtPIP2;8s
N.tomPIP2;8
NtPIP2;8t
SlPIP2;12
StPIP2;10


NIP
SlNIP1;1
StNIP1;2
N.sylNIP1;1
NtNIP1;1s
SlNIP1;2
StNIP1;1a
StNIP1;1b
N.sylNIP1;2
NtNIP1;2s
N.tomNIP1;2
NtNIP1;2t
AtNIP1;1
AtNIP1;2
SlNIP4;1
StNIP3;1
N.sylNIP4;1
NtNIP4;1s
N.tomNIP4;1
NtNIP4;1t
N.sylNIP4;3
NtNIP4;3s★
N.tomNIP4;3
NtNIP4;3bpseudo
N.sylNIP4;3pseudo
NtNIP4;3spseudo
SlNIP4;1pseudo
AtNIP4;1
AtNIP4;2
SlNIP4;2
StNIP3;2
N.sylNIP4;2
NtNIP4;2s
N.tomNIP4;2
NtNIP4;2t
SlNIP3;1
StNIP4;1
N.sylNIP3;1
NtNIP3;1s
N.tomNIP3;1
N.sylNIP3;1bpseudo
AtNIP3;1
AtNIP2;1
SlNIP3;2
StNIP4;2
SlNIP5;1
StNIP5;1
N.sylNIP5;1
NtNIP5;1s
N.tomNIP5;1
NtNIP5;1t
AtNIP5;1
SlNIP6;1
StNIP6;1
N.sylNIP6;1
NtNIP6;1s
N.tomNIP6;1
NtNIP6;1t
AtNIP6;1
SlNIP2;1
N.sylNIP2;1
NtNIP2;1s
SlNIP2;2pseudo
SlNIP7;1
StNIP7;1
N.sylNIP7;1
NtNIP7;1s
N.tomNIP7;1
NtNIP7;1t
AtNIP7;1
TIP
SlTIP1;1
StTIP1;3
N.sylTIP1;1
NtTIP1;1s
N.tomTIP1;1
NtTIP1;1t
N.tomTIP1;4
NtTIP1;4t★
SlTIP1;2
StTIP1;2
N.sylTIP1;2
NtTIP1;2s
N.tomTIP1;2
NtTIP1;2t
AtTIP1;3
SlTIP1;3
StTIP1;1
N.sylTIP1;3
NtTIP1;3s
N.tomTIP1;3
NtTIP1;3t
AtTIP1;1
AtTIP1;2
AtTIP3;1
AtTIP3;2
SlTIP3;1
StTIP3;1
N.tomTIP3;1
NtTIP3;1t
N.sylTIP3;1
NtTIP3;1s
SlTIP3;2
StTIP3;2
N.tomTIP3;2
NtTIP3;2t
SlTIP4;1
StTIP4;1
N.sylTIP4;1
NtTIP4;1s
N.tomTIP4;1
NtTIP4;1t
AtTIP4;1
SlTIP2;1
StTIP2;1
N.sylTIP2;1
NtTIP2;1s
N.tomTIP2;1
NtTIP2;1t
N.sylTIP2;4
NtTIP2;4s★
N.tomTIP2;4
SlTIP2;5
StTIP2;2
N.sylTIP2;5
NtTIP2;5s
N.tomTIP2;5
NtTIP2;5t
SlTIP2;2
StTIP2;3
AtTIP2;1
N.sylTIP2;2
NtTIP2;2s
N.tomTIP2;2
SlTIP2;3
StTIP2;4
N.sylTIP2;3
NtTIP2;3s
N.tomTIP2;3
NtTIP2;3t
AtTIP2;2
AtTIP2;3
SlTIP5;1
StTIP5;1
N.sylTIP5;1
NtTIP5;1s
N.tomTIP5;1
NtTIP5;1t
AtTIP5;1

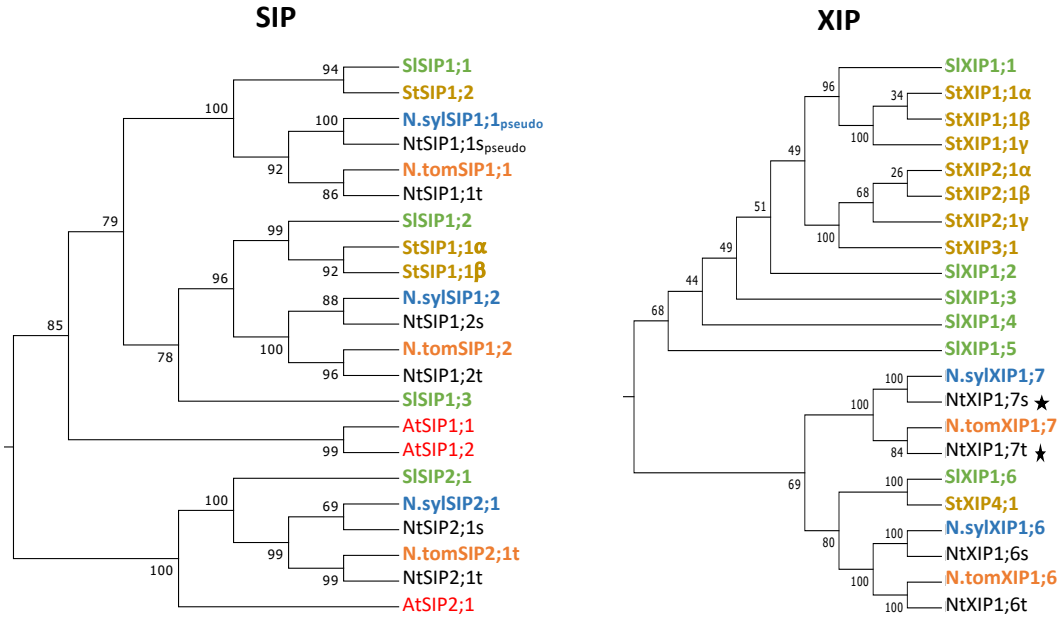


Figure S2. 3 Phylogeny of Arabidopsis and currently identified Solanaceae AQPs. *Phylogenetic trees for each AQP sub-family were generated using the neighbour-joining method from MUSCLE aligned protein sequences. Confidence levels (%) of branch points generated through bootstrapping analysis (n=1000). Solanaceae species included in this phylogeny include; N.sylvestris (orange), N.tomentosiformis (blue), tomato (green), potato (brown) and tobacco (black). Arabidopsis genes are coloured red. Black stars indicate NtAQPs which did not have a clear tomato ortholog.*

NIP C-terminal tail

```
>GmNOD26    IVRYTDKPLSETTKSASF---LKGRAASK*----------------------
>NtNIP2;1s  FIRVTDKPVHAIAPGQSFSFKLRRMKSNDEEQGV*-----------------
>NtNIP4;3s  LIRFTNKPLLQLVKSRSF---LPKLRE*------------------------
>NtNIP4;1s  LIRSTDKPLRELAKSASS---LRS*--------------------------
>NtNIP4;1t  LIRSTDKPLRELAKTASS---LRS*--------------------------
>NtNIP4;2s  LIRFTEKPLRELTKSSTF---LKSMSRSHT*---------------------
>NtNIP4;2t  LIRFTEKPLKELTKSSTF---LKSMSRSHA*---------------------
>NtNIP1;1s  IIRFTDKPLREITKSGSF---LKSKISNT-----------------------
>NtNIP1;2s  IIRFTDKPLREITKSGSF---LKSIRSSKSLRSST*----------------
>NtNIP1;2t  IIRFTDKPLREITKSGSF---LKSIRSST------------------------
>NtNIP7;1s  LLRLQGWSCKPNSTPTTT---HQHNPL*-------------------------
>NtNIP7;1t  LLRLQGWSCKPNSTPTTT---HQHNPL*-------------------------
>NtNIP5;1s  LVKLRGDDSSETPRQ------VRSFRR*-------------------------
>NtNIP5;1t  LVKLRGDDSTETPRQ------VRSFRR*-------------------------
>NtNIP6;1s  AVKLPNEDDNNHGKPSV----EHSFRR*-------------------------
>NtNIP6;1t  AVKLPNEDDNNHGKPSL----EHSFRR*-------------------------
>NtNIP3;1s  LMRLTNKSWGEAVKEISESQKVIEVSSKDKVICKCS...50...RSYKMYI*
```

PIP C-terminal tail

```
>SoPIP2;1    LRAAAIKALGSFRSNPTN*
>AtPIP2;1    LRASGSKSLGSFRSAANV*
>NtPIP1;5s   IRAIP------FHKSS*--
>NtPIP1;5t   IRAIP------FHKSS*--
>NtPIP1;3s   IRAIP------FKSKA*--
>NtPIP1;3t   IRAIP------FKSKA*--
>NtPIP1;1s   IRAIP------FKSKS*--
>NtPIP1;1t   IRAIP------FKSKS*--
>NtPIP1;7s   IRAIP------FKSK*---
>NtPIP1;7t   IRAIP------FKSK*---
>NtPIP1;2s   IRAIP------FKSGNLA*
>NtPIP1;2t   IRAIP------FRSGN*--
>NtPIP2;11s LRAQAAKTLSSFHSNPSI*
>NtPIP2;11t LRAQAAKTLSSFHSNSSI*
>NtPIP2;13s LRAGAAKALGSFRSSSQV*
>NtPIP2;13t LRAGAAKALGSFRSSSQV*
>NtPIP2;5s   LRAGAVKALGSFRSNA*--
>NtPIP2;5t   LRAGAVKALGSFRSNA*--
>NtPIP2;3t   LRAGAVKALGSFRSNA*--
>NtPIP2;2t   LRAGAVKALGSFRSNA*--
>NtPIP2;1s   LRAGAVKALGSFRSNA*--
>NtPIP2;1x   LRAGAVKALGSFRSNA*--
>NtPIP2;7t   LRAGAVKALGSFRSNA*--
>NtPIP2;6s   LRAGALKALGSFRSNA*--
>NtPIP2;6t   LRAGALKALGSFRSNA*--
>NtPIP2;4s   LRAGAIKALGSFRSNA*--
>NtPIP2;4t   LRAGAIKALGSFRSNA*--
>NtPIP2;9s   LRAGAVKALGSFRSNPTN*
>NtPIP2;9t   LRAGAVKALGSFRSNPTN*
>NtPIP2;8s   LRGSAIKALGSFRSNPTN*
>NtPIP2;8t   LRGSAIKALGSFRSNPTN*
```

Figure S2. 4. Sequence alignment of C-terminal tails of NtPIP and NtNIP proteins. *Serine residues in red are those predicted to be phosphorylated by NetPhos 3.1 (prediction score ≥ 0.8). Underlined red serine residues in GmNOD26, SoPIP2;1 and AtPIP2;1 have been experimentally confirmed as being phosphorylated in plants. Bold residues indicate the substitution of strongly conserved positively charged Lys(K)/Arg(R) residues to a His(H) residue (blue) occurring in NtPIP1;5 and NtPIP2;1 proteins.*

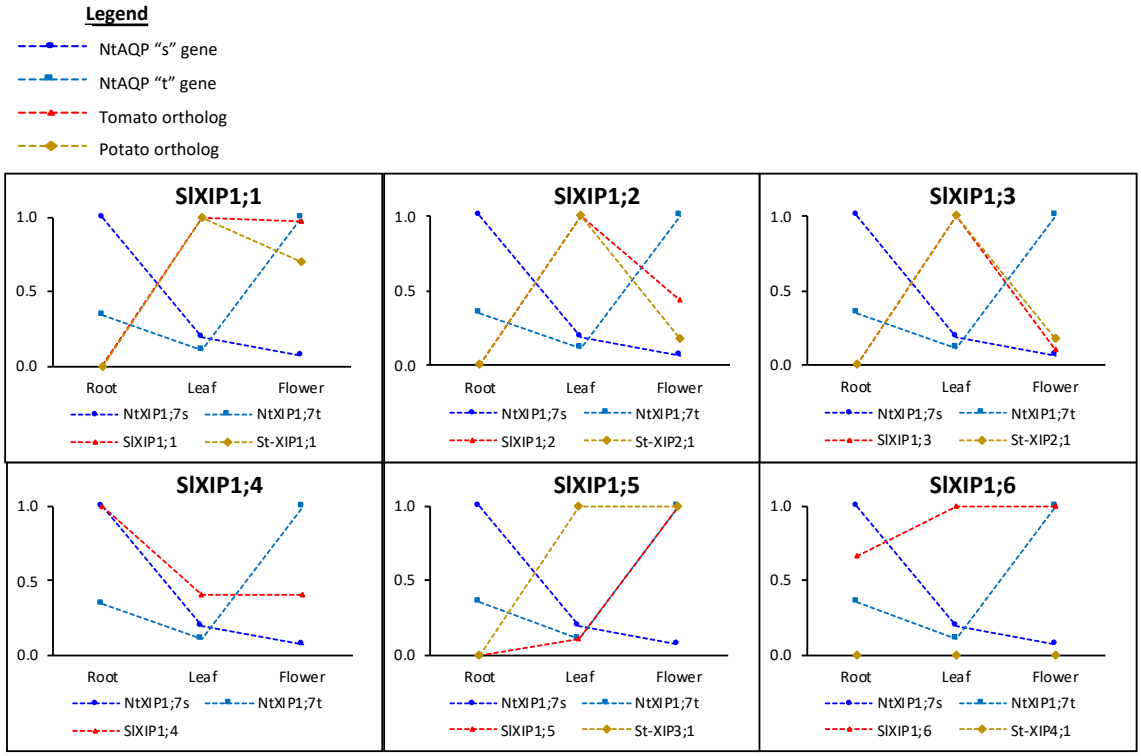


Figure S2. 5Tissue-specific expression patterns of tomato XIP isoforms (SlPXIP1;1-SlXIP1;6) and the tobacco NtXIP1;7 sister genes. *Comparison of relative gene expression in roots, leaves and flowers of tobacco NtXIP1;7 sister genes (blue) against all the tomato XIP isoforms (red, SlXIP1;1-SlXIP6), with potato orthologs (brown), trying to find a match between the XIP which are difficult to assign orthology.*

Chapter 3: Establishment of yeast-based assays for assessing AQP substrate specificities

3.1 Introduction

Replicable and accurate functional assays are essential in furthering our knowledge of AQP biology and maximising crop engineering efforts to generate higher yielding and more resilient crops. Yeast is a commonly used heterologous expression system for functional characterisation of AQPs, enabling substrate specificities to be assigned to candidate isoforms (Kaldenhoff et al. 2007). Characterisation of AQPs' permeating substrates can further our understanding of functional sites that might act as determinants for pore selectivity, and also allows us to decipher potential AQP functional roles within the plant.

We used yeast growth and survival assays and distinct yeast strains to test uptake of a variety of permeating solutes relevant to our AQP research: water, hydrogen peroxide, boric acid and urea.

Water permeability of AQPs can be tested in yeast using 'freezing and thawing' survival assays, exploiting the property of yeast cells that show increased freezing tolerance when they express functional water AQPs (Deshmukh et al. 2016). Water rapidly exits yeast cells as they freeze, then as the yeast thaws, the swift transport of water into the cell through the AQPs (integrated in the yeast plasma membrane) allows for its recovery and continued growth. The expression of water-permeable AQPs provides a survival advantage for yeast culture growth. Yeast has four AQP genes in total, two being orthodox water-permeable AQPs (AQY1 and AQY2 genes) and the other two are aquaglyceroporins (Fps1 and Ylf054c genes) (Hohmann et al. 2000). To maximise the effect of foreign AQP expression on yeast membrane water permeability, we used the *aqy1aqy2* yeast strain in our freeze-thaw survival assays which lacks its native *AQY1* and *AQY2* genes (Tanghe et al. 2002).

In order to test for hydrogen peroxide (H_2O_2) permeability, we used the *Δskn7* yeast strain, which has an increased sensitivity to H_2O_2 due to having a defective oxidative defence response (Bienert et al. 2007). Reactive Oxygen Species

(ROS), such as H_2O_2, have the ability to damage proteins, lipids and nucleic acids. Therefore, cells require scavenging mechanisms to maintain a relatively low and constant ROS concentration within the cell (Halliwell and Gutteridge 2015). *Skn7* is a transcription factor regulating expression of a number of oxidative stress response genes (Lee et al. 1999) and its mutation impairs yeast's survival upon exposure to H_2O_2. The yeast's survival is further compromised if AQPs are facilitating the diffusion (and accumulation) of H_2O_2 into the cell.

A similar assay system to H_2O_2 can be used to test if an AQP is permeable to boric acid. Boron is an essential micronutrient for plants, animals and fungi. It is a weak acid, occurring as boric acid (H_3BO_3, BA) at physiological pH. High concentrations of boric acid can negatively impact yeast growth by disrupting its cell wall synthesis and cell division (Schmidt et al. 2010). By exploiting this phenotype, we can test whether the expression of a foreign AQP impairs yeast growth when exposed to boric acid compared to the control. Impaired growth at a given boric acid concentration implies that the AQP enhances membrane diffusion of boric acid into the cell.

In order to test for urea transport, we used the *ynvwI* (*Δdur3*) yeast strain. When grown in media with urea as the sole nitrogen source, *ynvwI* (*Δdur3*) yeast has limited growth due to a deletion of the DUR3 urea transporter (Liu et al. 2003). The expression of urea-permeable AQPs in *ynvwI* (*Δdur3*) yeast would provide a growth advantage when yeast is exposed to media containing urea as the sole nitrogen source.

My PhD project aimed to functionally characterise a diverse set of tobacco AQPs across the three larger AQP sub-familes, PIP, TIP and NIP (NtPIP2;4s, NtPIP2;5t, NtPIP1;1t, NtPIP1;3t, NtPIP1;5s, NtTIP1;1s, NtTIP2;5t, NtNIP2;1s and NtNIP5;1t), using yeast as the heterologous expression system (Results presented in Chapter 4). These AQPs isoforms were selected based on the following criteria. Firstly, the PIP1 genes, NtPIP1;1t, NtPIP1;3t and NtPIP1;5s were chosen as they had more than 90% homology in gene sequences between them, with NtPIP1;5s (NtAQP1) being an established CO_2 permeable AQP isoform. Secondly, the PIP2 genes, NtPIP2;4s and NtPIP2;5t, were chosen as representative isoforms from

distinct phylogenetic sub-clades within the PIP2 phylogeny. Thirdly, NtTIP1;1s was chosen as a gene highly expressed throughout the plant (with the potential to be permeable to range of solutes, while NtTIP2;5t had high homology to AtTIP2;1, an established ammonium transporter (Loqué et al. 2005). Fourthly, within the NIP subfamily, NtNIP2;1s was selected due to its homology to rice NIP2;1, a Class III NIP thought to be a silicon transporter (Zhao et al. 2010) and NtNIP5;1s was chosen due to it's homology to AtNIP5;1, as Class II NIP characterised as boron-permeable (Takano et al. 2006). In addition to characterising AQPs permeable to a range of substrates (water, H_2O_2, boric acid and urea), comparisons between highly homologous isoforms could identify amino acid variations that might confer specificity to certain substrates. In order to test for water, H_2O_2 , boric acid and urea substrate permeabilities, we established and optimised growth based yeast functional assays. I was responsible for the establishment of the H_2O_2 and urea growth based yeast assays presented in this Chapter. The resulting assays provide a high throughput and reliable method that will benefit future functional characterisation of AQPs in yeast-based systems.

3.2 Materials and Methods

3.3.1 Yeast strains, genes and yeast expression vectors used

Three distinct yeast strains were used for the development of yeast-based experiments to ascertain aquaporin substrate specificities: *aqy1aqy2*, *skn7* and *ynvwI* (*Δdur3*) (details listed in Table 3.1).

Table 3.1 List of yeast strains used for aquaporin functional assays.

Yeast strain	Strain-specific mutations	Functional assay
aqy1aqy2	AQY1 and AQY2 aquaporin genes (Tanghe et al. 2002)	Freeze-thaw assay; testing for water permeability Boric Acid permeability assay
skn7	Skn7 transcription factor, required for ROS scavenging (Bienert et al. 2007)	H_2O_2 permeability assay
ynvwI (*Δdur3*)	DUR3 urea transporter (Liu et al. 2003)	Urea growth-based assay

Nine tobacco aquaporin genes were chosen for functional characterisation through yeast-based assays, and their accession numbers are listed in Table 3.2. Also listed are Arabidopsis aquaporins *PIP2;3* and *TIP2;3* that have already been characterised and shown to permeate H_2O_2 and urea, respectively (Hooijmaijers et al. 2012; Dynowski et al. 2008). The gene coding sequences were commercially synthesised and obtained in gateway-enabled entry vectors.Three Gateway destination vectors were used for yeast functional assays (illustrated in Figure 3.1): pRS423-GPD (Alberti et al. 2007) referred to as MG0515, pRS426-GPD-ccdB-ECFP (Alberti et al. 2007) referred to as MG0527 and pRS426-GPD-EGFP-ccdB vector (Alberti et al. 2007) referred to as MG0522. The MG0515 vector contains Histidine 3 (HIS3) marker gene required for selection in the *aqy1aqy2* and *skn7* yeast strains, whereas the MG0527 and MG0522 destination vectors contain the uracil 3 (URA3) gene required for selection in the *ynvwI* (*Δdur3*) yeast strain.

Table 3.25 Table of genes selected for functional characterisation studies in yeast.

Gene Name	Gene identifier
NtPIP1;1t	BK011393
NtPIP1;3t	BK011397
NtPIP1;5s	BK011398
NtPIP2;4s	BK011406
NtPIP2;5t	BK011409
NtTIP1;1s	BK011426
NtTIP2;5t	BK011440
NtNIP2;1s	BK011379
NtNIP5;1t	BK011387
AtPIP2;3 (H_2O_2 positive control)	At2g37180
AtTIP2;3 (urea positive control)	At5g47450

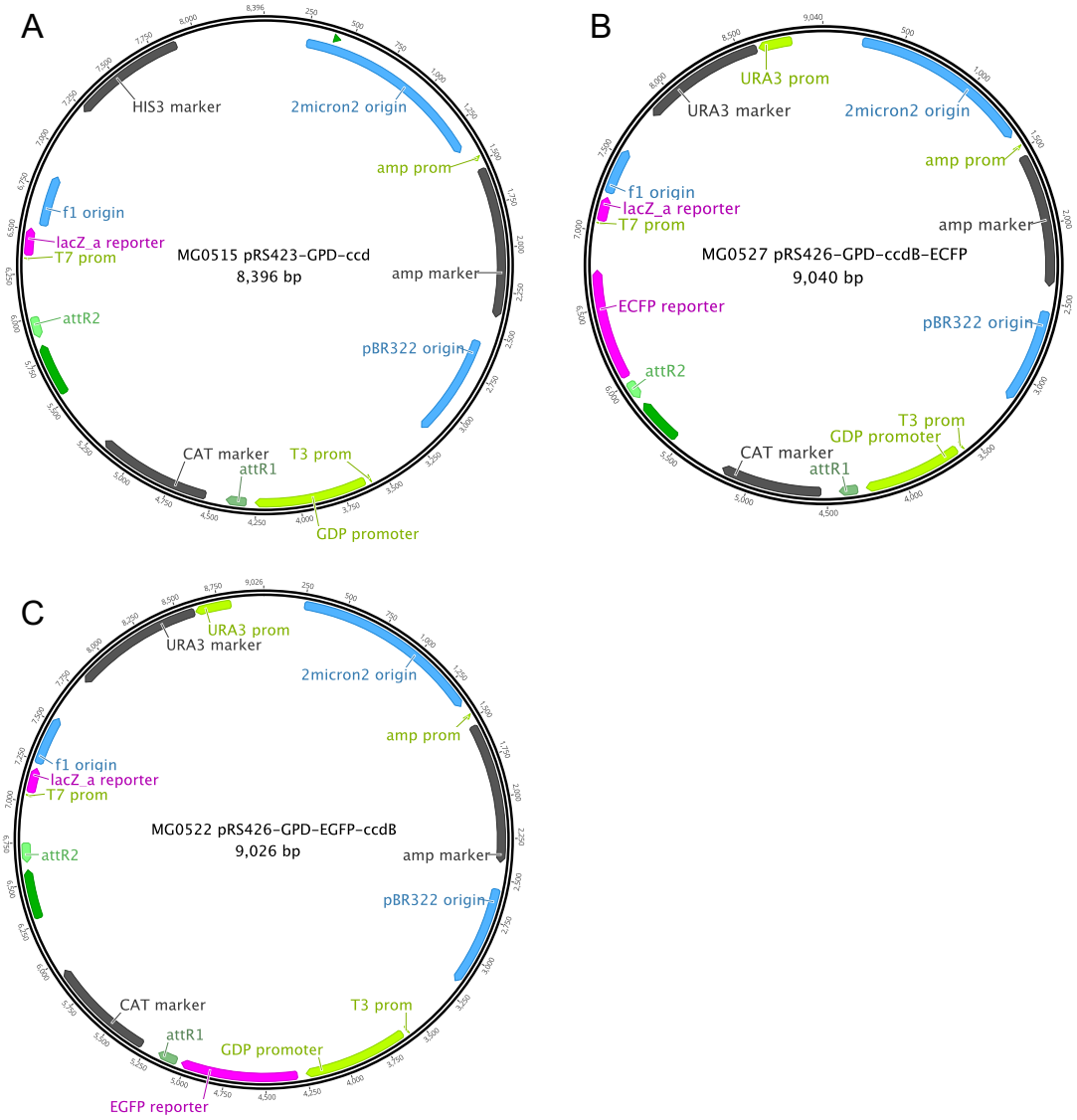


Figure 3. 1 Images of Expression vectors used for yeast functional characterisation studies. A. *MG0515 pRS423-GPD expression vector, containing histidine 3 (HIS3) marker gene for media selection and GPD promoter.* ***B.*** *MG0527 pRS426-GPD-ccdB-ECFP vector, containing uracil 3 (URA3) marker gene for media selection, GPD promoter, and a cyan fluorescent protein (CFP) reporter gene not translationally fused to gene insert.* ***C.*** *MG0522 pRS426-GPD-ECFP-ccdB vector, containing uracil 3 (URA3)) marker gene for media selection, GPD promoter, and a N-terminal Green fluorescent protein (GFP) reporter gene translational fusion to gene insert.*

3.3.2 Preparation of Yeast extract Peptone Dextrose (YPD) media

10g of Bacto Yeast extract, 20g of Bacto Peptone and 20g of Dextrose were dissolved in 1L of distilled water, the media was then autoclaved.

3.3.3 Yeast Nitrogen Base (YNB) agar and liquid media

Prior to the preparation of the YNB media (Table 3.3 and Table 3.4), the following solutions were prepared: 20x sterile Glucose (200g in 500mL water, filter sterilised); 10x Yeast Drop Out Solution (6.95g dissolved in 500mL water, filter sterilised); 100x Histidine solution (100mg of L-Histidine HCL monohydrate dissolved in 50mL water, filter sterilised); 100x Tryptophan solution (100mg of L-tryptophan dissolved in 50mL water, filter sterilised); 100x Leucine solution (500mg of L-Leucine dissolved in 50mL water and filter sterilised); and 100x Uracil solution (100mg of L-uracil dissolved in 50mL water and filter sterilised). Liquid 2x Yeast Nitrogen Base without amino acids (6.7g Yeast nitrogen base without amino acids in 500mL water, autoclaved and stored at room temperature) was used to make up liquid YNB media, see Table 3.4.

Table 3. 3 YNB agar media recipe for yeast functional characterisation experiments. *Listed are recipes for YNB agar media lacking either Histidine (-HIS) or Uracil (-URA) required for growth of yeast expressing the HIS3 (present in MG0515 vector) and URA3 (present in MG0527 and MG0522 vectors) marker genes respectively.*

	YNB (-HIS)	**YNB (-URA)**
2x Yeast nitrogen base without amino acids	3.35g	3.35g
Agar	7g	7g
Water	250mL	250mL
	adjust pH to 5.8 with 1M KOH	
	autoclave	
10x Yeast Drop Out solution (-HIS -TRP -LEU -URA)	35mL	35mL
x100 HIS	-	3.5mL
x100 LEU	3.5mL	3.5mL
x100 TRP	3.5mL	3.5mL
x100 URA	3.5mL	-
Water	37mL	37mL
20x Sterile Glucose (40% w/v)	17.5mL	17.5mL
	Pour media in plates and set for 90 mins.	

Table 3.4 YNB liquid media recipe for yeast functional characterisation experiments. *Listed are recipes for YNB liquid media lacking either Histidine (-HIS) or Uracil (-URA) required for growth of yeast expressing the HIS3 (present in MG0515 vector) and URA3 (present in MG0527 and MG0522 vectors) marker genes respectively.*

	YNB liquid media (-HIS)	YNB liquid media (-URA)
2x Yeast nitrogen base without amino acids (liquid media)	50mL	50mL
10x Yeast Drop Out solution (-HIS -TRP -LEU -URA)	10mL	10mL
x100 HIS	-	1mL
x100 LEU	1mL	1mL
x100 TRP	1mL	1 mL
x100 URA	1mL	-
20x Sterile Glucose	5mL	5mL
Sterile Water	32mL	32mL
Total	100mL	100mL

3.3.4 Preparation of yeast competent cells

The Zymo "Frozen-EZ yeast transformation Kit II" was used for yeast competent cell preparation, with slight modifications to the protocol. Yeast glycerol stocks for *aqy1aqy2*, *skn7* and *ynvw1* yeast strains were streaked on freshly made YPD plates and incubated for 1 day at 30°C. For each strain, 2-3 Colonies were inoculated from plates, and grown in 5mL of YPD media at 30°C, 250rpm until an OD_{600} of 2 was reached (overnight). Three fresh cultures were set up for each strain (1 for OD testing and 2 'working cultures') by making up 10mL cultures and adding overnight culture from each strain to a starting OD_{600} of 0.3. The 10mL cultures were shaken at 250rpm and 30°C for 3.5-4 hours (until OD_{600} of 0.8-1.0). Once desired OD was reached, two 'working cultures' were merged (20mL total) and pelleted at 500g for 4 min. Supernatant was discarded and 10mL of EZ1 solution was added. Cells were pelleted at 500xg for 4 mins and 1mL of EZ 2 solution was added. 25µL aliquots of suspended cells were distributed to 1.5mL tubes and stored in a -80°C freezer.

3.3.5 Yeast transformation with AQP construct

The "Frozen-EZ yeast transformation Kit II" was used for yeast transformation as per manufacturer's instructions. Competent cells (prepared as per Section 3.3.4) were thawed on ice and 0.2-1µg DNA was added (maximum of

5µl). 250µL of EZ3 solution was added and mixed gently but thoroughly using the pipette. Yeast cell mixture was then incubated at 30°C for 3 hours; with vigorous mixing 2-3 times during incubation period (inverting tubes and flicking). 150µL of transformation mixture was spread on appropriate YNB media selection plates and incubated for 3-4 days at 30°C (until colonies appeared). Each transformed yeast culture was then grown for 24 hours at 30°C with shaking at 250rpm. 10 x 10µL spots of each expression construct were distributed on YNB media with either -HIS or -URA selection for either the MG0515 or for MG0527 & MG0522 selection respectively. Spotted plated were incubated at 30°C for 2 days and then placed in the fridge (4°C). Spotted plates were used for the starting cultures of functional assays (described below), using roughly ½ of a full spot as starting material for each culture.

3.3.6 H_2O_2 Yeast spotting assay

In order to prepare media plates containing various H_2O_2 treatments, a 100mM H_2O_2 stock solution was freshly made and filter sterilised. 100µL, 300µL, 500µL and 750µL of 100mM H_2O_2 stock solution were applied to a total of 100mL YNB (-HIS) media to make 0.1mM, 0.3mM, 0.5mM and 0.75mM final H_2O_2 concentrations respectively. A 0mM H_2O_2 treatment was used as an "Untreated" control. Yeast expressing the tobacco AQPs cloned in MG0515 vector and the MG0515 empty vector control were grown overnight to OD_{600} 1.5-2. OD_{600} was then standardised to 1 and each culture diluted to 1/10, 1/100, 1/1000 and 1/10000 dilutions. Yeast cultures were spotted out on plates containing various H_2O_2 treatments, incubated at 30°C and grown for 7 days. Plates were scanned for image collection.

3.3.7 Using the SpectroStar microplate reader for yeast functional assays

The SpectroStar Nano absorbance microplate reader (BMG Labtech) was used to monitor the growth of yeast cultures. Yeast culture volumes of 190µL were aliquoted in 96-well microculture plates and treatments (10µL) were added accordingly. Shaking of the plate was set to a moderate speed (400rpm) to ensure

an even distribution of the yeast suspension within the well and their eventual even settling to the bottom of the well. Yeast growth was measured through OD_{650} absorbance readings collected by the instrument at 10 minute intervals. Multi-point readings across the entire well were collected and averaged to give a single representative value. This ensured a more representative OD value, as opposed to a measurement obtained via a single beam location of the well.

OD readings from yeast microcultures were processed in order to standardise growth measurements and compare growth of yeast expressing various constructs, with each exposed to different treatments. The blank media (media not containing yeast cells) OD_{650} absorbance readings were subtracted from the individual OD_{650} readings for each yeast culture. There was a short period of rapid increase in OD_{650} values accompanying the settling of the yeast to the bottom of the wells, which was followed by the characteristic flat consistent reading associated with the 'lag phase'. Subsequently, the OD_{650} readings began to increase again, indicating yeast growth (increasing OD resulting from increased cell number).

The output readings were transformed to compensate for the non-linearity in increasing cell density and OD_{650} reading. The correction formula was derived by plotting OD_{650} reading of known concentrations of yeast cells through a dilution series and fitting these with a polynomial equation. Small fluctuations in the reading between time intervals were smoothed by fitting sliding splines. The final readings were then converted to the natural log of OD/initial OD ($Ln(OD/OD_i)$).

After the growth curve data was processed, we could calculate a biologically meaningful and consistent measuring time point for data collection of yeast growth phenotypes across various treatments. From $Ln(OD/OD_i)$ curves we calculated the slope (growth rate), identifying the region where the rate of growth was the highest (Max growth rate). The max exponential growth period was defined as the time range at which growth rate was within a 10% of the Max growth rate. A linear line of best fit was applied to this max growth period, with the x-axis intersect defining the point at which the lag-phase ended and the accelerated growth period began. Stationary phase was defined as having commenced for the 'Untreated' yeast cultures at the time point where growth rate slowed to below 5% of the Max growth rate. This time point was selected as the measuring time point. The Area Under the

Curve (AUC) was calculated from the commencement of the assay up to the measuring time point of the 'Untreated' yeast line for each construct.

3.3.8 H_2O_2 Yeast micro-culture toxicity assay

Yeast (*skn7* strain) cultures expressing MG0515 empty vector (negative control) or MG0515-NtAPQs (listed in Table 3.2) were grown for 28 hours in 1.25mL YNB(-HIS) liquid media at 30°C with shaking at 250rpm. Culture OD was measured (expected OD_{650} readings between 1-2) and two 2mL sub-cultures (two biological replicates per construct) were made having a cell concentration of $0.6x10^7$ cells/ml. Diluted cultures were incubated at 30°C whist H_2O_2 solutions were prepared. 1M H_2O_2 was prepared by adding 5.1mL of 9.97M H_2O_2 to 44.9mL sterile water. Ten 10mL tubes were labelled (1-10) and were prepared as listed in Table 3.5.

10µL of H_2O_2 treatments (Tubes 1-10) and 190µL of yeast cultures were distributed to their respective wells in the 96-well plate; yeast growth was monitored using the SpectroStar nano microplate reader. The program consisted of incubation at 30°C, with measurement of OD_{650} at the end of each 10 minute kinetic cycle (running for 24 hours). Each experimental batch consisted of 2 biological replicates; 3 experimental batches were conducted in total (resulting in 6 replicates per constructs).

Table 3. 5 Preparation of H_2O_2 stock solutions for H_2O_2 toxicity assay.

Tube number	Stock Concentration	Final assay concentration **(10uL Stock in 200µL well)**	Preparation instructions
1	160mM	8mM	Add 1.6mL of **1M H_2O_2 stock** to 8.4mL sterile water
2	80mM	4mM	Add 5mL from **Tube 1** to 5mL sterile water
3	40mM	2mM	Add 5mL from **Tube 2** to 5mL sterile water
4	20mM	1mM	Add 5mL from **Tube 3** to 5mL sterile water
5	10mM	0.5mM	Add 5mL from **Tube 4** to 5mL sterile water
6	5mM	0.25mM	Add 5mL from **Tube 5** to 5mL sterile water
7	4mM	0.2mM	Add 4.8mL from **Tube 6** to 1.2mL sterile water

8	2mM	0.1mM	Add 3mL from **Tube 7** to 3mL sterile water
9	1mM	0.05mM	Add 3mL from **Tube 8** to 3mL sterile water
10	-	Untreated/Water	6mL sterile water

3.3.9 Urea yeast micro-culture growth based assay

Half yeast spots of *ynvwI* yeast strain expressing MG0527 empty vector (negative control) or MG0527-NtAPQs (listed in Table 3.2) were resuspended in 1.25mL of YB + Glucose media (culture medium without nitrogen source; see Table 3.6). Culture OD was measured and two 2mL cultures (two biological replicates per construct) were made using YB +Glucose media to have a total cell number of $1.2x10^7$ cells. 500mM urea stock was made up (300mg urea dissolved in 10mL sterile water; filter sterilised) to prepare the 400mM, 320mM, 240mM, 160mM and 80mM stock solutions (required for 20mM, 16mM, 12mM, 8mM and 4mM urea treatments respectively; see Table 3.7). 10µL of urea stock solutions and 190µL of diluted yeast cultures were distributed to their respective wells in the 96-well plate; yeast growth was monitored using the SpectroStar nano microplate reader. The program consisted of incubation at 30°C, with measurement of OD_{650} at the end of each 20 minute kinetic cycle (running for 48 hours).

Table 3. 6 Preparation basic YB + Glucose media.

	Basic YB + Glucose media
2x Yeast nitrogen base without amino acids (liquid media; see Section 3.3.3)	25 mL
20x Sterile Glucose	4.25 mL
Sterile Water	21.75 mL
Total	50mL

Table 3. 7 Preparation of urea stock solutions for urea growth-based functional assays.

Tube number	Urea stock Concentration	Final assay concentration **(10µL of stock in 200µL well)**	Preparation instructions
1	400mM	20mM	Add 800µL of **500mM urea** stock to 920µL sterile water
2	320mM	16mM	Add 640µL of **500mM urea** stock to 360µL sterile water
3	240mM	12mM	Add 480µL of **500mM urea** stock to 520µL sterile water
4	160mM	8mM	Add 320µL of **500mM urea** stock to 680µL sterile water
5	80mM	4mM	Add 160µL of **500mM urea** stock to 840µL sterile water
6	Untreated	Untreated/Water	1mL sterile water

3.3 Results

3.3.1 Development of H_2O_2 toxicity assay: yeast spotting experiment

As a first attempt to establish a fast and reliable screen for assessing AQP H_2O_2 permeability, we used the *Δskn7* yeast strain (hypersensitive to ROS) in a spotting assay on H_2O_2 containing agar medium. Yeast expressing tobacco AQPs (NtPIP1;1t, NtPIP1;3t, NtPIP1;5s, NtPIP2;4s, NtPIP2;5t, NtTIP1;1s, NtTIP2;5t, NtNIP2;1s and NtNIP5;1t) and MG0515 empty vector control were spotted out on media plates containing 5 different H_2O_2 concentrations (0mM, 0.1mM, 0.3mM, 0.5mM and 0.75mM), and grown for 7 days at 30°C (Figure 3.2).

We were interested in determining whether any NtAQP-expressing yeast showed a reduced rate of growth compared to the empty vector control when grown on media containing H_2O_2. A hypersensitive growth phenotype suggests that expression of that AQP increases membrane permeability H_2O_2 resulting in a toxicity response and therefore the slower growth.

The toxicity of H_2O_2 in these assays is proportional to the absolute number of H_2O_2 molecules, the number of yeast cells, and the permeability of the yeast membrane. Therefore, spotting different dilutions of yeast (1/10, 1/100, 1/1000 and 1/10000 dilutions of OD_{600} 1) provides a semi-quantitative means of determining differential sensitivity to H_2O_2 for the different AQP-expressing yeast.

Figure 3.2 shows representative growth trends of the diverse set of NtAQPs tested for the H_2O_2 yeast spot assays (showing results for NtPIP2;4s, NtPIP2;5t, NtNIP2;1s and NtNIP5;1t). It was immediately apparent that yeast expressing different NtAQPs had different rates of growth compared to the empty vector control on 0mM H_2O_2 (Untreated yeast). Beyond the growth inconsistencies at 0mM H_2O_2, some differential yeast growth was also seen in response to the different H_2O_2 treatments. For example, NtNIP5;1 and NtPIP1;5s, appeared to be less sensitive to H_2O_2 exposure, exhibiting some growth even at the highest level tested (0.75mM H_2O_2), while the empty vector control had minimal growth at 0.5mM and no growth at 0.75mM H_2O_2. Other NtAQPs, such as NtNIP2;1, appeared to have slightly less growth than the empty vector control on all H_2O_2 media treatments, having no growth at 0.75mM H_2O_2. NtPIP2;4s and NtPIP2;5t were particularly remarkable, in that they appeared to have the least growth across all treatments, having minimal growth at 0mM H_2O_2, with no growth at all at 0.5mM and 0.75mM H_2O_2 (Figure 3.2). Although some constructs appeared to show reduced growth compared to empty vector at higher H_2O_2 concentrations, it was difficult to attribute this to variation in H_2O_2 sensitivity due to AQP substrate specificity, because of the inconsistencies in growth on 'Untreated' media.

The inconsistent growth of yeast cultures at 0mM H_2O_2 (despite multiple attempts) suggests that there might be inherent problems in the robustness of this methodology. Additionally, other limitations could impact our ability compare yeast sensitivity to H_2O_2 exposure, such as differences in growth rates of a yeast expressing a given AQP, lack of knowledge of yeast culture growth stage at measurement collection and potential for variability in yeast cell exposure to media containing the H_2O_2 treatments. Collectively, such limitations in gathering growth response data at a single time point on agar media, sparked our interest in monitoring yeast growth at frequent intervals in liquid microcultures.

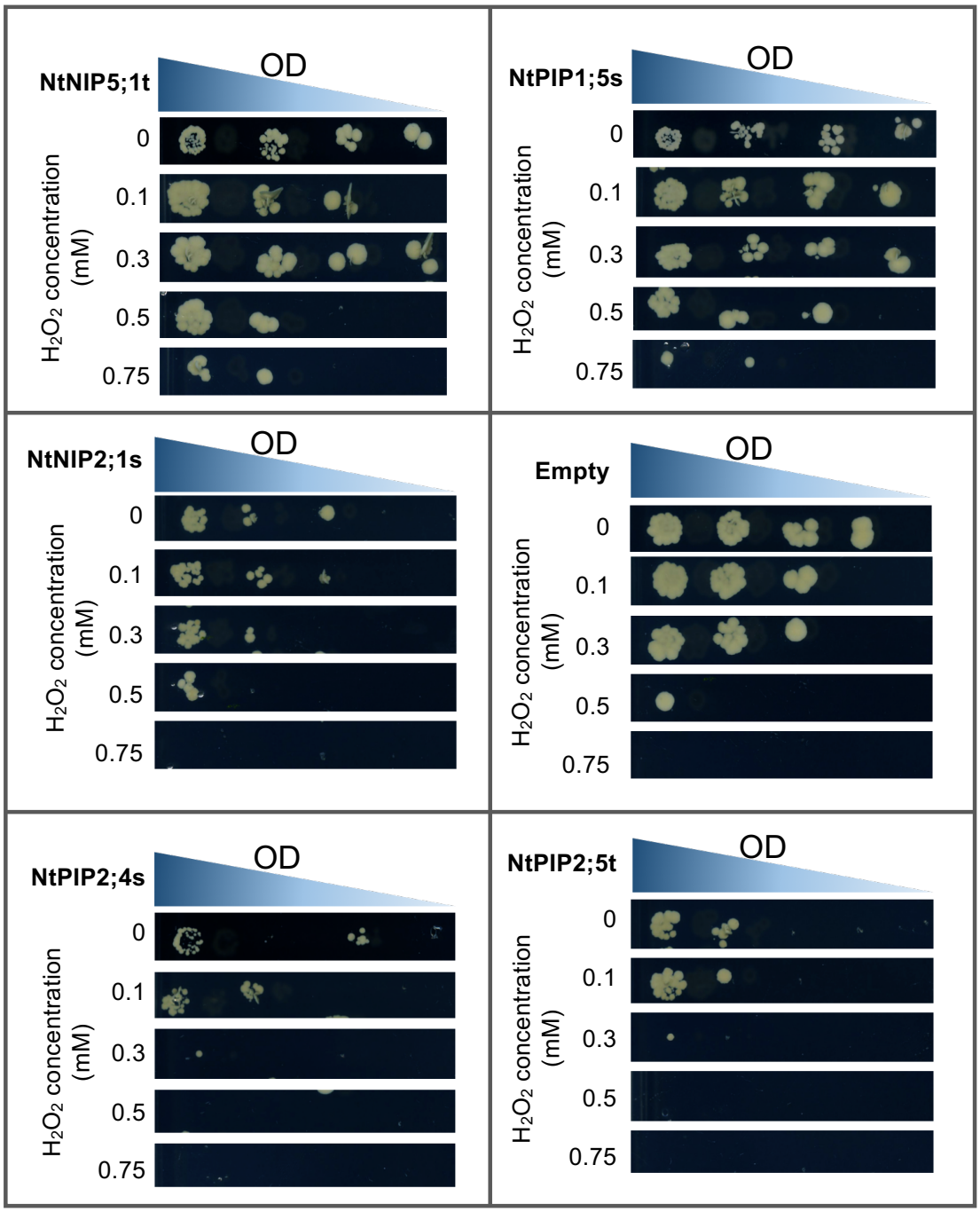


Figure 3. 2 Yeast spotting assay on media with increasing H_2O_2 concentrations. *Each panel shows yeast growth on plates containing, 0mM, 0.1mM, 0.3mM, 0.5mM and 0.75mM H_2O_3. Spotted yeast has decreasing cell numbers from left to right, starting at OD_{600} of 1, then 1/10, 1/100, 1/1000 and 1/10000 dilutions.*

3.3.2 Influence of recording time on interpretation of yeast assay results

The spotting assay results showed variation in yeast growth of NtAQP-expressing cultures (on untreated media) and highlighted difficulties in detecting sensitivity of the *Δskn7* yeast to H_2O_2 exposure. To overcome these problems and increase the monitoring resolution of the yeast growth, we changed to a yeast microculture based assay using a plate reader shaking incubator to measure OD at 10min intervals throughout the duration of the experiment.

Figure 3.3 shows theoretical growth curves of two yeast cultures over time, one is an 'Untreated' yeast culture, and the other has been treated with a yeast stress (e.g. dose of H_2O_2; 'Treated yeast'), appearing to have an growth delay/impairment. Three time points are selected as examples to illustrate the potential limitations and errors that can arise when only a single non-standardised time point is used to interpret yeast assay results. If 'Time 1' was used to infer the effect of the treatments, (e.g. by looking at a yeast spot size difference or measuring a single OD culture reading), we would be sampling when the untreated yeast is in exponential growth, but the treated yeast was still in lag phase making it appear as if the treatment lethal to the AQP expressing yeast. At 'Time 2', the untreated is no longer in exponential growth, the curve is starting to plateau and its growth measurement will not change much over the remainder of the experiment. By contrast, the treated yeast is now in exponential growth, having grown significantly since 'Time 1' (Figure 3.3). If we were to use 'Time 2' to interpret the yeast culture growth difference, we could infer that the treatment has led to a slight decrease of yeast culture growth. If we were to leave the yeast to grow until 'Time 3' and assess the treatment response then, it would appear as if there was no difference in growth between the untreated and treated yeast, and from this information only, we would not be aware of any difference in growth between treatments.

The interpretation of growth phenotypes differed significantly at each time point, highlighting the value of monitoring yeast growth by repeatedly recording growth at short time intervals to improve detection of yeast culture responses to specific treatments and the yeast responses to generally expressing a foreign protein.

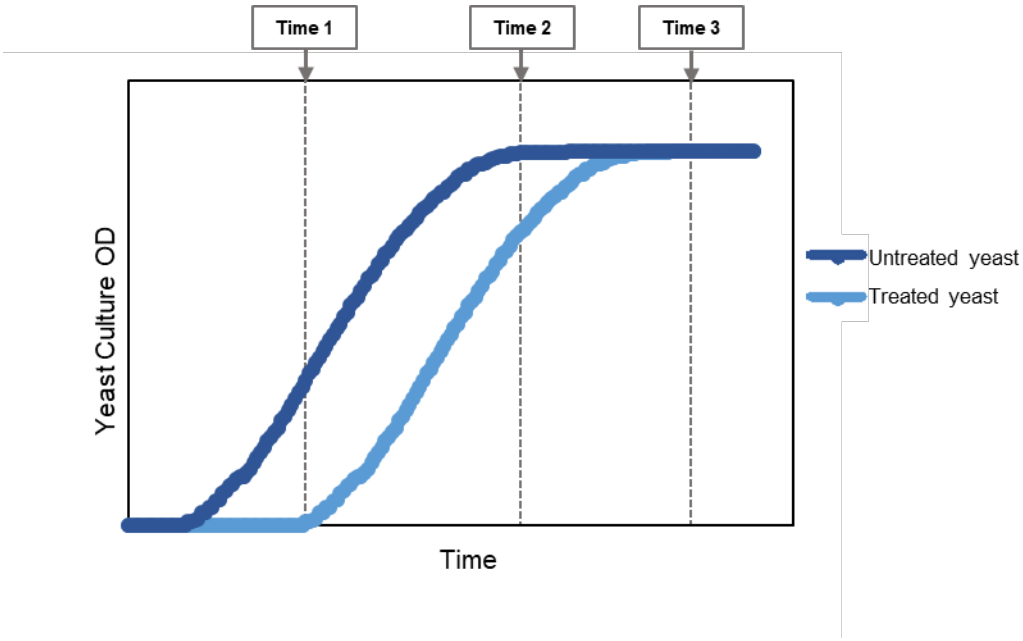


Figure 3. 3 Theoretical comparison of yeast growth curves of a treated and untreated yeast culture and the varying interpretations which could be derived based on the measuring time. *Yeast culture Optical Density (OD) measurements over time of an 'untreated' (dark blue) and 'treated' (light blue) yeast culture, with 'treated' yeast having impaired growth over time. Time 1, Time 2 and Time 3, were marked to highlight potential differences in interpretation which can arise when a single time point is used to infer treatment effects.*

3.3.3 Characteristics of a yeast growth curve

We can more accurately compare growth characteristics between different AQP expressing yeast, by analysing full growth curves. At its optimum, yeast growth is exponential in nature and so the collection of OD reading over time can be plotted as $Ln(OD/OD_{initial})$ vs. time. The $OD/OD_{initial}$ function allows for slight variations in starting OD to be normalised. There are different phases of a yeast culture growth curve, which are functions of the yeast's ability to replicate, response to the medium, and tolerance of waste by-product accumulation (Hall et al. 2013). These are known as the Lag, Log (or exponential), Deceleration and Stationary (curve plateau) phases.

During the 'Lag phase', there is little to no detectable growth, and the yeast transcriptome and metabolome acclimate to the growth environment (i.e. medium,

temperature, etc.). Lag phase ends when yeast growth accelerates and enters the 'Log phase'. During the 'Log phase', yeast cells multiply rapidly leading to exponential growth; nutrients within the media are in excess relative to the cell number and toxicity effects of waste accumulation are negligible due to dilution effects. The 'Log phase' is used to determine the maximum growth potential of the yeast as either maximum growth rate (OD time^{-1}) or the doubling time (inversely related to growth rate), which are derived from the steepest slope (maximum slope points) of the exponential phase (Blomberg 2011). Once the yeast cell population reaches a certain point, the growth rate begins to decline and enters the 'Deceleration phase', and eventually a plateau is reached where further yeast division essentially stops due to an exhaustion of nutrients and/or accumulation of waste in the 'Stationary phase' (Hall et al. 2013).

The assessment of growth characteristics between different AQP expressing yeast means we can analytically compare lag times and growth rates between treatments, but also between replicates, to monitor the health of individual yeast cultures within and between replicate experiments, providing us with consistency between experiments and greater confidence in the data. Furthermore, it allows us to choose a standardised measuring point which is biologically appropriate and consistent across each AQP-expressing yeast culture.

Examining many different growth curves, we decided the most robust measuring point to standardise across our experiments occurs at the end of the active growth phase of untreated yeast, when the growth rate has declined to 5% of the maximum and the yeast are just about to enter the stationary phase (Figure 3.4). This carefully considered measuring point is applicable across different yeast strains and treatments, and allows us to capture periods of lag phase and maximal growth, both of which are informative in interpreting effects of treatments. It also excludes the period of non-active growth (i.e. stationary phase), therefore minimising the excessive disproportionate growth that occurs when the control untreated culture has fully plateaued, but the 'treated' cultures are often in a period of rapid growth (e.g. Time point 3, Figure 3.3). The Area Under the Curve (AUC) provides a cumulative measure capturing growth characteristics occurring during both the lag and growth phases of the yeast culture.

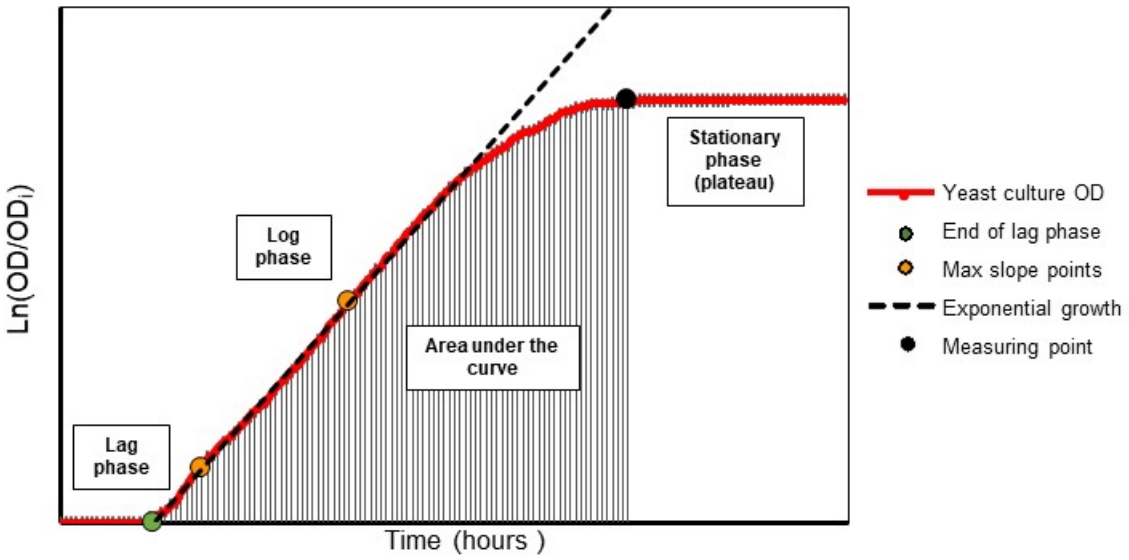


***Figure 3. 4 Parameterization of a yeast growth curve**. Characteristic parameters of yeast growth curve, plotted as measurements of yeast culture OD (Ln(OD/OD$_i$)) over time, include the Lag phase, Log phase and Stationary phase. Exponential growth (black dashed line) fitted through the "Max slope points" (orange circles), is used to define the end of Lag phase (green circle, intercept with the x-axis). The "Measuring point" (black circle) is placed where growth rate (slope of the growth curve) has declined to 5% of the maximum growth rate (growth at the maximum slope), indicating the start of the stationary phase. The Area under the curve (grey) can then be calculated as a measure of yeast growth incorporating dynamics in lag phase, log phase and stationary phase.*

3.3.4 Measured differences in growth between constructs grown in same conditions

AQP expression in yeast can lead to alteration in yeast growth (Pettersson et al. 2006). Therefore, it is important to consider that expression of a foreign AQP might alter parameters of the yeast growth curves, with each construct potentially imposing slight variation in yeast lag times, growth rates and end of log phase.

Our microculture growth assays enable us to individually monitor and measure yeast growth of each expression construct, obtaining independent measuring points for each AQP tested and experimental batch (three experimental batches, each with 2 biological replicates). Conducting several experimental batches accounts for any variation which might arise from potential media

preparation/treatment differences. As anticipated, the growth curves of various NtAQP and Empty vector yeast cultures (*skn7* yeast, Untreated) differed when their OD was monitored over a 24 hour period (Figure 3.5a). This was evident in the differing maximum growth rates ($d(OD_{600}/OD_{600i})/dt$) of the yeast cultures (Figure 3.5b). We did see some correlation between spotting assay results and growth phenotypes of yeast in liquid microcultures, with NtNIP5;1t having a faster growth rate in liquid media compared to Empty vector (Figure 3.5b), and also being one of the better growing constructs on the 'Untreated' media in the spotting assay (Figure 3.2). The AQP-specific variation in yeast growth reinforces the importance of quantifying relative sensitivities of cultures for particular treatments to the 'Untreated' sample of each construct, and not the 'Empty vector'/negative control.

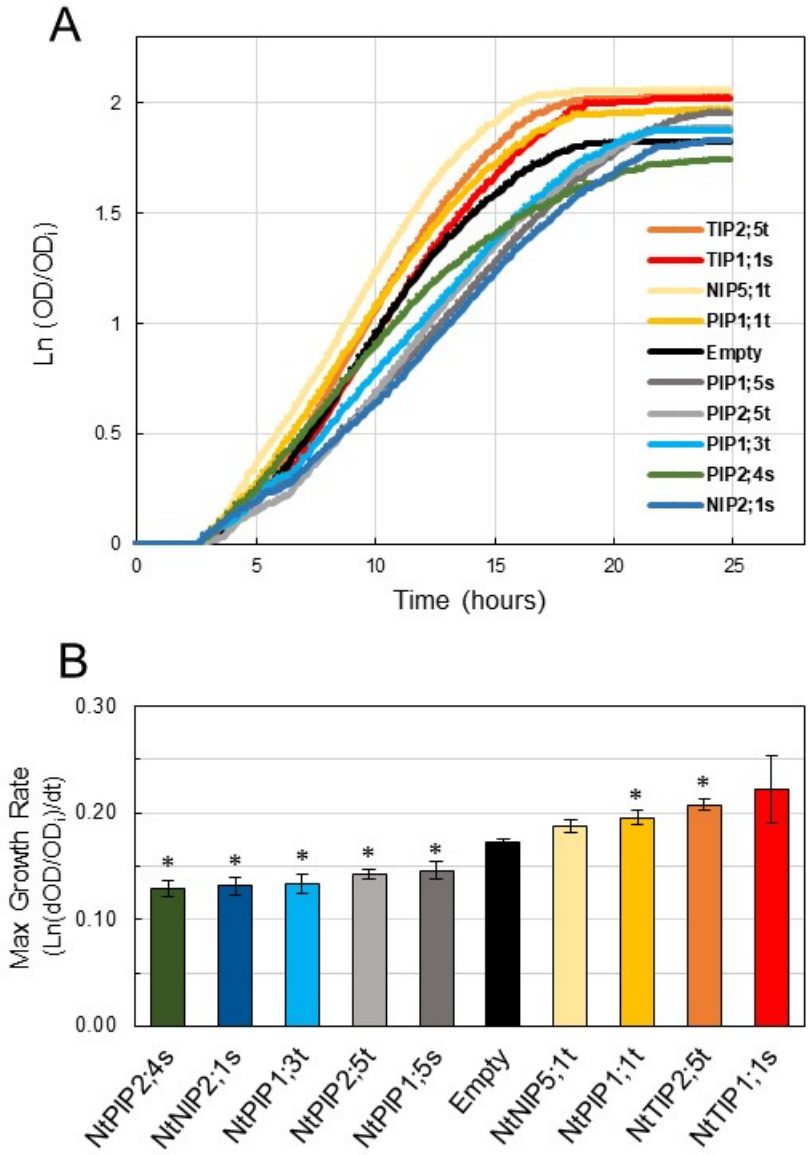


Figure 3.5 Comparison of growth over time of skn7 yeast expressing Empty vector and tobacco AQPs. ***A.*** *Growth curves Ln(OD/OD$_i$) vs. time (hours) of tobacco AQPs (NIP5;1t, PIP1;5s, PIP2;4s, TIP1;1s, TIP2;5t, PIP2;5t, PIP1;3t, PIP1;1t, NIP2;1s) and Empty vector control.* ***B.*** *Histogram of Maximum growth rate (dLn(OD/OD$_i$)/d***t***) of each AQP construct and the Empty vector control for the H_2O_2 permeability assays. N=6, Error bars= Standard error. Asterisks (*) indicate Students' t-test results of Max Growth Rates statistically different ($p<0.05$) to Empty vector control.*

3.3.5 Development of H_2O_2 toxicity microculture assay: optimising treatment concentrations

An optimisation experiment was required to determine the appropriate H_2O_2 concentrations for use in a toxicity assay. H_2O_2 treatment is toxic to yeast and will reduce culture growth, with the effective toxicity determined by the concentration of H_2O_2 and the relative native permeability of the *Δskn7* plasma membrane (PM) to H_2O_2. Ideally, we desire a range of H_2O_2 concentration with little to moderate effects on native *Δskn7*, which would be substantially increased if a H_2O_2 permeable AQP is heterologously expressed.

H_2O_2 treatments of 0.05mM, 0.1mM, 0.25mM, 0.5mM, 1mM, 2mM, 4mM and 8mM H_2O_2 were tested using *Δskn7* yeast harbouring the MG0515 Empty vector (Figure 3.6), so they could be grown in the same selective medium as the AQP expressing constructs. The effects of the different H_2O_2 concentrations on *Δskn7* yeast growth ranged from undetectable (0.05mM and 0.1mM) to severe toxicity (≥2mM) (Figure 3.6). Minor to moderate impacts on growth were seen at 0.25mM, 0.5mM and 1mM; these concentrations were chosen for future H_2O_2 screens of NtAQP expressing yeast. Using 3 treatment concentrations for the H_2O_2 liquid microculture assays should improve the resolution in assessing and comparing the sensitivity of the NtAQP expressing yeast to H_2O_2.

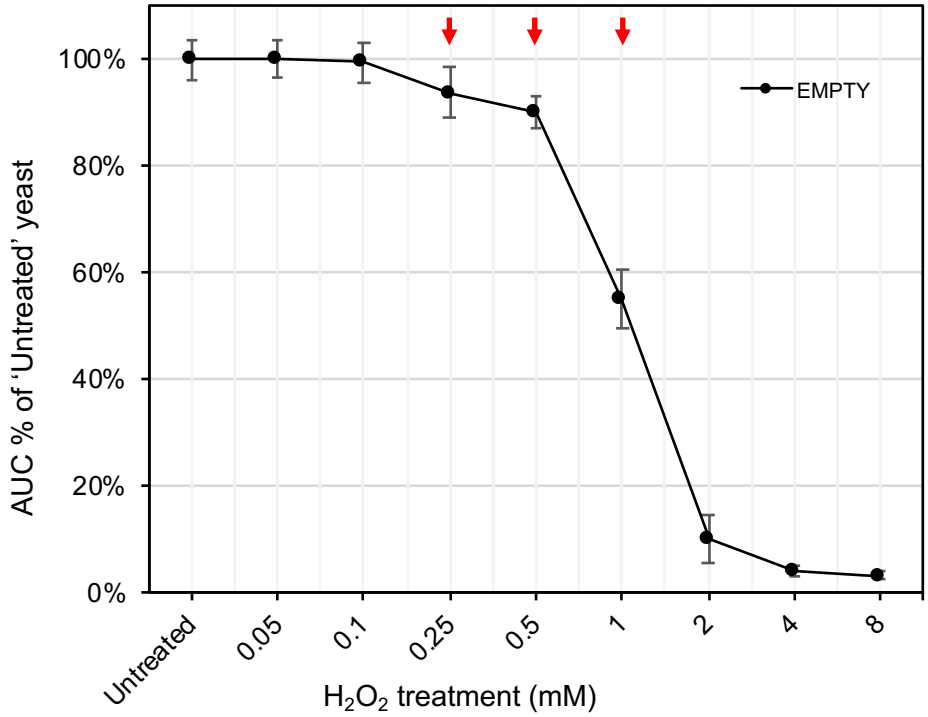


Figure 3. 6 The effect of H_2O_2 concentration on yeast growth relative to untreated controls. *Growth is represented as relative area under the growth curves (AUC, %) for snk7 yeast expressing Empty vector, exposed to various H_2O_2 treatments (0.05mM, 0.1mM, 0.25mM, 0.5mM, 1mM, 2mM, 4mM and 8mM H_2O_2). N= 6-4, Error bars=SE. Red arrows indicate treatment concentrations selected for future H_2O_2 toxicity assays.*

3.3.6 Development of Urea Growth based assay: optimising cell count for growth assay

Following the establishment of the H_2O_2 toxicity assay, we applied similar experimental techniques to set up a Urea growth-based assay. For this assay we used the *ynvwI* (*Δdur3*) yeast strain, which carries a deletion of its native DUR3 urea transporter and is incapable of growing on medium contains <5mM urea as the sole nitrogen source (Liu et al. 2003). Therefore, screening AQPs for improved *Δdur3* yeast growth on urea limited medium, should enable the identification of efficient urea transporting NtAQPs.

The *ynvwI (Δdur3*) yeast expressing the empty vector control was used to establish the growth requirements in media with urea as a sole nitrogen source.

For the H_2O_2 toxicity assay (Figure 3.6), we used a starting cell concentration of $0.6x10^7$ cells/ml, or ~$1.14x10^6$ cells, for each microculture growth assay. This cell population was sufficient for the yeast cultures to reach growth plateau within a reasonable 24h period of growth (in untreated conditions). Furthermore, $0.6x10^7$/ml cell density were also used for other microculture growth based assays established within our laboratory such as Freeze-thaw assay (testing for water permeability) and the Boric Acid toxicity assay.

However, because yeast grow more slowly when metabolising urea as a sole nitrogen source (Godard et al. 2007), using the $0.6x10^7$ cells/ml (or ~1.14 million cells) took ~53 hours before growth plateaued (Figure 3.7a). The *ynvwI* yeast expressing the empty vector (grown in 16mM Urea) had a growth rate of 0.065 OD/hr (data not shown), which is 2.6 times slower than *skn7* yeast expressing the empty vector and grown in medium containing the yeast preferred nitrogen source of ammonia sulfate (0.172 OD/hr; Figure 3.5b). To shorten the duration of the run-time, for convenience, assay turnover and also reliability, we doubled the starting cell density to $1.2x10^7$ cells/ml. This resulted in the culture reaching stationary phase in a more favourable ~40 hours (Figure 3.7b).

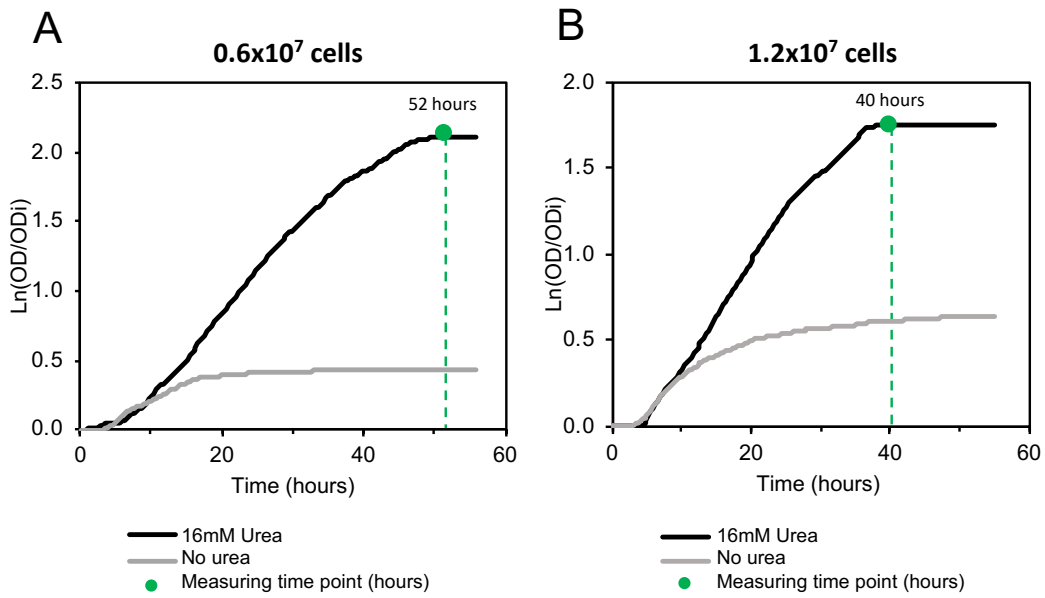


Figure 3.7 Effect of cell number on growth over time for ynvwI (Δdur3) yeast growing on media with no urea and 16mM Urea. *Yeast growth, Ln(OD/OD$_i$), was measured over time (53 hours) and the reading point of yeast cultures with $0.6x10^7$ (A) and $1.2x10^7$ (B) cells derived from growth curves.*

Doubling initial cell number from 0.6x10^7 to 1.2x10^7 reduced the time taken to reach the reading point by 12 hours.

3.3.7 Development of the Urea growth-based microculture assay: optimising treatment concentrations

Once the starting cell number was optimised for yeast growth in urea-only media, we examined the response of growth to urea concentration for yeast expressing either the empty vector control or the Arabidopsis TIP2;3, a urea-permeable AQP (Dynowski et al. 2008). The AtTIP2;3 positive control was included to ensure a differential response when a urea-permeable AQP is present in the yeast, thereby enhancing yeast growth compared to the empty vector control in urea-limited conditions.

The empty vector control and AtTIP2;3-expressing yeast were grown in microcultures with either 0mM, 4mM, 8mM, 12mM or 16mM urea concentrations; with 16mM having been established as a sufficient concentration to allow even the empty vector control to grow (Figure 3.7).

The growth curves of the *ynvwI* (*Δdur3*) empty vector (Figure 3.8a) and *ynvw1* (*Δdur3*) AtTIP2;3 expressing yeast (Figure 3.8b), both showed an obvious growth response to changing urea concentrations in the medium. Equal minimal growth at 0mM indicated a dependence on urea as a sole nitrogen source, with no advantage at this deficient level to either yeast lines. Both yeast lines also grew equally well at 16mM urea, indicating a sufficient saturating concentration to overcome the limited permeability of the native *ynvwI* (*Δdur3*) plasma membrane to urea. Yeast expressing the empty vector showed an almost linear response from 0mM urea through the evenly incremented increases to 16mM urea (Figure 3.8c). By contrast, the AtTIP2;3 expressing yeast showed a superior growth response that was most prominent at the lower 4mM and 8mM concentrations (Figure 3.8c), with a 30-35% greater AUC value than the Empty vector. In essence, expressing AtTIP2;3 provided a growth response at 4mM urea equivalent to supplying *ynvwI* yeast expressing Empty vector with 12mM urea (i.e. ~3x improvement in membrane permeability to urea). Overall, the enhanced growth of the urea permeable AtTIP2;3

expressing *ynvwI* (*Δdur3*) yeast, indicates that our assay is sufficiently sensitive to screen for other urea permeable AQPs when minimal concentrations of urea provide the sole nitrogen source. To stream-line the assay, and maximise samples processed in a given run, we selected 0mM, 4mM and 12mM urea concentrations as treatments for future yeast heterologous functional characterisation experiments.

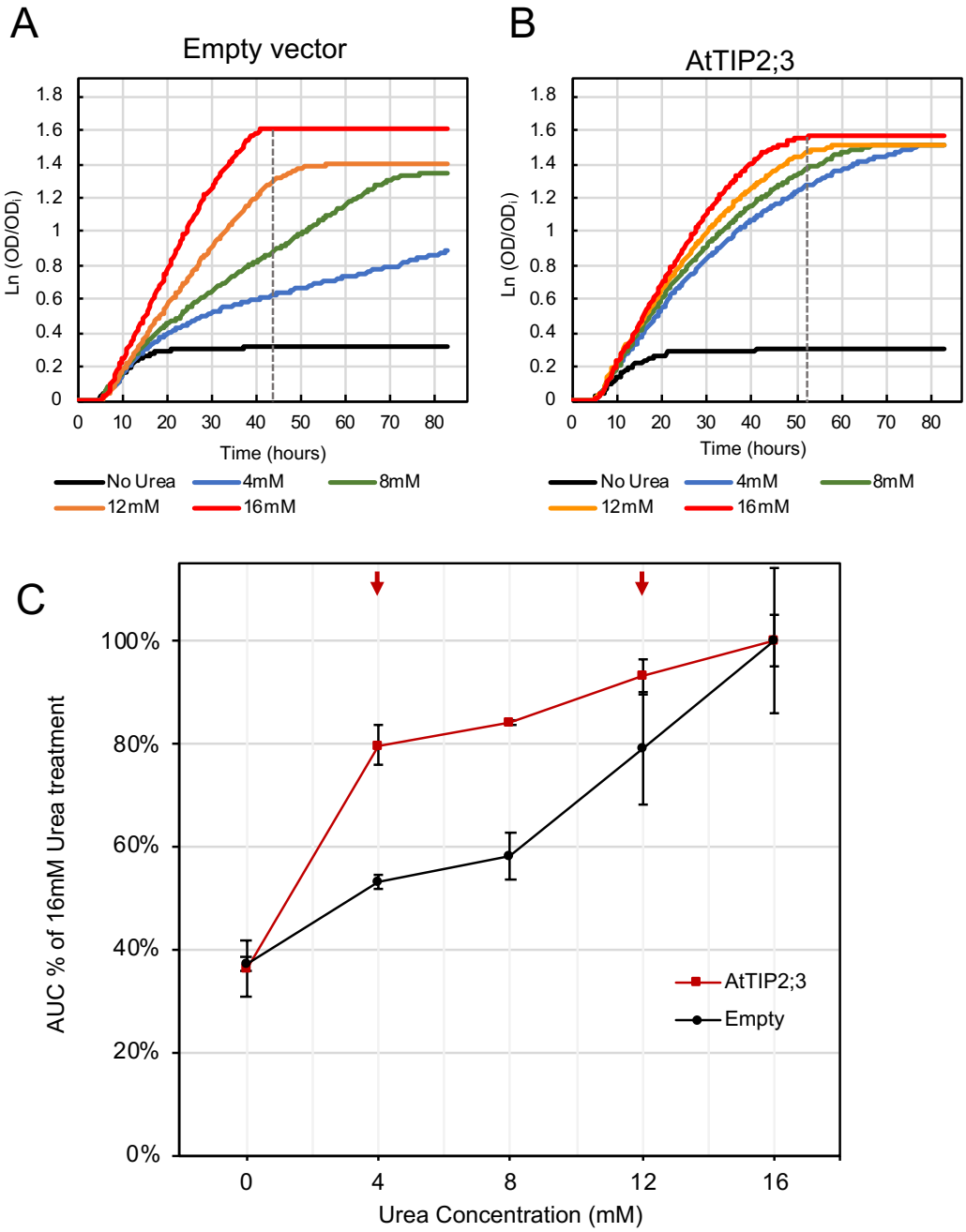


Figure 3. 8 Growth differences between ynvwI (Δdur3) yeast expressing Empty vector and AtTIP2;3 (urea transporter). *Yeast growth (Ln(OD/OD$_i$)) over time for yeast expressing Empty vector (**A**) or AtTIP2;3 (**B**), grown with no urea (0mM) or with 4mM, 8mM, 12mM or 16mM urea treatments.*

Grey dotted line indicates the measuring point calculated for the 16mM urea treatment growth curves, used for the area under curve comparisons of each treatment. ***C.*** *Differences in percent yeast growth (area under curve, AUC) compared to 16mM Urea treatment, for ynvw I (Δdur3) yeast expressing Empty vector (black) or AtTIP2;3 (dark red) and grown on 0 mM, 4 mM, 8 mM and 12 mM Urea. Increased growth of AtTIP2;3-expressing yeast at 4mM and 8mM urea concentrations. N=2; Error bars=SE. Red arrows indicate concentrations which were chosen for future Urea functional growth assays.*

3.3.8 Impact of GFP fused to AQP on substrate permeability - Urea assay

In initial attempts to measure growth response to different urea concentrations, we used an expression vector which had an N-terminal GFP translational fusion (vector MG0522; see Materials and Methods) to the AQP of interest. This vector was chosen for the urea growth-based assay as it provides expression of the URA3 gene which complements the mutation in the native URA3 gene in the *ynvwI (Δdur3*) yeast strain (Liu et al. 2003). Previous research had not indicated that GFP fusions could greatly affect an AQPs' functionality (Fetter et al. 2004; Gao et al. 2005). We believed that the MG0522 expression vector could allow us to screen for potential growth enhancements of yeast expressing a urea-permeable AQP (AtTIP2;3), when grown in N-limiting media. Empty vector and GFP:AtTIP2;3 *ynvwI* (*Δdur3*) yeast were grown in liquid microcultures with no urea (0mM), or with 4mM, 8mM, 12mM and 16mM urea concentrations. Figure 3.9 shows urea-dependent growth responses for yeast expressing the Empty vector and GFP:AtTIP2;3, where increases in nitrogen source increased the growth of the yeast cultures. However, unlike results from AtTIP2;3 alone (i.e. no GFP fusion; Figure 3.8), we did not see growth enhancement of yeast expressing GFP:AtTIP2;3 compared to Empty vector at low urea concentrations. This is shown in the growth curves (Figure 3.9a,b) and AUC % of 16mM Urea treatment calculations (Figure 3.9c) where both constructs show similar responses of growth to urea concentration. These results suggest that the N-terminal GFP fusion might be impairing the urea permeability of AtTIP2;3.

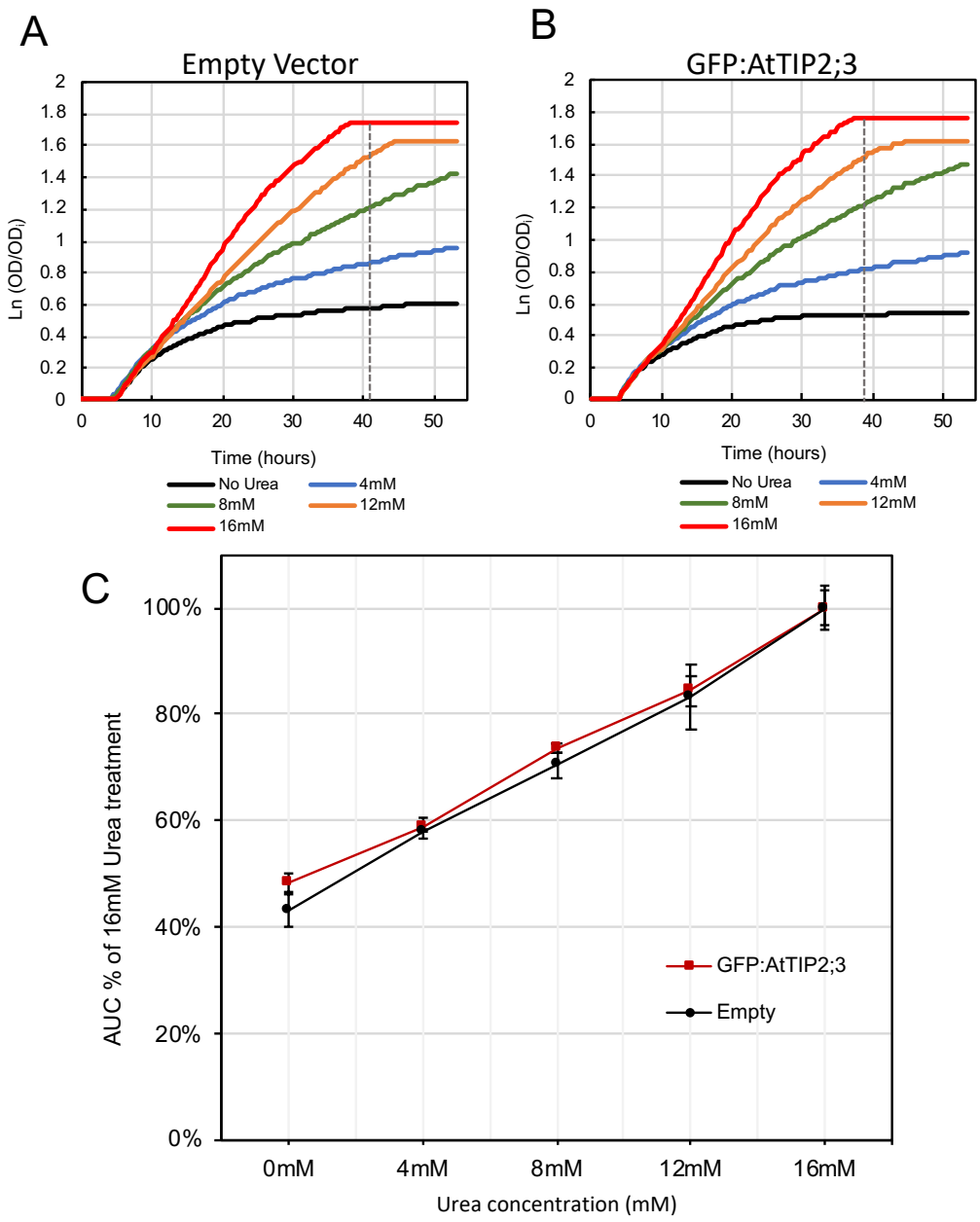


Figure 3. 9 ***Impaired Urea permeability of ynvwI yeast expressing AtTIP2;3 (urea transporter) with N-terminal GFP translational fusion.*** *Yeast growth (Ln(OD/OD$_i$)) over time for yeast expressing Empty vector* ***(A)*** *or GFP:AtTIP2;3* ***(B)****, grown with no urea (0mM) or with 4mM, 8mM, 12mM or 16mM urea treatments. Grey dotted line indicates the measuring point calculated for the 16mM urea treatment growth curves, used for the area under curve comparisons of each treatment.* ***C.*** *Percent yeast growth (area under curve, AUC) relative to 16mM Urea treatment, for ynvw I (Δdur3) yeast expressing Empty vector (black) and GFP:AtTIP2;3 (grey) grown in 0 mM, 4 mM, 8 mM and 12 mM Urea. Urea-dependent increases in growth were observed across each treatment as nitrogen source (urea) increased. No enhancements in growth were observed for GFP:AtTIP2;3 compared to Empty vector control. N=2; Error bars=SE.*

In order to validate our hypothesis, we investigated whether AtTIP2;3 N-terminally fused to GFP successfully integrated into the yeast cell plasma membrane. We imaged yeast cells expressing the MG0522 empty vector and the GFP:AtTIP2;3 (MG0522-AtTIP2;3) constructs (Figure 3.10). Expression of MG0522 Empty vector resulted in cytoplasmic distribution of GFP in the yeast cells, whereas yeast expressing MG0522-AtTIP2;3 had GFP signal around the periphery of the yeast cell, in addition to ER and vacuole (Figure 3.10). These subcellular localisation results illustrate the successful integration of GFP:AtTIP2;3 into the yeast cell plasma membrane. Therefore, we could conclude that the lack of growth enhancement in GFP:AtTIP2;3 expressing yeast at low urea concentrations was not due to its inability to integrate into the plasma membrane. Also a potential explanation for the impairment of urea permeability through GFP:AtTIP2;3, could be that the the GFP fusion might be reducing protein abundance of AtTIP2;3 in the yeast cells; but that this is less likely given that the confocal images (Figure 3.10) show bright fluorescence signal, indicating an abundance of protein. Instead, it is possible that the fusion of GFP to AtTIP2;3's N-terminus is impacting its ability to facilitate urea diffusion across yeast cell membranes.

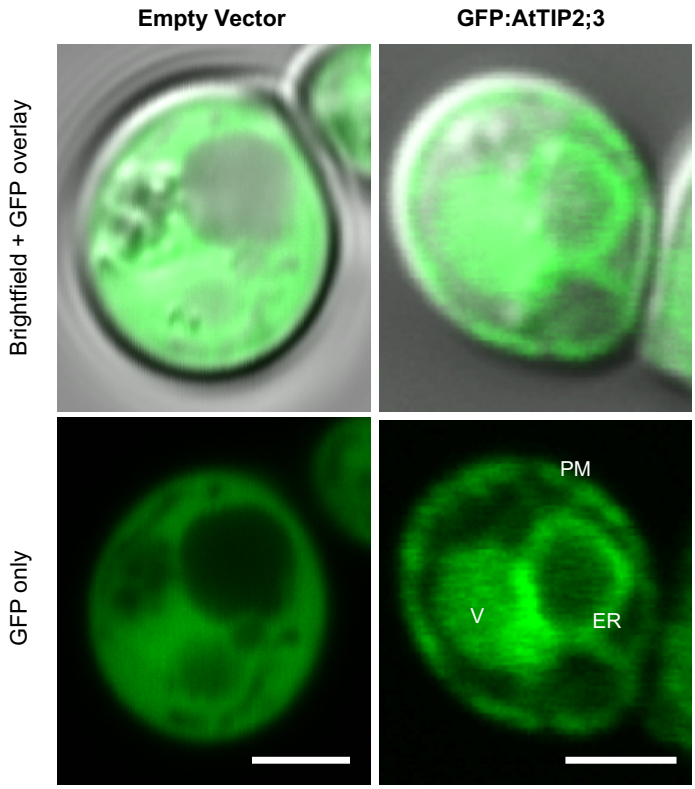


Figure 3. 10 Subcellular localisation of MG0522 Empty vector and GFP:AtTIP2;3 (MG0522-AtTIP2;3) in yeast cells. *For each construct we report a Brightfield + GFP overlay image of a yeast cell and a GFP only image . Plasma membrane (PM), ER and vacuole (V) are labelled. Scale bar 2 µm.*

3.4 Discussion

3.4.1 Advantages of yeast-based liquid microculture assays

In planta characterisation of AQPs has been limited due to the high number of isoforms occurring in plants (Abascal et al. 2014) and also because plant experimental methods are time consuming (e.g. plant transformations and phenotyping of transgenic plant lines).

A pressing issue for heterologous assays is their replicability between different labs, to ensure that high quality and comparable data is generated (Blomberg 2011). Yeast is a valuable heterologous expression system to investigate which substrates are permeable through each AQP, allowing us to efficiently survey AQP isoforms and thereby furthering our understanding of AQP biology.

Spotting assays are a common technique for heterologous expression of AQPs in yeast. AQP-expressing yeast is grown and diluted to a specific OD, and then

set micro-volume (e.g. 10μL) spots of culture are applied onto agar-media containing different concentrations of the molecule being tested (e.g. different H_2O_2 concentrations, Figure 3.2). Following an incubation, differences in spot size between treatments can be linked to sensitivity to a particular molecule i.e. a smaller spot size/decreased growth indicating higher sensitivity to a specific molecule. Spotting assays are a relatively cheap and easy technique, not requiring any specialised equipment. While appearing to be a simple method that is easy to interpret, the data collected from spotting assays can often lead to detection of false positives (Blomberg 2011). This is predominantly due to mis-interpretation of potential inconsistencies in yeast growth (arising both from variation across replicates and/or resulting from AQP-specific expression) and assigning such phenotypes to a particular treatment effect. Additionally, interpreting results from a yeast spot at a single time point can be limiting and subjective, having no quantitative information from which to estimate the growth stage of the yeast culture, i.e. whether it has entered exponential phase or has plateaued (single time point-based interpretation issues are addressed in Figure 3.3). A yeast spot which might have delayed growth in response to a specific treatment, but that has a faster growth rate, might catch up in size to a non-sensitive yet slow growing spot (appearing to have similar response to a treatment). Previous studies have tackled this issue on agar-media based assays, by acquiring sequential images of yeast spots over time, quantifying changes in spot area and extrapolating their growth rate (Shah et al. 2007; Blomberg 2011). Although monitoring growth of a yeast spot over time is more informative than knowing the area of a spot at a single time, growth rates estimated from agar-based experiments have a greater variance compared to those estimated from liquid cultures (Shah et al. 2007).

Our assay methods aim to improve issues surrounding the reliability and replicability of results, presenting a novel strategy to screen for substrate permeability of AQPs. The shaking plate reader incubator used for our liquid microculture assays enables us to monitor yeast growth by measuring the OD of the culture at short (10 minute) intervals. These measurements are automated and their frequency provide detailed information on subtle variations in growth

characteristics between various AQP expressing yeast lines. Additionally, we can identify and often account for instances of human or/instrumental error, that may influence the data. For instance, minor inconsistencies in yeast cell number allocation across the microculture wells can be corrected by normalising with the initial OD of each well. In spot assays, final yeast cell number actually spotted onto the agar plate cannot be conclusively determined. Variations in cell numbers could have substantial implications on subsequent density of colony forming units (CFU) and exposure to agar-medium treatment. An added advantage of liquid culture assays is the improved exposure of yeast to the treatment substrate through the continuous shaking of cells within the culture media, unlike the static contact of yeast spots with agar media containing the various treatments in spotting assays.

Our liquid microculture assays further allows for the selection of a consistent and biologically meaningful 'measuring time' point. This is done by calculating the start of the stationary phase for the 'untreated' sample's growth curve (Figure 3.4) for each AQP-expression construct. A tailored detection of measuring point is particularly important as yeast cultures expressing various AQP constructs might enter the stationary phase at differing times due to reasons such as differences in growth rates, sensitivity to accumulated toxins within the media or alterations in metabolism resulting from permeation of certain substrates. Therefore, tailoring the measuring time point is crucial to provide an appropriate interpretation of growth phenotypes in response to various treatments (as demonstrated in Figure 3.3). Having established the measuring time point obtained from the 'untreated' sample, the area under the curve (AUC) can be calculated for each concentration. This captures changes occurring during both the lag and growth phases of the culture growth. By having an accurate, repeatable and flexible method of measuring yeast growth, we can efficiently measure and detect phenotypes. The method provides a novel and valuable assay technique.

3.4.2 Impacts of H_2O_2 toxicity on Δskn7 *yeast growth*

H_2O_2 is a reactive oxygen species (ROS) involved in key processes regulating the physiology of cells and organisms (Bienert and Chaumont 2014). In plants, H_2O_2 transport across membranes is necessary for signal transduction, influencing gene

expression and triggering tolerance to environmental stresses (Bienert et al. 2007). AQPs have been implicated in facilitating transmembrane diffusion of H_2O_2; with its cellular redistribution allowing for fine adjustment of concentrations in the cytoplasm, intracellular organelles, apoplast and extracellular space (Hooijmaijers et al. 2012; Bienert and Chaumont 2014). Identifying AQPs that might be involved in facilitating H_2O_2 transport could enrich our understanding of ROS signalling and its implications within the plant.

The *Δskn7* yeast strain has a defective ROS scavenging system, resulting in reduced growth and survival upon exposure to H_2O_2 (Bienert et al. 2007), as such it is a valuable system to screen for AQPs that might be enhancing membrane permeability to H_2O_2.

Our newly established liquid microculture assay allowed us to monitor the growth effects on *Δskn7* yeast expressing the MG0515 empty vector control upon exposure to increasing concentrations of H_2O_2. The 1mM H_2O_2 treatment appeared to be the maximal H_2O_2 concentration that the *Δskn7* yeast (expressing empty vector) could tolerate, having a growth reduction to 55% of the untreated cultures; with all other higher H_2O_2 treatments being essentially fatally toxic, impairing growth to only 9-6% relative to untreated. The growth response of control *Δskn7* yeast provided us with a baseline toxicity response to H_2O_2, from which we can screen for AQPs that lead to further decreases in growth at each concentrations. This will enable us to functionally assess the H_2O_2 permeability of novel tobacco AQPs; results are shown in Chapter 4.

3.4.3 Impacts of Urea on ynvwI (Δdur3) yeast growth

Nitrogen is a key essential nutrient required for growth of all organisms which can be taken up and metabolised in a variety of different forms (Godard et al. 2007). Urea is a commonly used nitrogen source in plant fertilisers. It can be taken up by plants from the soil through specific transporters (e.g. DUR3), or through facilitated diffusion via urea-permeable aquaporins (Kojima et al. 2007; Gao et al. 2018). As well as being absorbed from the soil, urea can also be derived within cells

from precursor metabolites, and loaded in and out of vacuoles (e.g. via AQPs) when nitrogen might be in excess or limiting, respectively (Kojima et al. 2006).

In yeast, urea is one of the 27 distinct nitrogen sources which can be utilised in order for it to grow (Godard et al. 2007). Following entry into yeast cells, urea can be degraded into ammonia and subsequently assimilated into glutamate and other compounds involved in central nitrogen metabolism (Godard et al. 2007; Cooper 1982). Although urea can be used as a sole nitrogen source to sustain yeast culture growth, it is not a preferred source and results in slower yeast growth rates (Godard et al. 2007). In order to shorten the urea growth-based assay, we doubled the starting microculture cell number in the urea-only media, which meant that the cultures reached their stationary phase sooner (40 hours vs. 50+ hours, Figure 3.7), allowing for higher assay turnover.

We verified the efficacy of our urea growth-based liquid microculture assay, by testing growth differences of *ynvwI* yeast (*Δdur3*) expressing either an empty vector control or the urea-permeable AtTIP2;3. AtTIP2;3 has previously been reported to enhance membrane permeability of yeast to urea, providing a growth advantage on urea-limited agar media (Dynowski et al. 2008). Indeed, we saw enhanced growth of yeast expressing AtTIP2;3 compared to empty vector control in liquid media containing 4mM and 8mM urea, confirming the urea permeability of AtTIP2;3 (Figure 3.8). We also saw a decline in growth enhancement as urea concentration increased (Figure 3.8), suggesting that sufficient quantities of urea are able to diffuse into the *ynvwI* (*Δdur3)* yeast cells expressing the Empty vector with an external urea concentration of 16mM. Our liquid microculture growth-based assay will enable us to accurately identify new candidate urea-permeable AQPs to better understand AQPs' involvement in the nitrogen metabolism in plants.

3.4.4 GFP translational fusion affects urea permeability of GFP:AtTIP2;4 expressing yeast

GFP translational fusions have been crucial for advancements in AQP research, enabling membrane localisation of AQPs within cells to be visualised. It can be used to show which cell types a particular AQP might be functioning in and even shed light on AQP polarisation within cells. Such advancements have been

possible as the fusion of GFP (or other reporter proteins) to AQP protein does not appear to impair AQPs from reaching their sub-cellular destination within cells, nor does not it appear to alter AQP tetrameric composition (Umenishi et al. 2000).

Although AQP sub-cellular localisation is not affected by GFP translational fusion, we found that N-terminal GFP fusion appeared to disrupt the function of the urea-permeable AtTIP2;3 (Figure 3.10). Previously, impacts of GFP translational fusion on AQP function have been looked at in *Xenopus leavis* oocytes (Fetter et al. 2004; Umenishi et al. 2000), in mammalian cell lines (Gao et al. 2005) and in tobacco protoplast suspension cells (Reisen et al. 2003) . These studies measured impacts on the water permeability of AQPs from GFP translational fusion (to either N- or C-terminals) and identified that while AQP fusion with the GFP reporter protein did not interfere with its overall ability to transport water, it did appear to reduce transport efficiency (Fetter et al. 2004; Umenishi et al. 2000; Gao et al. 2005; Reisen et al. 2003). The total lack of urea permeation resulting from the GFP translational fusion observed here (Figure 3.9), as opposed to the reduction in water permeability observed in previous studies, could potentially be due to the molecular size difference of the two molecules. The combination of partial pore occlusion by GFP together with urea being a larger molecule, severely restricting urea permeability. The impact of AQP-GFP translational fusion has not been tested on permeability of substrates other than water, and has not been looked at previously in yeast expression systems. Our results indicate the importance of testing for AQP function in yeast without a GFP translational fusion.

3.5 Conclusions

Yeast based growth assays are a valuable system for assaying AQP permeating substrates, allowing for an improved understanding of substrate specificities. Interpretation of AQP-induced growth phenotypes in yeast spotting assays can be subjective, limited by lack of information regarding growth characteristics of yeast cultures and the detection of yeast growth at a single time point. We established liquid microculture yeast based assays allowing an improved and more reliable assessment of the permeability of AQPs to substrates such as urea

and H_2O_2. These novel assays involve continuous collection of yeast growth data over a 24-40 hour period (at 10-20 min intervals), the detection of a biologically meaningful measuring time point and utilisation of AUC measurements to compare yeast growth across various treatments.

Chapter 4: Functional characterisation of tobacco Aquaporins in *Saccharomyces cerevisiae*

4.1 Introduction

Plant AQPs are permeable to wide range of substrates with some isoforms acting as multifunctional channles. Elucidation of AQP permeating substrates through functional characterisation studies is a key component of understanding their potential roles in in plants, as well as enabling us to select candidate genes for crop engineering purposes. AQPs have a very conserved structure across organisms. They assemble as tetramers, with each monomer holding a functional pore created by 6 membrane spanning helices, 5 connecting loops and 2 shorter helices (Pommerrenig et al. 2015; Kirscht et al. 2016; Törnroth-Horsefield et al. 2006). Although the overall structure is highly conserved, slight deviations in structural and functional characteristics between isoforms contribute to differences in their transport selectivity. Such characteristics include pore dimension parameters (pore diameter and overall morphology); identities and flexibilities of pore-lining residues; and chemical configurations of pore constriction regions (Luang and Hrmova 2017).

Subfamily-specific substrate specificities in plant AQPs have been largely attributed to diversity in the ar/R Selectivity filter (SF) which forms the first constriction site towards the extracellular side of the pore. Variation in this site potentially delimits the substrates able to permeate across the membrane through the AQP (Hove and Bhave 2011; Sui et al. 2001). Additionally, the NPA motifs located at the centre of the pore act as a second constriction, with variation in residue composition contributing to selectivity for substrates such as ammonia, boric acid and urea (Wu and Beitz 2007; Hove and Bhave 2011).

Utilising the yeast functional assays described in Chapter 3 (testing for water, H_2O_2, boric acid and urea permeabilities), we aimed to functionally characterise a diverse subset of tobacco AQPs spanning the 3 largest AQP subfamilies; selecting 5 PIPs (NtPIP1;1t, NtPIP1;3t, NtPIP1;5s, NtPIP2;4s and NtPIP2;5t), 2 NIPs (NtNIP2;1s and NtNIP5;1t) and 2 TIPs (NtTIP1;1s and NtTIP2;5t). Our characterisation began by examining sub-cellular localisations in yeast with GFP tagged AQPs to confirm

incorporation of the expressed AQP in the plasma membrane. The confirmation of plasma membrane integration of the tested NtAQP isoforms in yeast was essesntial in order to correctly interpret the yeast permeability assay results. The plasma membrane integraton enabled us to assign enhanced or reduced yeast growth in the functional assays to AQP-facilitated diffusion across the yeast plasma membrane. Subsequently, each AQP construct was screened for substrate specificity in yeast, and 3D protein homology modelling was used to investigate variation in width and physico-chemical properties of the pores. Sub-cellular localisation *in planta* was also completed. The combination of these results present a comprehensive approach for a functional characterisation of AQPs.

4.2 Methods

4.2.1 Yeast strains, genes and yeast expression vectors used

aqy1ayq2, skn7 and *ynvwI* (listed in Chapter 3 Materials and Methods) yeast expressing NtAQPs listed in Table 4.1 (coding sequences were commercially synthesised, and cloned in appropriate expression vectors: MG0522 & MG0515 for *aqy1ay2* and *skn7* strain, MG0527 for *ynvwI* strain) was grown for 24 hours at 30^{o}C with shaking at 250rpm. 10 x 10µL spots of each expression construct were distributed on YNB media with either -HIS or -URA selection for either the MG0515 or for MG0527 & MG0522 selection respectively. Spotted plated were incubated at 30^{o}C for 2 days and then placed in the fridge (4^{o}C). Spotted plates were used for the starting cultures of functional assays (described below), using roughly ½ of a full spot as starting material for each culture.

Table 4. 1. List of genes selected for functional characterisation studies

Gene Name	Gene identifier
NtPIP1;1t	BK011393
NtPIP1;3t	BK011396
NtPIP1;5s	BK011398
NtPIP2;4s	BK011406
NtPIP2;5t	BK011409
NtTIP1;1s	BK011426

NtTIP2;5t	BK011440
NtNIP2;1s	BK011379
NtNIP5;1t	BK011387

4.2.2 Confirming plasma membrane integration of NtAQPs in yeast

Tobacco AQP GFP fusion constructs were generated via gateway cloning of coding sequences of 9 NtAQPs listed in table 4.2.1 from pZEO entry vector into destination vector MG0522 (pRS426-GPD-EGFP-ccdB vector (Alberti et al. 2007), vector map included in Section 3.2.3 -*Materials and Methods of Chapter 3*); which produced N-terminal GFP:NtAQP fusion proteins driven by the constitutive GPD promoter. The GFP only yeast expression was obtained using the empty MG0522 vector (no GOI fusion) which leaves the eGFP alone to be constitutively expressed via the GPD promoter. Yeast was grown overnight in 2mL YNB (-URA) liquid culture until an OD_{600} of 1-1.5 was reached. Then a 1mL aliquot of overnight culture was sub-cultured and diluted with fresh YNB (-URA) media to a total volume of 2mL and grown for 3-4 hours. A yeast aliquot was transferred into a 1.5mL Eppendorf tube for easier transportation. 10µL of yeast was mounted on polysine slide, coverslip was added and sealed with nail polish to avoid evaporation during imaging. Yeast cells were visualised with the Zeiss LSM 780 Confocal microscope using a 40x oil immersion objective (1.2 NA). Light micrographs of yeast cells were visualised using Differential Interference Contrast (DIC), with GFP fluorescence captured using excitation at 488 nm and emission detection across the 490-526 nm range. Images were processed using Fiji (ImageJ) program (Schindelin et al. 2012).

4.2.3 Water permeability ('Freeze-thaw') assay

Yeast (*aqy1aqy2* strain) cultures expressing MG0515 Empty vector (negative control) and MG0515-NtAQPs (listed in Table 4.1) were grown for 24-28 hours in 1.25mL YNB (-HIS) liquid media at 30°C with shaking at 250rpm until an OD_{650} of 0.5-1 was reached. Cultures were diluted to $0.6x10^7$ cells using YPD medium, and incubated at 30°C for 60 mins. 250µL of each culture was aliquoted to 2 Eppendorf tubes. One tube was placed on ice, being the 'Untreated' control, and

the other was used for the 'Freeze-thaw' cycles. Each 'Freeze thaw' cycle consisted of yeast cultures (in 250µl aliquots) being frozen in liquid nitrogen for 30 seconds, and subsequently thawed in a water bath at 30°C for 20 mins. Each 'treated' culture was exposed to two freeze thaw cycles. For each construct, 200µL aliquots of 'Untreated' and 'Freeze-thawed/treated' yeast were transferred into a 96 well plate and yeast growth was monitored using the SpectroStar nano microplate reader. The program consisted of incubation at 30°C, with measurement of OD_{650} at the end of each 10 minute kinetic cycle (running for 18 hours). Each experimental batch consisted of 2 biological replicates; 3 experimental batches were conducted in total (resulting in 6 replicates per construct).

4.2.4 H_2O_2 permeability assay

Yeast (*skn7* strain) cultures expressing MG0515 empty vector (negative control) or MG0515-NtAPQs (listed in Table 4.1) were grown for 28 hours in 1.25mL YNB(-HIS) liquid media at 30°C with shaking at 250rpm. Culture OD was measured (expected OD_{650} readings between 1-2) and two 2mL cultures (two biological replicates per construct) were made to have a total cell number of $0.6x10^7$ cells. Diluted cultures were incubated at 30°C whilst H_2O_2 solutions were prepared. 1M H_2O_2 was prepared by adding 5.1mL of 9.97M H_2O_2 to 44.9mL sterile water. A 40mM H_2O_2 stock (400µL 1M H_2O_2 added to 9.6mL sterile water) was then prepared to make up the H_2O_2 treatment solutions. Four 5mL tubes were labelled (1-4) and were prepared as listed in Table 4.2.

Table 4. 2 Preparation of H_2O_2 stock solutions for H_2O_2 toxicity assay.

Tube number	Stock Concentration	Final assay concentration **(10uL Stock in 200µL well)**	Preparation instructions
1	20mM	1mM	Add 5mL of 40mM H_2O_2 stock to 5mL sterile water
2	10mM	0.5mM	Add 5mL from Tube 1 to 5mL sterile water
3	5mM	0.25mM	Add 5mL from Tube 2 to 5mL sterile water
4	-	Untreated/Water	5mL sterile water

10µL of H_2O_2 treatments and 190µL of yeast cultures were distributed to their respective wells in the 96-well plate; yeast growth was monitored using the SpectroStar nano microplate reader. The program consisted of incubation at 30°C, with measurement of OD_{650} at the end of each 10 minute kinetic cycle (running for 24 hours). Each experimental batch consisted of 2 biological replicates; 3 experimental batches were conducted in total (resulting in 6 replicates per constructs).

4.2.5 Boric Acid permeability assay

Yeast (*aqy1aqy2* strain) cultures expressing MG0515 empty vector (negative control) or MG0515-NtAPQs (listed in Table 4.1) were grown for 28 hours in 1.25mL YNB(-HIS) liquid media at 30°C with shaking at 250rpm. Culture OD was measured (expected OD_{650} readings between 1-2) and two 1.5mL cultures (two biological replicates per construct) were prepared to have a total cell number of $0.6x10^7$ cells. Diluted cultures were incubated at 30°C whilst boric acid solutions were prepared from a 1M boric acid stock solution (30mM, 20mM and 10mM boric acid) as per Table 4.3.

Table 4. 3. Preparation of Boric acid stock solutions for Boric Acid toxicity assay.

Tube number	Stock Concentration	Final assay concentration **(20µL Stock in 200µL well)**	Preparation instructions
1	300mM	30mM	Add 300µL of 1M Boric Acid stock to 700µL sterile water
2	200mM	20mM	Add 200µL of 1M Boric Acid stock to 800µL sterile water
3	100mM	10mM	Add 100µL of 1M Boric Acid stock to 900µL sterile water
4	-	Untreated/Water	1mL sterile water

20µL of boric acid stocks (or Water) and 180µL of diluted yeast cultures were distributed to their respective wells in the 96-well plate; yeast growth was monitored using the SpectroStar nano microplate reader. As per H_2O_2 toxicity assay, the growth monitoring program consisted of incubation at 30°C; and kinetic cycles

of 10mins (running for 24 hours) with measurement of OD_{650} at the end of each cycle. Each experimental batch consisted of 2 biological replicates; 3 experimental batches were conducted in total (resulting in 6 replicates per constructs).

4.2.6 Urea permeability assay

Half yeast spots of *ynvwI* yeast strain expressing MG0527 empty vector (negative control) or MG0527-NtAPQs (listed in Table 4.1) were resuspended in 1.25mL of YB + Glucose media (culture medium without nitrogen source; see Table 4.4). Culture OD was measured and two 2mL cultures (two biological replicates per construct) were made using YB +Glucose media to have a total cell number of $1.2x10^7$ cells.

Table 4. 4. Preparation of basic YB + Glucose media.

	Basic YB + Glucose media
2x Yeast nitrogen base without amino acids (liquid media)	25 mL
20x Sterile Glucose	4.25 mL
Sterile Water	21.75 mL
Total	50mL

Table 4. 5. Preparation of urea stock solutions for urea growth-based functional assays.

Tube number	Urea stock Concentration	Final assay concentration **(10µL of stock in 200µL well)**	Preparation instructions
1	40mM	2mM	Add 80µL of 500mM urea stock to 920µL sterile water
2	80mM	4mM	Add 160µL of 500mM urea stock to 840µL sterile water
3	240mM	12mM	Add 480µL of 500mM urea stock to 520µL sterile water
4	-	Untreated/Water	1mL sterile water

500mM urea stock was made up (300mg urea dissolved in 10mL sterile water; filter sterilised) to prepare the 40mM, 80mM and 240mM stock solutions (required for 2mM, 4mM and 12mM urea treatments respectively; see Table 4.5). 10µL of urea stock solutions and 190µL of diluted yeast cultures were distributed to their respective wells in the 96-well plate; yeast growth was monitored using the SpectroStar nano microplate reader. The program consisted of incubation at 30°C,

with measurement of OD_{650} at the end of each 20 minute kinetic cycle (running for 48 hours). Each experimental batch consisted of 2 biological replicates; 3 experimental batches were conducted in total (resulting in 6 replicates per constructs).

4.2.7 Characterising in planta sub-cellular localisation of NtAQPs

Tobacco AQP GFP fusion constructs were generated via Gateway cloning of a *NtPIP1;1t, NtPIP1;3t, NtPIP1;5s, NtPIP2;4s, NtPIP2;5t, NtTIP1;1s, NtTIP2;5t, NtNIP2;1s* and *NtNIP5;1t* coding sequences from pZeo entry vectors into the pMDC43 destination vector (Curtis and Grossniklaus 2003); producing N-terminal GFP:NtAQP fusion proteins driven by the constitutive 2x35S CaMV promoter. Arabidopsis transgenic lines were generated via agrobacterium (GV3101) floral dipping plant transformation method (Clough and Bent 1998).

Arabidopsis seeds were liquid sterilised using hypochlorite, washed several times and sown on Gamorg's B5 medium containing 0.8% Agar and the antibiotic hygromycin for selection of transformants. After 8 days of growth, Arabidopsis seedlings were gently removed from the agar, mounted in Phosphate Buffer (100mM $NaPO_4$ buffer, *pH* 7.2) on a standard slide and covered with coverslip, and visualised with a Zeiss LSM 780 Confocal microscope using a 40x water immersion objective (1.2 NA). Light micrographs of cortical cells in the root elongation zone were visualised using Differential Interference Contrast (DIC), with GFP fluorescence captured using excitation at 488 nm and emission detection across the 490-526 nm range. Autofluorescence was detected in the 570-674 nm range and excluded from the GFP detection channel. Images were processed using Fiji (ImageJ) program (Schindelin et al. 2012).

4.2.8 3D protein homology modelling and characterisation of NtAQP pores

3D protein homology modelling was used to generate 3D models of the 9 NtAQPs listed in Table 4.1. The resolved crystal structure for Spinach PIP2;1 (Törnroth-Horsefield et al. 2006) and Arabidopsis TIP2;1 (Kirscht et al. 2016), were used as templates for our modelling analyses. For Spinach PIP2;1 (SoPIP2;1), two structural conformations were available as templates; differing in orientation of

Loop D either tucking in and occluding the pore cavity (Closed) or being displaced opening up the pore (Open). The SWISS-MODEL software (Waterhouse et al. 2018) was used to construct homology models of NtAQPs. The Open (PDB ID:2bf5.1.A) and Closed (PDB ID:1z98.1.A) conformations of SoPIP2;1 differed in resolution, having 4.9 Å and 2.1 Å resolutions respectively. We were interested in modelling NtAQPs in an Open conformation.We compared the modelled pore profile (results outputs from ChexVis program; described below) of a representative NtPIP (NtPIP2;4s) using both the Closed and Open SoPIP2;1 conformations as templates to ensure that the lower resolution of the Open structure was not inadvertently impacting on pore shape and width. No differences in pore properties were observed between the two NtPIP templates (data not shown).

For some NtAQPs, MUSCLE alignments were conducted in Geneious software, truncating N- and C-terminal regions which were overhanging from template in order to improve model homology scores (truncations listed in Table 4.7). Sequence similarity against the chosen template (SoPIP2;1 Open conformation or AtTIP2;1) and Global Model Quality Estimation (GMQE) scores (Guex et al. 2009) were obtained for each model generated . GMQE scores estimate the quality of the generated protein model, combining properties of the target-template alignment. 3D templates (SoPIP2;1 or AtTIP2;1) and the SWISS-MODEL-generated NtAQP 3D models were aligned and visually inspected in UCSF Chimera software (Pettersen et al. 2004). The NtAQP 3D model generated by SWISS-MODEL was the uploaded to ChexVis online software (Masood et al. 2015) to obtain pore characteristics: pore radius, pore lining residues, and pore lining residues' hydrophobicity and flexibility. Selectivity Filters and NPA regions (characterised in Chapter 2), were mapped along the pore profile of each NtAQP.

4.3 Results

4.3.1 Tobacco Aquaporin localisation in Yeast

The membrane integration of the tobacco aquaporins being tested in the yeast functional assays was determined using N-terminal GFP::AQP translational fusions (Figure 4.1), with confirmation of plasma membrane integration being

critical in validating results obtained from the yeast functional assays (results presented in following sections).

Optically thin sliced focal plane images were taken of the yeast to visualise the GFP signal localisation. These images were further processed using surface profiling and lines scans of signal intensity, to better assess the distribution of GFP::AQP signal within the yeast cells. Yeast expressing GFP alone, showed GFP signal throughout the majority of the cell, with signal excluded from the vacuole. The surface and line signal scans show that the intensity was relatively equally distributed. This pattern is consistent with the expected cytosolic localisation of GFP (Figure 4.1j). The fusion of NtAQPs to GFP, results in the redistribution of GFP fluorescence to different yeast sub-cellular compartments including; plasma membrane (PM), endoplasmic reticulum (ER), and/or the tonoplast (vacuolar membrane). The NtPIP1-GFP fusion localised to the periphery of the cell and the ER. Although signal intensity in the periphery varied, likely due to co-localisation in peripheral ER, the signal remained continuous around the cell, consistent with PM integration (Figure 4.1a-c). The ER signal for NtPIP1;5s, was frequently observed to include bright spots in the periphery of the cell (Figure 4.1a-c). The NtPIP2 proteins also integrated into the PM, with clearly defined peaks present in the line scans (Figure 4.1 d-e). GFP signal was also localised to the the ER and faintly inside the vacuole. NtNIP2;1s localised to the PM and ER, similar to the NtPIPs. NtNIP5;1 weakly localised to the PM and ER, with signal predominantly associated with integration into the tonoplast (Figure 4.1f-g). The NtTIPs had strong signal that was distributed between the ER, tonoplast and notably, the PM (Figure 4.1h-i).

The sub-cellular localisation in yeast differed between the NtAQPs, but importantly for the yeast functional assay, we could confirm plasma membrane integration for all NtAQP constructs tested.

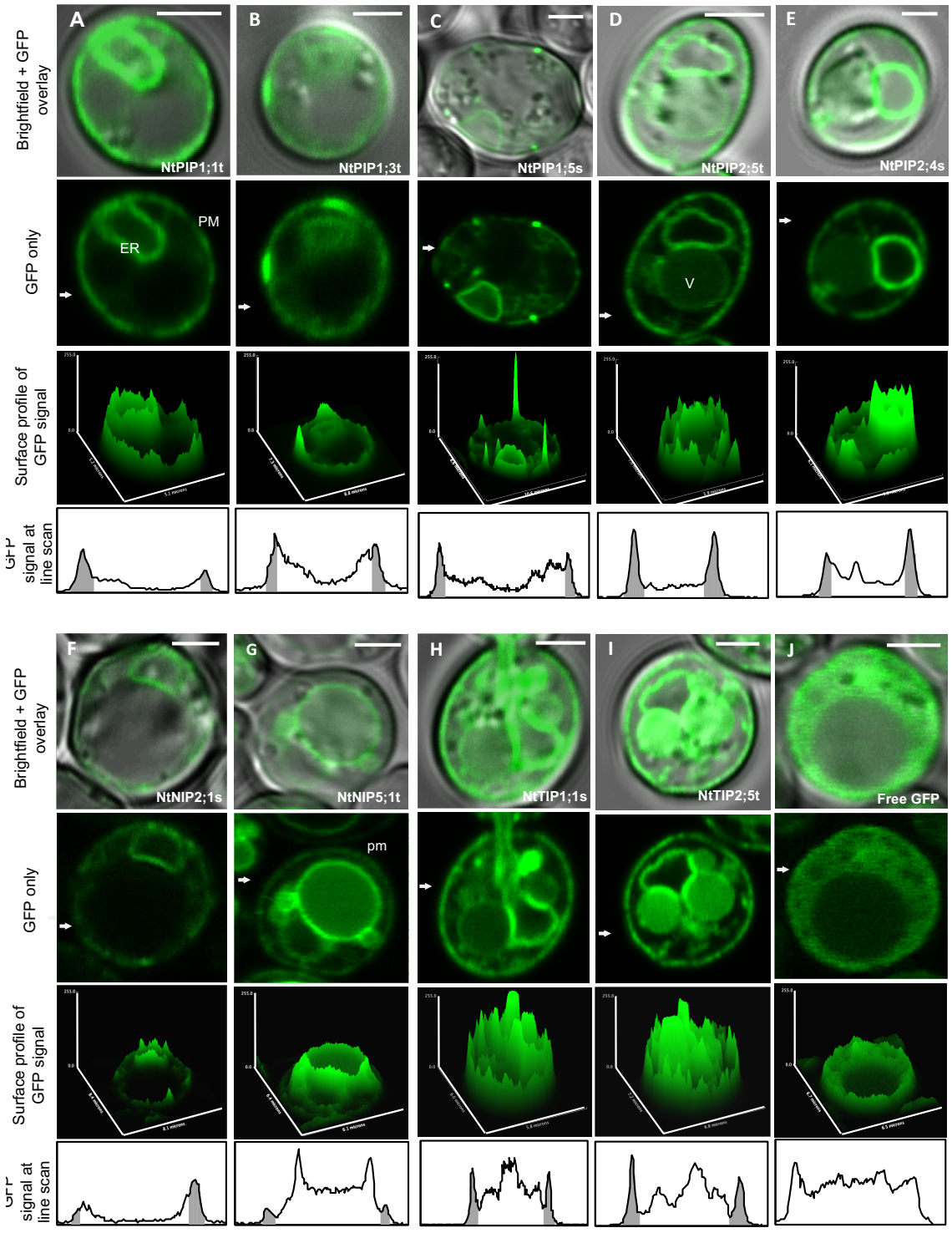


Figure 4. 1 Subcellular localisation of GFP tagged aquaporins expressed in yeast. Confocal microscopy images of yeast expressing NtAQP-GFP translational fusions of (A) NtPIP1;1t, (B) NtPIP1;3t, (C) NtPIP1;5s, (D) NtPIP2;4s, (E) NtPIP2;5t, (F) NtNIP2;1s, (G) NtNIP5;1t, (H) NtTIP1;1s, (I) NtTIP2;5t and (J) Free GFP localisation. For each construct we report a Brightfield + GFP overlay image of a yeast cell; a GFP only image; a surface plot profile of GFP signal intensity at the imaged focal plane; and a line scan of signal intensity traversing the cell (indicated by white arrow in GFP

only image). Grey shading in GFP signal line scan corresponds to regions which align with plasma membrane. Plasma membrane (PM), ER and vacuole (V) are labelled. Scale bar 2μm

4.3.2 Water permeability "Freeze-thaw" assay

In order test for water permeability of NtAQPs, we quantified the recovery of yeast cultures after exposure to Freeze/Thaw treatments. Improved survivorship of the snap-frozen yeast populations translated to improved growth of cultures (i.e. greater viable starting population). This was interpreted as evidence for rapid AQP-mediated movement of water across the plasma membrane, thereby avoiding cell death. Figure 4.2 shows the differences in growth between yeast expressing either a particular NtAQP or an empty vector (MG0515). Growth following exposure to two freeze-thaw cycles is expressed relative to Untreated yeast . The *aqy1aqy1* yeast expressing the Empty vector was unable to survive exposure to two freeze-thaw cycles and essentially failed to grow after treatment (2% growth relative to the untreated Empty vector culture, Figure 4.2). NtPIP2;4s and NtPIP2;5t had the greatest growth following the freeze-thaw treatment, achieving 70% of the untreated growth (Figure 4.2). NtTIP1;1s and NtTIP2;5t achieved 62% and 30% growth, respectively, relative to untreated controls. Thus, four out of the nine NtAQPs tested were able to increase the permeability of the plasma membrane to water sufficiently to allow the yeast to survive freeze-thaw treatments.

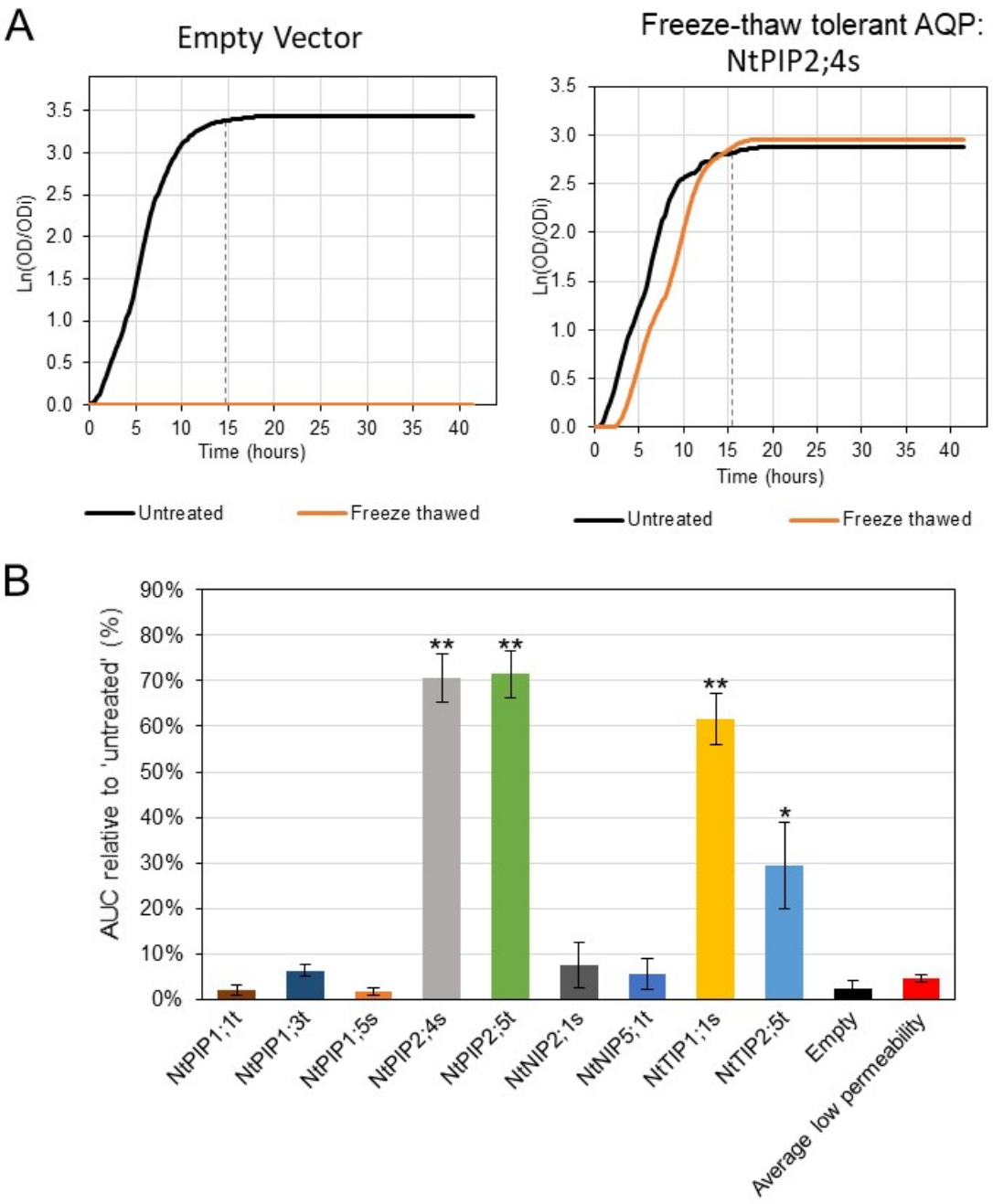


Figure 4. 2 NtAQP water permeability assessed with the 'Freeze-thaw' assay. **A.** Yeast growth curves, Ln(OD/OD$_i$) vs. time, of *aqy1aqy2* yeast expressing MG0515 Empty vector control or a freeze-thaw tolerant AQP (NtPIP2;4s), exposed to Freeze-thaw treatments. **B.** Yeast culture growth following the Freeze–thaw treatment (AUC relative to untreated yeast control, %) for *aqy1aqy2* yeast expressing Empty vector or one of the 9 NtAQPs (see legend). "Average low permeability", final red bar, represents the average value of all AQPs which appeared impermeable to water (PIP1;1t, PIP1;3t, PIP1;5s, NIP2;1s, NIP5;1t). Asterisk denotes significantly greater growth following the Freeze – thaw treatment compared against "Average low permeability" from a Student's t-test: "*" $p<0.05$ and "**" $p<0.01$, N=6, Error bars = SE.

4.3.3 H_2O_2 toxicity assay

For the H_2O_2 toxicity assay, *skn7* yeast expressing one of the 9 tobacco AQPs or an MG0515 Empty vector control was exposed to 3 different H_2O_2 treatments: 0.25mM, 0.5mM and 1mM H_2O_2. Exposing yeast carrying the empty vector to 0.25mM or 0.5mM H_2O_2, had no significant impact on growth, however, 1mM H_2O_2 caused a 37% decrease in growth relative to the untreated control (Figure 4.3). For 6 of the 9 NtAQPs tested (NtPIP1;3t, NtPIP1;5s, NtTIP1;1s, NtTIP2;5t, NtNIP2;1s and NtNIP5;1t), the yeast had H_2O_2 toxicity responses similar to the empty vector. These were classified as 'H_2O_2 – impermeable'. Their averaged growth values were used to establish a second negative control value termed the 'Average low permeability' (Figure 4.3). While the H_2O_2 toxicity response curve of the H_2O_2 impermeable NtAQPs was similar to yeast with the MG0515 Empty vector, the absolute growth was reduced by about 10% at each H_2O_2 concentration. Statistical comparisons were made against the "Average low permeability" as a more like-for-like comparison (i.e. all AQP expressing yeast).

Three NtAQPs increased sensitivity to H_2O_2 exposure: NtPIP2;4s, NtPIP2;5t and NtPIP1;1t. The most dramatic decline in culture growth was found for *skn7* yeast expressing NtPIP2;4s (growth reductions of 66% , 86% and 80% at 0.25mM, 0.5mM and 1mM H_2O_2, respectively, Figure 4.3). The lowest concentration of H_2O_2 that resulted in a dramatic reduction in growth varied from 0.25mM for NtPIP2;4s, 0.5mM for NtPIP2;5t and 1mM for NtPIP1;1t. We were unable to establish

whether this reflected different H_2O_2 permeability between AQP monomers, or arose due to other mechanisms.

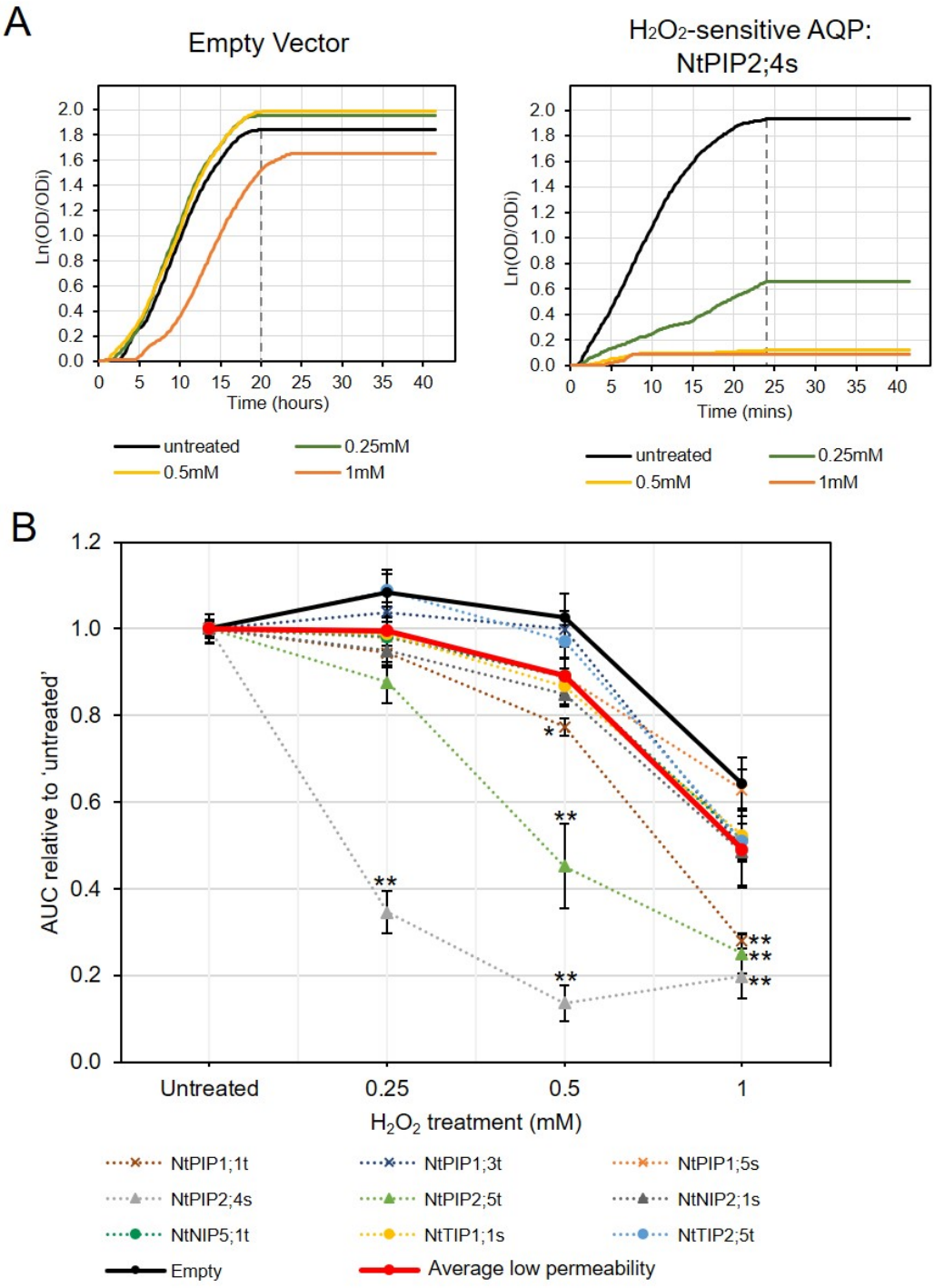


Figure 4. 3 NtAQP H_2O_2 permeability assay. **A.** Yeast growth curves, Ln(OD/OD_i) vs. time, of *snk7* yeast expressing MG0515 Empty vector control or an H_2O_2-sensitive AQP (NtPIP2;4s), exposed to 0.25mM, 0.5mM and 1mM H_2O_2 treatments. **B** Yeast culture growth relative to 'Untreated' control, for *skn7* yeast expressing an Empty vector or one the 9 NtAQPs surveyed with functional assays.

"Average low permeability", red line, indicates the average value of 6 AQPs which were impermeable to H_2O_2 (PIP1;3t, PIP1;5s, NIP2;1s, NIP5;1s, NtTIP1;1s and NtTIP2;5t). Asterisks denote Student's t-test results comparing H_2O_2-treated yeast growth against "Average low permeability"; "*" indicates $p<0.05$ and "**" indicates $p<0.01$, N=6, Error bars=SE.

4.3.4 Boric acid toxicity assay

To screen for AQPs which might enhance boric acid diffusion across the yeast plasma membrane, a boric acid toxicity assay was developed within our laboratory. The *aqy1aqy2* yeast expressing one of the 9 NtAQPs (in Table 4.1) or MG0515 Empty vector were exposed to different concentrations of boric acid (10, 20 and 30mM). Exposure of yeast expressing the Empty vector control to 10mM boric acid did not affect growth (Fig 4.4). However, growth was progressively reduced at greater concentrations (by 33 and 64% at 20 and 30mM Boric acid concentrations, respectively).

Three different phenotypes were observed across NtAQP-expressing yeast in response to boric acid treatments, noticeable at 20mM and 30mM boric acid concentrations. The first phenotype had a boric acid toxicity response that was similar to the MG0515 Empty vector (NtNIP2;1s, NtNIP5;1t and NtPIP1;5s). The second phenotype resulted in a greater tolerance to Boric acid, with the average growth being 20% greater than the yeast with the MG0515 Empty vector at 20 and 30mM boric acid (NtPIP1;1t, NtPIP1;3t, NtPIP2;4s, NtPIP2;5t)(represented by red line, Figure 4.4). The reduced toxicity response of these 4 NtAQPs is consistent with a reduction in the permeability of the plasma membrane to boric acid associated with the incorporation of the AQP protein. If so, the NtAQPs which had similar growth to the MG0515 Empty vector (NtNIP2;1s, NtNIP5;1t and NtPIP1;5s) could in fact be enhancing membrane permeability to boric acid, when compared to the "Average low permeability" group (classifying these as boric acid-permeable AQPs). The third phenotype displayed a 20-30% reduction in in growth relative to the empty vector yeast at 20 and 30mM Boric acid (NtTIP1;1s and NtTIP2;5t, Figure

4.4). This third phenotype is consistent with the AQPs being classified as boric acid-permeable.

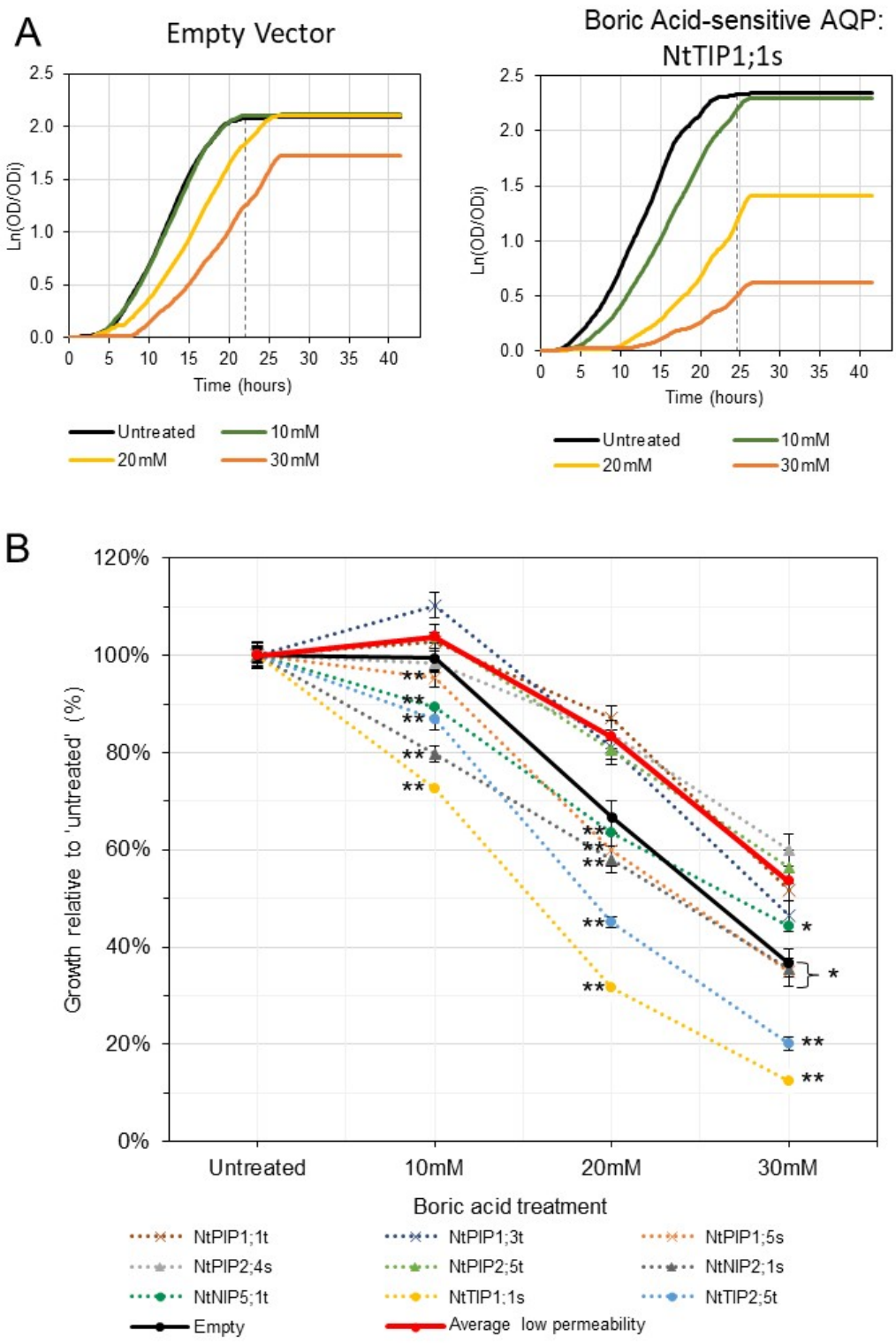


Figure 4. 4 Boric acid permeability assay for yeast expressing NtAQPs. **A.** Yeast growth curves, Ln(OD/OD$_i$) vs. time, of *aqy1aqy2* yeast expressing MG0515 Empty vector control or a boric acid-sensitive AQP (NtTIP1;1s), exposed to 10mM, 20mM and 30mM boric acid treatments. **B** Yeast

culture growth relative to untreated control (AUC, %) of *aqy1aqy2* yeast expressing either an Empty vector or one of the 9 NtAQPs exposed to boric acid. "Average low permeability" (red line) indicates the average value for the four AQPs which were apparently impermeable to boric acid (PIP1;1t, PIP1;3t, PIP2;4s and PIP2;5t). Asterisks denote growth that was significantly less than "Average low permeability" using a Student's t-test, "*" indicates $p<0.05$ and "**" indicates $p<0.01$, N=6, Error bars=SE.

4.3.5 Urea growth-based assay

In order to identify NtAQPs which might be permeable to urea, we used the growth-based assay established in Chapter 3. *ynvwI y*east expressing one of the 9 NtAQPs listed in Table 4.1 or the MG0527 Empty vector control were grown in No urea, 2mM, 4mM and 12mM urea treatments. 12mM Urea provides sufficient nitrogen for the growth of yeast cultures to reach a plateau within a 50 hour incubation. The 2mM urea treatment concentration was added to detect any subtle growth enhancements that might arise at a low urea concentration. Yeast expressing the MG0527 Empty vector exhibited a urea dose-dependent growth response, with 12mM urea increasing growth to by 110% compared to when grown in 'no urea' media (Figure 4.5). The dose-dependent growth responses for 6 of the 9 NtAQPs tested (NtPIP1;1t, NtPIP1;5s, NtPIP1;3t, NtPIP2;4s, NtPIP2;5t, NtNIP5;1t), were similar to the *ynvwI* yeast expressing the Empty vector (illustrated by the "Average low permeability", red line in Figure 4.5). As such, we assign these AQPs as being "urea impermeable" against which urea-permeable AQPs could be assessed. The expression of NtTIP1;1s, NtNIP2;1s and NtTIP2;5t increased plasma membrane permeability to urea, resulting in a 50% growth advantage at both 2 and 4mM Urea, compared to expression of urea-impermeable AQPs in *ynvwI* yeast (Figure 4.5).

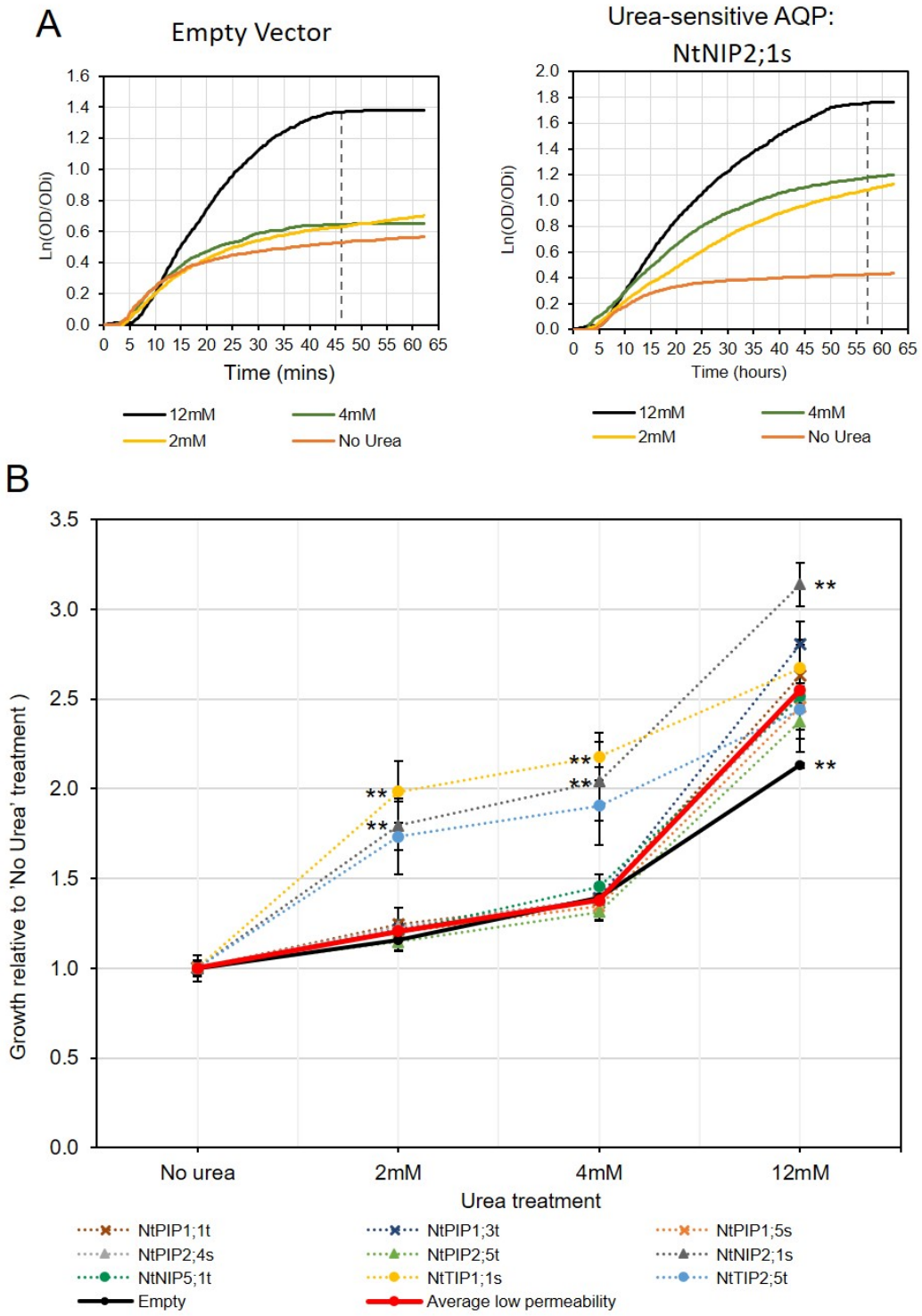


Figure 4. 5 **Urea permeability assays for yeast expressing NtAQPs . . A.** Yeast growth curves, Ln(OD/OD$_i$) vs. time, of *ynvwI* yeast expressing MG0515 Empty vector control or a urea-sensitive AQP (NtNIP2;1s), exposed to no Urea or 2mM, 4mM and 12mM urea treatments. **B.** Yeast culture growth relative to "0mM Urea" control of *ynvwI* yeast expressing either an Empty vector or one of the 9 NtAQPs. "Average low permeability", red line, indicates the average value of 6 AQPs which were classed as urea impermeable (PIP1;1t, NtPIP1;3t, , NtPIP1;5s, NtPIP2;4s, NtPIP2;5s, NtNIP5;1t).

Asterisks denote growth that was significantly greater or less than "Average low permeability" using a Student's t-test, "*" indicates $p<0.05$ and "**" indicates $p<0.01$, N=6, Error bars=SE.

4.3.6 In planta *sub-cellular localisation of tobacco AQPs*

In addition to assigning NtAQP substrate specificities, we screened for diversity in subcellular localisation using *Arabidopsis thaliana* as an *in planta* expression system. Confocal images of root cortical cells from Arabidopsis plants expressing GFP:NtAQP constructs were obtained (Figure 4.6). To enhance interpretation, surface plots of a region of GFP intensity near the cell wall are shown at greater magnification (indicated by white dashed box).

In Chapter 2, Figure 2.4, we utilised organelle-specific marker lines, established in Nelson et al. (2007), to guide our interpretation of AQP subcellular localisations. The cellular localisation results of the NtAQPs characterised in this chapter were concordant with those described in Chapter 2, Figure 2.4. We observed diversity in AQP membrane integration across the PIP, TIP and NIP subfamilies. The PIPs localised to the plasma membrane (PM), with the PIP1s (PIP1;1t, PIP1;3t and PIP1;5s) appearing to have a weaker and more diffuse PM integration when compared to the PIP2s (PIP2;4s and PIP2;5t) which had a sharp and defined GFP signal around the cell's periphery (Figure 4.6).

The NtNIPs (NIP2;1s and NIP5;1t) also localised to the cell's periphery. However, their GFP signal was speckled in appearance, with distinct localised spots of brighter fluorescence. This pattern was highlighted in the surface plots, which portray a wider spread in GFP signal, arising from adjacent PM and ER. The ER contributes the bright spots/peaks in fluorescence (indicated by white arrow on NIP2;1s surface plot profile of GFP intensity, Figure 4.6). The localisation of TIPs (NtTIP1;1s and NtTIP2;5t) was consistent with integration in the tonoplast, described in Chapter 2, showing a uniform yet diffuse localisation with a wavy topology; also denoted by the presence of internal membranes resembling transvacuolar strands (V, Figure 4.6).

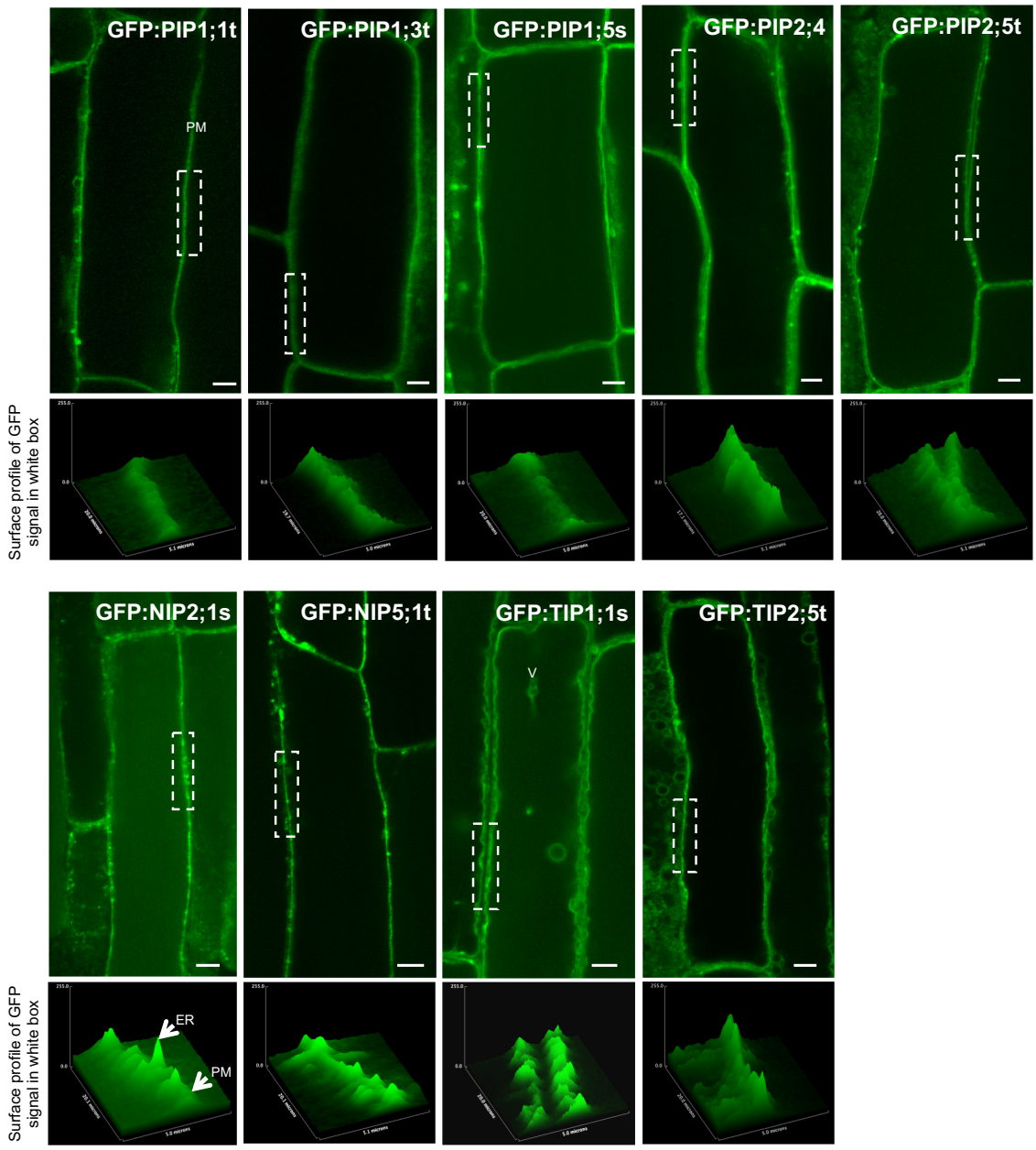


Figure 4. 6 . In planta sub-cellular localisation of NtAQPs. Confocal images of root cortical cells of transgenic 8-day old Arabidopsis seedlings expressing GFP translational fusions with PIP (NtPIP1;1t, NtPIP1;3t, NtPIP1;5s, NtPIP2;4s and NtPIP2;5t); TIP (NtTIP1;1s and NtTIP2;5t) and NIP (NtNIP2;1s and NtNIP5;1t) AQPs. Surface profiles of a region of the membrane (indicated by white dashed boxes, 5µm x 20µm dimension) is magnified in the panel below each confocal image. Plasma membrane and transvacuolar strands are denoted by PM and V respectively. Highlighted by the white arrows are peak intensity discrepancies present in the NIPs assigned to AQP integration into the ER and PM. Scale bar 5µm.

4.3.7 Summary of functional characterisation and sub-cellular localisation of NtAQPs

In order to characterise the function of each NtAQP, a combination of experimental approaches and sources of information have been combined. Firstly, it was necessary to establish that the expressed AQP was successfully incorporated in the plasma membrane of yeast. Then, the effect of each AQP on altering the permeability of the plasma membrane to four substrates was determined. *In planta* sub-cellular localisation in Arabidopsis root cortical cells and gene expression localisation are combined with permeability data to form a comprehensive array of information for these AQPs (Table 4.6).

Overall we saw similarities in substrate specificities and sub-cellular localisation within AQP subfamilies. The PIPs all localised to the PM, with the PIP2s having a more pronounced PM integration compared to the PIP1s. Our functional studies identified that NtPIP2s improved yeast survivorship after freeze-thaw treatments (water permeability assay), increased susceptibility to H_2O_2 exposure, and did not alter the response to boric acid or urea. These results suggest that NtPIP2s enhance plasma membrane permeability to both water and H_2O_2, but are impermeable to boric acid and urea. By contrast, the PIP1s were impermeable to water, but NtPIP1;1t enhanced plasma membrane permeability to H_2O_2 and NtPIP1;5s enhanced permeability to boric acid. None of the PIP1s were permeable to urea.

For the NIPs, GFP tagging demonstrated that both NtNIP2;1s and NtNIP5;1s were incorporated into the plasma membrane as well as accumulating in the endoplasmic reticulum. While both enhanced permeability to boric acid, neither altered the permeability to water or H_2O_2. NtNIP2;1s also enhanced permeability to urea. (Table 4.6).

The NtTIPs (NtTIP1;1s and NtTIP2;5t) localised to the plant cell tonoplast. When NtTIPs were expressed in yeast cells we observed an altered localisation, also incorporating in the plasma membrane and ER of yeast cells (Figure 4.1). Improved survival following freeze-thaw treatment suggests they both enhanced the

permeability of the plasma membrane to water. Their expression in yeast also increased plasma membrane permeability to boric acid and urea.

The gene expression patterns differed between members of the PIP1, PIP2 and TIP subfamilies. Both PIP1s and PIP2s show diverse expression throughout the plant as do the TIPs. However, expression of the two NIPs is restricted to the flowers.

Table 4. 6 Results summary of NtAQPs tested for water, H_2O_2, boric acid and urea permeability; in planta sub-cellular localisations and gene expression localisations (reported in Chapter 2).
Red tick (✓) indicates a positive permeability result for a specific substrate. The number of ticks represent the magnitude of phenotypic response observed, with one, two and three ticks representing a small, medium and high phenotypic growth alteration relative to the other NtAQPs, respectively. Sub-cellular localisations for the NtAQPs tested include plasma membrane (PM), endoplasmic reticulum (ER) and tonoplast.

		Water	H_2O_2	Boric Acid	Urea	Sub-cellular localisation	Expression localisation
PIPs	PIP1;1t	-	✓	-	-	PM	Roots
	PIP1;3t	-	-	-	-	PM	Whole plant
	PIP1;5s	-	-	✓	-	PM	Leaves, stem, roots, flowers
	PIP2;4s	✓✓	✓✓✓	-	-	PM	Roots, flowers (low)
	PIP2;5t	✓✓	✓✓	-	-	PM	Leaves
NIPs	NIP2;1s	-	-	✓	✓✓	ER + PM	Young flowers
	NIP5;1t	-	-	✓	-	ER + PM	Young flowers
TIPs	TIP1;1s	✓✓	-	✓✓	✓	Tonoplast	Leaves, flowers
	TIP2;5t	✓	-	✓✓	✓	Tonoplast	Roots, leaves (low), flowers (low)

4.3.8 Protein modelling of aquaporin pores

Tertiary structure homology modelling was used to compare pore width and the physico-chemical properties between the 9 NtAQPs functionally characterised in this chapter. In order to evaluate NtAQP monomer characteristics, we firstly conducted preliminary analyses to guide our selection of which template crystal structure (either Spinach PIP2;1, SoPIP2;1, or Arabidopsis TIP2;1, AtTIP2;1) was most suitable for generating 3D models of the NtAQP isoforms. Assessment of template suitability consisted of comparing model quality scores (GMQE scores) generated through the SWISS-MODEL software , with higher scores indicating better quality of the model (Table 4.7), as well as comparing pore radius profiles of 3D models generated from each template (Figure 4.7). AQP regions known to be of importance for substrate selectivity and structural integrity, such as the Selectivity Filter (SF) and the NPA regions, were also annotated on the pore radius profiles. The NtNIP5;1t and NtNIP2;1s pore radius profiles were plotted separately due to observed variations in pore radius between in these two isoforms (Figure 4.7).

Overall, we saw that regardless of which crystal structure template was used (either SoPIP2;1 or AtTIP2;1) for modelling the NtTIP, NtNIP2;1s and NtNIP5;1t isoforms, the pore radius profiles of the two 3D models generated were not very different. Additionally, it is interesting to note that in modelling the NtTIPs, although the sequence identity decreases when mapping isoforms from different sub-families (~80-60% for NtTIP isoforms modelled to AtTIP2;1 models, vs. ~35-39% sequence ID when modelling NtTIP isoforms to SoPIP2;1, Table 4.7), general trends of the radius profiles were not affected. Homology in pore radius profiles between the two generated models suggests that the variation we see across the different isoforms is gene specific and not due to variations resulting from the selected model. However for the PIPs we observed differences in the pore radius profiles modelled to either the SoPIP2;1 or AtTIP2;1 3D structures, suggesting that in modelling the PIPs, it is prefereable to use a 3D model from the same subfamily.

For the tobacco PIP and TIP isoform homology models, we decided to use the SoPIP2;1 and AtTIP2;1 crystal structures, respectively, as templates. This was due to the higher sequence identity and GMQE scores (Table 4.7), and the consideration

that modelling AQP isoforms to templates of the same AQP subfamily would result in a more accurate 3D homology model. In modelling the NtNIP isforms, we were constrained in using crystal structure templates which had overall lower sequence identity (~26-35% seq ID) and model quality scores (0.58-0.7 GMQE scores), belonging to different AQP subfamilies. For NtNIP5;1t, we chose the 3D model generated based on homology to the AtTIP2;1 template, due to the higher sequence identity to the AtTIP2;1 template and GMQE score. Whereas for NtNIP2;1s, we selected the model generated based on homology to the SoPIP2;1 template, even though the sequence identity and GMQE scores were slightly lower that those of AtTIP2;1 (Table 4.7). We found that key pore lining residue of the NPA1 region was not present in the NtNIP2;1s model generated from the AtTIP2;1 crystal structure, indicating that there was a lower resolution of of the 3D model in a region of importance for our analyses.

Table 4. 7. SWISS-MODEL parameters for 3D protein homology modeling of NtAQPs. *For each gene, the crystal structure templates used (SoPIP2;1 Open conformation PDB ID: 2b5f.1.A or AtTIP2;1 PDB ID: 5i32.1.A), the truncations used to align each sequence against the template (amino acids ,aa, truncated from N-terminus and C-terminus), sequence identity (Seq ID %) against the template and the GMQE scores for each NtAQP 3D model generated are listed. Outputs highlighted in red indicate the 3D model selected for subsequent characterisation.*

Gene name	Truncation against template alignments	Template used	Seq ID %	GMQE
PIP1;1t (*truncated*)	41 aa N-term 10 aa C-term	**SoPIP2;1 OPEN**	76.27%	0.94
		AtTIP2;1	34.82%	0.72
PIP1;3t (*truncated*)	42 aa N-term 10 aa C-term	**SoPIP2;1 OPEN**	76.69%	0.94
		AtTIP2;1	35.71%	0.72
PIP1;5s (*truncated*)	41 aa N-term 10 aa C-term	**SoPIP2;1 OPEN**	75.45%	0.97
		AtTIP2;1	35.27%	0.71
PIP2;4s (*truncated*)	29 aa N-term 16 aa C-term	**SoPIP2;1 OPEN**	78.57%	0.81
		AtTIP2;1	37.99%	0.69
PIP2;5t (*truncated*)	27 aa N-term 16 aa C-term	**SoPIP2;1 OPEN**	80.93%	0.97
		AtTIP2;1	45.28%	0.70
TIP1;1s (*truncated*)	5 aa N-term	**SoPIP2;1 OPEN**	39.32%	0.68
		AtTIP2;1	57.32%	0.80
TIP2;5t	none	**SoPIP2;1 OPEN**	38.89%	0.70
		AtTIP2;1	84.49%	0.93
NIP2;1s (*truncated*)	26 aa N-term 17 aa C-term	**SoPIP2;1 OPEN**	26.8%	0.64
		AtTIP2;1	35.53%	0.70
NIP5;1t (*truncated*)	23 aa N-term 2 aa C-term	**SoPIP2;1 OPEN**	27.60%	0.58
		AtTIP2;1	35.5%	0.64

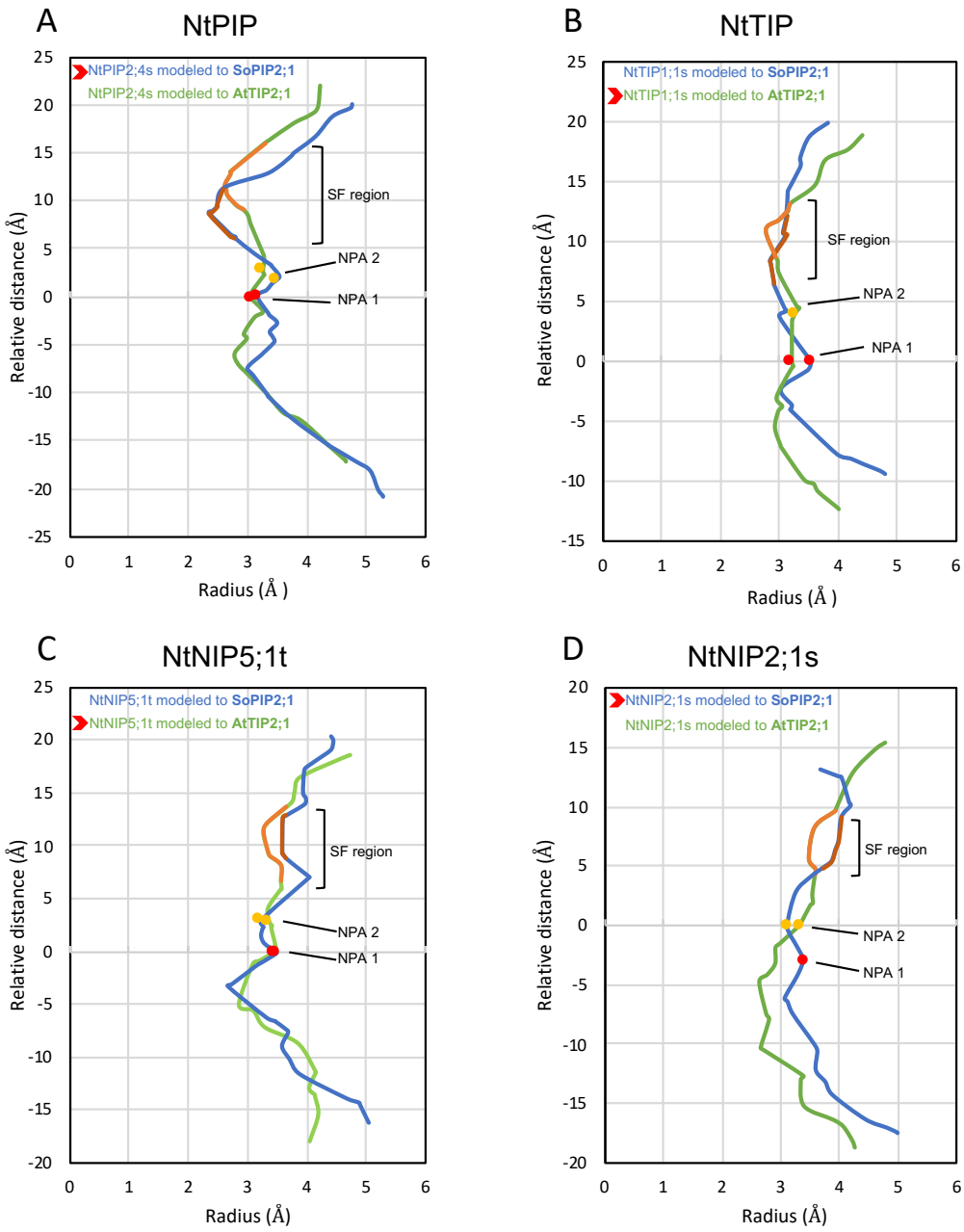


Figure 4. 7. Pore profiles of NtNIP5;1t, NtNIP2;1s, NtPIP and NtTIP isoforms modelled to either Spinach PIP2;1 or Arabidopsis TIP2;1 crystal structure templates*. Pore radius profiles of 3D monomer structures for* ***A.*** *a representative NtPIP (NtPIP2;4s) ,* ***B.*** *a represenatative NtTIP (NtTIP1;1s),* ***C.*** *NtNIP5;1t and* ***D.*** *NtNIP2;1s isoforms modeled to both Spinach PIP2;1 (blue, PDB:2b5f.1.A) and Arabidopdis TIP2;1 (green PDB:5i32.1.A) crystal structures. Also labeled are the Selectivity filter (SF, dark and light orange), the NPA 1 (red) and NPA 2 (yellow) regions which could be identified in the pore contact residues. . Relative distance was standardised to the NPA1 region in for each model, with the exception of the pore raius profiles of the NtNIP2;1s models, which were standardised to NPA2. Red arrow next to each legend indicates the 3D homology model which was selected for further functional characterisations.*

Once the highest-confidence 3D homology model was selected for each of our NtAQP isoforms, we could compare pore radius profiles and physico-chemical characteristics of AQPs across the different subfamilies (Figure 4.8-4.9). The pore radius profile overlaying all NtAQP models (Figure 4.8a), illustrates homology in pore radius across the 5 NtPIPs characterised (blue line, Figure 4.8a), with the SF region being the narrowest point along the pore (radius of 2.4Å). The SF residue composition of the NtPIPs is conserved, having Phe-His-Thr-Arg composition in Helix 2 (H2), Helix 5 (H5), Loop E position 1 (LE1) and Loop E position 2 (LE2), respectively (Figure 4.8b). The TIPs also have a conserved pore radius profile, with the SF region being the narrowest point/ bottle neck (dark and light purple lines, Figure 4.8a). However, the pore shape of the NtTIPs was less undulating in comparison to the PIPs, with the SF region being wider (2.75 Å radius) and the NPA region being slightly narrower (4.26 Å radius in TIPs vs. the 4.48 Å of the PIPs; Figure 4.7a,b). The SF residue composition differed between the 2 NtTIPs; with NtTIP1;1s having His- Ile- **Ala**- **Val** vs. NtTIP2;5t having His-Ile-**Gly**-**Arg** at H2-H5-LE1-LE2 positions (Figure 4.8b). The NtNIPs characterised showed some variation in pore radius profiles; with NtNIP2;1s having a SF radius ranging from **3.5-4** Å, vs. NtNIP5;1t's SF radius of **2.6-3.5** Å (Figure 4.8a). The remainder of the pore toward the cytosolic side was similar in profile between these two genes. The differences in SF pore width between the NtNIPs were matched with variation in residue composition in this region, with NtNIP2;1s having SF residues: **Gly**-**Ser**-**Gly**-Arg and NtNIP5;1t: **Ala**-**Ile-Ala**-Arg at H2-H5-LE1-LE2 positions respectively. The small size of the residues at the H2 and H5 position (Ser and Gly, respectively) in NtNIP2;1s, contributing to the larger pore diameter at this constriction site (Figure 4.8b). The NPA motifs composition (NPA1 and NPA2 motifs) is conserved across all of the NtPIPs, NtTIPs and NtNIP2;1s, with the exception of NtNIP5;1t which has NPS at NPA1 and NPV at NPA2 (Figure 4.8c).

As well as comparing the profiles of each pore radius and residues for the SF and NPA motifs, we characterised physico-chemical properties of the NtAQP pores (Figure 4.9). The NtPIPs showed high homology in residue hydrophobicity and flexibility (Blue box, Figure 4.9). Conserved regions with hydrophylic residues are

present across the PIP1s and PIP2s, occurring around the narrowest part of the pore (the SF region) and towards the cytosolic pore mouth; with the pores' centre being lined with hydrophobic residues. The flexibility of the PIPs' pore lining residues was also conserved; having a localised band of low-flexibility residues in their SF constriction region and the remainder of the pore having medium-high flexibility residues (blue box, Figure 4.9). The NtTIPs pore profiles highlighted their more cylindrical and less hourglass shape, lined with mostly hydrophobic and less flexible residues (purple box, Figure 4.9). Overall, the NtNIPs' pores were lined with more hydrophobic residues compared to the NtPIPs, and had hydrophilic regions towards the apoplastic pore entrance (green box, Figure 4.9). Some differences in residue flexibility were observed between the 2 NIPs. Pore lining residues in NtNIP5;1t were more flexible than in NtNIP2;1, with the exception of the central bottleneck region where both had low flexibility. Also, the apoplastic pore entrance of NtNIP2;1 was very flexible (green box, Figure 4.9).

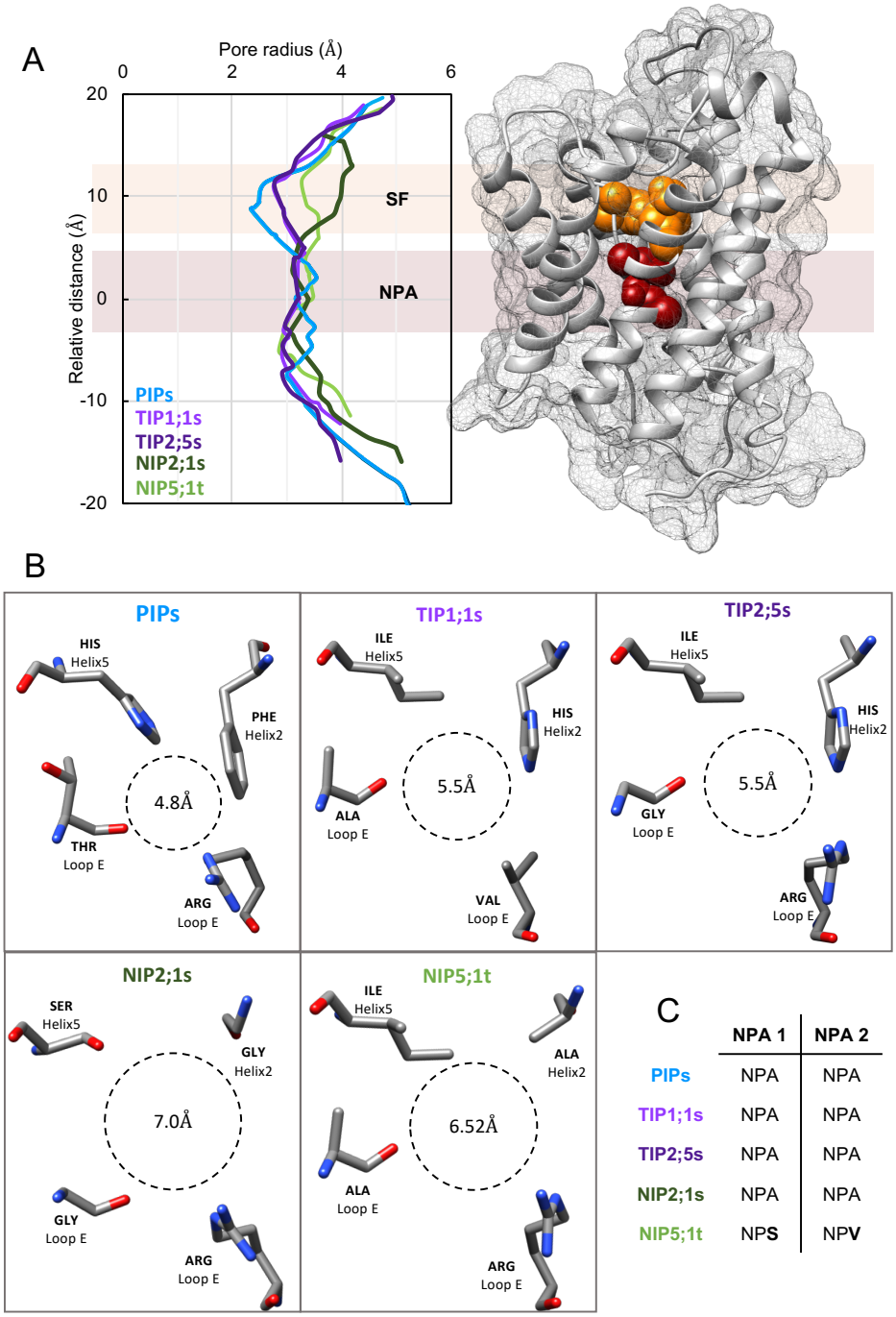


Figure 4. 8 . Modelled NtAQP pore features. **A.** Pore profiles of PIPs (PIP1;1 t, PIP1;3t, PIP1;5s, PIP2;4s, PIP2;5t; blue), TIP1;1s (light purple), TIP2;5t (dark purple), NIP2;1s (dark green) and NIP5;1t (light green). In the 3D protein model, the Selectivity filter region (SF, orange residue in 3D Protein model) and NPA region (dark red residues in 3D protein model) are highlighted. **B.** Amino acid residues forming the selectivity filter and its diameter at the narrowest point of the SF filter, viewed perpendicular to the membrane plane from the extracellular side. **C.** NPA motifs: NPA 1 and NPA 2 composition of PIPs, TIP1;1s, TIP2;5t, NIP2;1s and NIP5;1t.

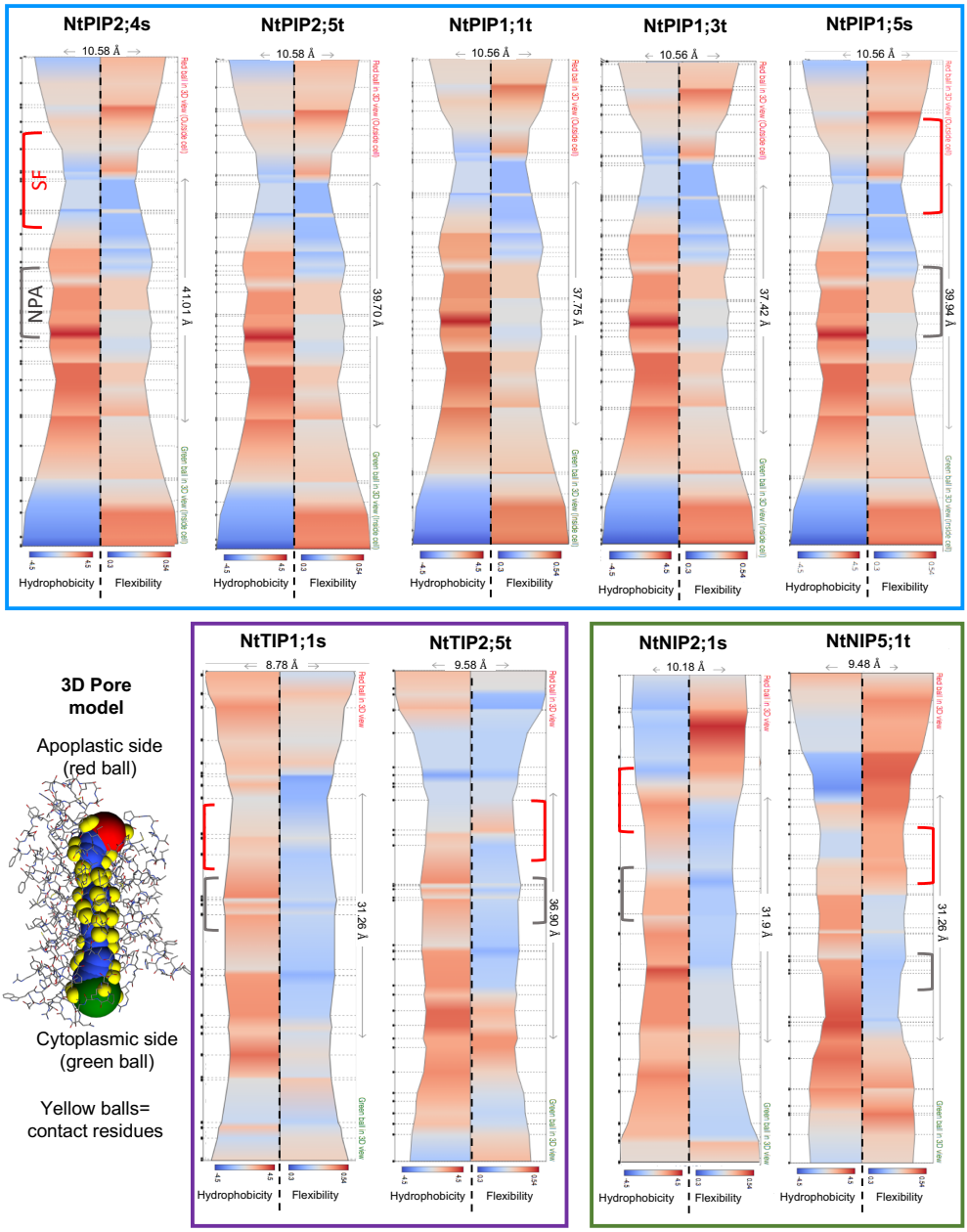


Figure 4. 9 . Hydrophobicity and flexibility profiles of NtAQP pores. Pore profile output from ChexVis software, illustrating Hydrophobicity (Left hand, Blue to Red indicating low to high hydrophobicity, respectively) and Flexibility (Right hand, Blue to Red indicating low to high flexibility, respectively) properties of NtPIP1;1t, NtPIP1;3t, NtPIP1;5s, NtPIP2;4s, NtPIP2;5t, NtNIP2;1s, NtNIP5;1t, NtTIP1;1s and NtTIP25t. Each pore profile includes maximum pore diameter and the measured pore length. Black dots to the left of each pore profile correspond to contact residue interactions for which hydrophobicity and flexibility outputs were generated. Red brackets indicate Selectivity Filter (SF) region and grey brackets indicate NPA region. 3D Pore model illustrates orientation of pore profile, top (red ball) being apoplastic side and bottom (green ball) being the cytosolic side. Yellow balls indicate the contact residues and bue balls represent the area inside the AQP pore.

4.4 Discussion

4.4.1 Confirming NtAQP integration into the plasma membrane of yeast

Using the yeast functional assays established in our laboratory, we were able to screen 9 NtAQPs for water, H_2O_2, boric acid and urea substrate permeabilities. Prior to conducting the functional assays, we visualised NtAQP subcellular localisation in yeast through GFP translational fusions. We observed some inherent homology in cellular localisation between plants and yeast cells, such as TIP isoforms which localised to the tonoplast in both plant (Figure 4.6) and yeast cells (Figure 4.1). However, an advantage of using yeast for AQP functional assays is that all AQP isoforms integrated into the plasma membrane (Figure 4.1). This confirmation enabled us to effectively interpret permeability results arising from yeast functional assays. If an AQP was able to successfully integrate into the yeast cell plasma membrane, we could infer that growth alterations upon exposure to specific substrates were due to AQP-related substrate specificities (Bienert and Chaumont 2014). Similarly, if an AQP was unable to integrate into a yeast plasma membrane, we could not assign a 'negative' permeability result to that AQP as it would not be in contact with the cell's exterior to permeate substrates present in the growth media. We therefore propose that confirmation of plasma membrane integration should be a necessary checkpoint when functionally testing AQPs in heterologous expression systems. We also suggest that although GFP translational fusions enable the visualisation AQP subcellular localisation in yeast cells, that functional assays in yeast should not be conducted with GFP translational fusions due to the resulting interference on AQP function (as seen in Chapter 3, Figure 3.9).

4.4.2 Phenotypic growth differences of yeast expressing Empty vector control vs. 'non-permeable' AQPs

Our functional study on NtAQPs from various subfamilies reported a range of yeast growth phenotypes in response stress-inducing or growth-enhancing treatments. Relative differences in growth between specific AQPs provides context to potential growth impairments or enhancements resulting from their heterologous expression in yeast. In several cases, phenotypes of yeast expressing

the Empty vector control differed from those expressing an AQP which was 'impermeable' to that particular substrate.

For the H_2O_2, boric acid and urea assays, yeast cultures were exposed to their respective treatments for the entire duration of the experiment. In the H_2O_2 and boric acid permeability assays, we observed slight differences in growth between the yeast expressing the Empty vector control and AQPs impermeable to the respective substrates tested ('Average low permeability' red line in Figures 4.3-4.5).

Such differences include those observed in the H_2O_2 assay, where 6 of the 9 NtAQP cultures consistently showed 5-10% less growth than Empty vector when exposed to increasing H_2O_2 concentrations. This phenotype has been previously encountered in studies characterising AQPs permeating H_2O_2 in yeast, where it was found that aquaglyceroporin from the malaria parasite (*Plasmodium falciparum*), PfAQP1, had a slight but consistent reduction in growth compared to non-AQP expressing yeast upon exposure to increasing H_2O_2 treatments. This phenotype was in contrast to orthodox water transporting AQPs which had high water and H_2O_2 permeability (Almasalmeh et al. 2014). It was hypothesised that this phenotype was due to an intrinsic property of all AQPs to have some degree of H_2O_2 permeability, with orthodox AQPs having a higher H_2O_2 conductance (Almasalmeh et al. 2014). Therefore, reduced growth at various H_2O_2 treatment concentration of the NtAQPs compared to the Empty vector control suggests they all have some degree of H_2O_2 permeability. It could be possible that an ubiquitous background-level permeation of H_2O_2 through AQPs could have a negligible physiological significance in plants due to the low magnitude. To better detect AQPs which are H_2O_2-permeable above potentially background/intrinsic levels, we attributed H_2O_2 permeability to the NtAQPs which had a statistically significant reduction in growth compared to the 'calculated-negative' values, rather than to yeast expressing the Empty vector control (Table 4.6).

In contrast to the H_2O_2 assay, a "shielding" of the toxicity effect was observed in the boric acid assay, with expression of boric acid impermeable AQPs (4 of the 9 NtAQPs) exhibiting a reduced toxicity reponse compared to yeast expressing Empty vector control. Variation in the growth phenotypes observed in the boric acid functional assay (relative to Empty vector expression) could potentially be due to a

reduced area in the plasma membrane for free diffusion. The over expression of a foreign membrane protein (impermeable to that particular substrate) in the yeast cell could be occluding the cell periphery, and as such less boric acid is able to freely diffuse across the membrane, reducing the phenotypic response at increasing treatment concentrations compared to empty vector control.

4.4.3 Functional characterisation of NtPIPs

The PIP subfamily in plants generally has the largest number of isoforms. All PIPs exhibit a narrow pore structure typical of water selective AQPs (Maurel et al. 2008). PIPs have 2 major phylogenetic groups, PIP1s and PIP2s, from which we selected 3 PIP1 and 2 PIP2 isoforms to characterise. Differences in substrate specificities were observed within these phylogenetic subgroups, concordant with broad literature (Maurel et al. 2015; Otto et al. 2010). Our results show that both PIP2s tested were permeable to water and H_2O_2, whereas no PIP1s were water-permeable. However, PIP1;1t and PIP1;5s conferred some permeability to H_2O_2 and boric acid, respectively. While we were unable to detect any enhanced permeability to the four substrates associated with the expression of NtPIP1;3t, it is possible that NtPIP1;3t could permeate other substrates such as CO_2, due to it's high homology to NtAQP1/NtPIP1;5s (an established CO_2–permeable AQP).

Although the substrate selectivity we observed differed between the NtPIPs, they all had the same SF filters (matching that of water-permeable AQPs), NPA motifs, pore radius profile and they had highly homologous physico-chemical properties of pore-lining residues (hydrophobicity and flexibility). These results highlight the challenges in predicting PIP substrate selectivity based on primary sequence and 3D pore structure and suggest that pore lining residues externally to the SF and NPA regions and/or domains regulating pore dynamics (i.e. open/closed state) also influence selectivity.

In addition to the yeast assay results, the *in planta* subcellular localisation results of the NtPIPs also show some slight variations between the PIP1 and PIP2s, with the PIP2s appearing to have a sharper integration in the plasma membrane. Studies have previously reported variation in sub-cellular localisation between

these two PIP sub-groups, with PIP1 interacting with PIP2s in order to more efficiently integrate into the PM through hetero-tetramer formation (Bienert and Chaumont 2014; Otto et al. 2010). The localisation of PIP1s appeared somewhat more diffuse and the GFP signal less bright, potentially reflective of their need to interact with PIP2s in order to efficiently integrate in the PM. Future experiments co-expressing fluorescent marker tagged PIP1 and PIP2 isoforms could investigate whether PIP1 integration in the plant cell membrane improves upon interaction with PIP2 isoforms.

Combining our functional, subcellular localisation and tissue specific expression analysis, can help elucidate possible physiological roles for these NtAQPs. The results implicate NtPIP2;4s and NtPIP2;5t as having a role in regulating water transport across cell membranes in roots and leaves, respectively. Furthermore, their dual specificity for H_2O_2 suggests a role in H_2O_2 diffusion between cells, for example, during stress responses through ROS signalling (Hachez et al. 2006). Since NtPIP1;1s is expressed in roots and is permeable to H_2O_2, it could also be associated with ROS signalling in response to plant roots experiencing stresses. NtPIP1;5s (also known as NtAQP1) has been previously studied in plants, being the first plant AQP shown to permeate CO_2 (Uehlein et al. 2003) and as such had been associated with photosynthesis through facilitating diffusion of CO_2 into the chloroplast (Flexas et al. 2006). Our yeast functional experiments further implicate NtPIP1;5s in boric acid transport. Being a highly expressed gene throughout the whole plant, it is likely that NtPIP1;5s is involved in a variety of plant processes. Because NtPIP1;1t and NtPIP1;3t are highly homologous to NtPIP1;5s (permeable to CO_2), the potential for NtPIP1;1t and NtPIP1;3t to influence CO_2 permeability in leaf mesophyll cells *in planta* is examined in Chapter 5.

4.4.4 Functional characterisation of NtTIPs

The TIPs sub-family has important physiological roles within plants, regulating the diffusion of water, ammonia, urea and metalloids across the tonoplast (Maurel et al. 1993; Loqué et al. 2005; Liu et al. 2003). Five specialised TIP subgroups have evolved in higher plants (TIP1-TIP5), differing in ar/R filter (SF) composition and substrate specificities (Anderberg et al. 2012; Kirscht et al. 2016).

In this study we characterised a TIP1 (*NtTIP1;1s*) and a TIP2 (*NtTIP2;5t)*; and both were found to increase permeability to water, boric acid and urea. Their SF region was found to be wider than that of the NtPIPs, which correlated with the passage of a larger substrate such as urea. An extended selectivity filter has been characterised for the TIP subfamily, containing an additional contact residue in the Loop C of the AQP monomer (Kirscht et al. 2016), with NtTIP1;1s and NtTIP2;5t having a Phe and His at this position, respectively (Listed in Chapter 2, Table 2.2).The NtTIP1;1 SF composition of a Phe in Loop C and a Val in Loop E2 creates a more hydrophobic environment in the SF compared to that of NtTIP2;5s which as His and Arg in the same two postions respectively. This difference in SF filter composition is reflected in the pore profile with TIP1;1s being more hydrophobic than TIP2;5t in its constriction region (purple box, Figure 4.9). Although the SF composition and hydrophobic properties varied between these two genes, they had identical substrate specificities. It is possible that these SF composition differences might affect permeation of other substrates which were not tested for in this study, e.g. ammonia (Kirscht et al. 2016). NtTIP1;1s is expressed in leaves and flowers whereas NtTIP2;5t is predominantly expressed in the roots (having low expression in leaves and flowers). Their proposed functional roles include the unloading of urea from vacuolar storage, the translocation of boron and equilibration of water in their respective tissues, concordant with characterised roles assigned to AQPs in the TIP subfamily (Maurel et al. 2015).

4.4.5 Functional characterisation of NtNIPs

NIP aquaporins are known to facilitate the transport of small uncharged solutes, such as glycerol, urea and metalloids (Wallace et al. 2006). They can be divided into 3 sub-classes (NIP I-III), based on ar/R selectivity filter and NPA motifs composition (Mitani et al. 2008). The NtNIPs characterised in this study were *NtNIP2;1s* and *NtNIP5;1t*, belonging to the NIP III and NIP II sub-classes, respectively.

Overall, NIPs have a more hydrophobic ar/R selectivity filter, which has been linked to their poorer water permeability and preference for other substrates such

as ammonia, urea and metalloids (Wu and Beitz 2007; Hove and Bhave 2011). The physico-chemical pore properties of the NIPs characterised in our study show an abundance of hydrophobic residues lining the pore (additionally to the ar/R filter) when compared to the NtPIPs, especially towards the cytosolic pore mouth (Figure 4.8).

NIP II aquaporins tend to have a larger pore diameter than those found in the NIP I sub class, having a substitution of the highly conserved and bulky Trp at the ar/R H2 position for a smaller Ala (Wallace and Roberts 2004). We also observed this characteristic in pore shape of NIP II AQPs, with NtNIP5;1, having a wider pore diameter compared to that of the tobacco PIP and TIP isoforms (Figure 4.8). NIP II aquaporins have been shown to permeate boric acid, glycerol and urea, and were described to be completely impermeable to water (Wallace et al. 2006; Takano et al. 2006; Hanaoka et al. 2014; Tanaka et al. 2008). Although we reported increased toxicity response to boric acid in yeast expressing NtNIP5;1t (concordant with the broad literature), we did not detect any improved growth of *ynvwI* yeast at low urea concentrations (Table 4.6). Since *NtNIP5;1t* expression is highly targeted to young flowers, this gene could be involved in boron redistribution during flower development . In support of this, the orthologous gene in Arabidopsis (AtNIP5;1) has an established role in boron transport and flower development (Takano et al. 2006).

NIP III aquaporins (such as NtNIP2;1s) are characterised by an ar/R filter composed of smaller residues (Gly-Ser-Gly-Arg), resulting in an even wider pore diameter than other NIP II sub-class (Bansal and Sankararamakrishnan 2007; Mitani-Ueno et al. 2011). Our 3D homology modelling results showed that NtNIP2;1s indeed has a wider pore with that all the other NtAQP isoforms characterised (Figure 4.8). NIP III aquaporins have been studied for their ability to permeate larger substrates, such as silicic acid (4.38 Å diameter), and additionally have been reported to permeate boric acid, urea and lactic acid (Mitani-Ueno et al. 2011). As the ar/R filter composition of NtNIP2;1s matches that of other characterised NIP III aquaporins, it is likely that it is able to permeate other substrates (such as silicilic acid) as well as boric acid and urea. As *NtNIP2;1* expression is localised to young flowers, it is likely that this NIP isoform is involved

in translocation of small molecules, such as urea and boric acid in developing flowers.

4.5 Conclusions

This Chapter presents a comprehensive characterisation of a diverse set of isoforms in the tobacco AQP family, screening for water, H_2O_2, urea and boric acid permeabilities. Elucidation of substrate specificities allowed us to incorporate gene expression data and *in planta* subcellular localisation results in order to propose functional roles for several PIP, TIP and NIP isoforms. Additionally, 3D protein homology modelling analyses enriched our characterisation these tobacco AQPs by linking pore size and physico-chemical properties of pore lining residues to the assigned substrate specificities of each isoform.

Our results can be used toward building a catalogue of AQPs with known subtrate specficities and assist with improving our understanding of AQP biology. Furthermore, by characterising diverse isoforms within the same species and by using a similar experimental set-up (yeast functional assays), we can better compare regions that might confer substrate specificity (avoiding intra-species variation) while reducing the probability of identifying false positive/negative permeabilities.

Chapter 5: *In planta functional characterisation of tobacco PIP isoforms*

5.1 Introduction

In order to engineer higher yielding crops, we need to further explore the role of CO_2-permeable AQPs in enhancing mesophyll conductance and photosynthetic efficiency. As such, in attempts to elucidate the role of PIP AQPs in facilitating CO_2 diffusion to the chloroplast, we measured photosynthesis and mesophyll conductance of tobacco plants Over Expressing (OE) PIP1 genes with high homology to the characterised CO_2 pore: *NtPIP1;5s/ NtAQP1* (Figure 5.1). NtPIP1;5s/NtAQP1 is an established CO_2 pore that has been shown to enhance mesophyll conductance and photosynthetic capacity *in planta* (Flexas et al. 2006). Studying AQP isoforms with high sequence identity to an NtPIP1;5s could help up tease apart potential differences in measured photosynthetic parameters, and identify residues that might confer AQP specificity to CO_2.

We previously identified 10 PIP1s in the tobacco genome, occurring in two distinct phylogenetic clades (Chapter 2). Three sister gene pairs occur in the clade containing *NtPIP1;5s*. As such, we selected one gene from each pair (*NtPIP1;1t*, *NtPIP1;3t* and *NtPIP1;5t*) to constitutively over-express in tobacco; testing whether these highly homologous genes could also enhance photosynthetic efficiency. In addition to these PIP1 genes, we also selected a more distantly related PIP2, *NtPIP2;4s*, to constitutively over-express in tobacco.

This chapter explores the effect of over expression of PIP1 genes highly homologous to PIP1;5s/NtAQP1 on tobacco photosynthesis. Unexpectedly, no increases in mesophyll conductance were observed in the PIP1 OE lines despite observed increases in photosynthetic rates. These findings suggest potential pleiotropic effects that might occur upon altered PIP expression *in planta*.

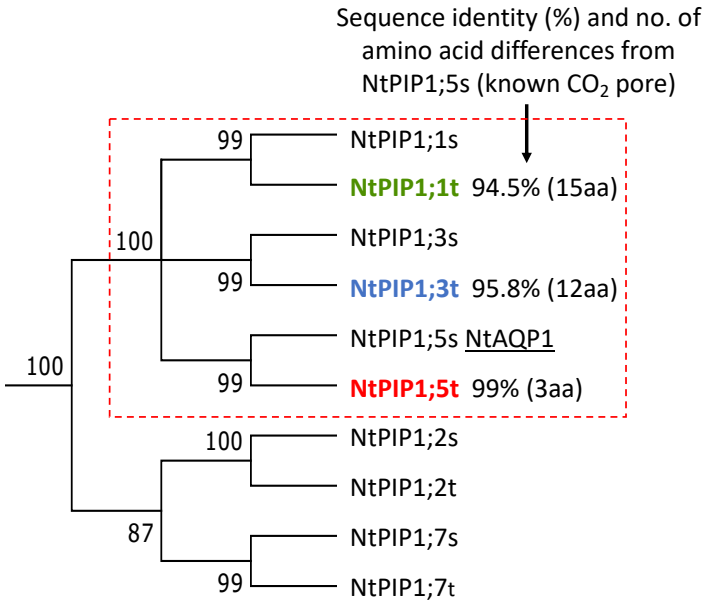


Figure 5. 1. Phylogenetic representation of PIP1 genes selected for in planta characterisation; and their homology to NtPIP1;5s (NtAQP1), an established CO_2 pore. *The Red dashed box highlights the sub-clade of NtPIP1 genes with high homology to NtPIP1;5s. NtPIP1;1t (green), NtPIP1;3t (blue) and NtPIP1;5t (red) were selected as candidates to over-express in tobacco, having 94.5%, 95.8% and 99% sequence identity to NtPIP1;5s respectively. Also included are the number of amino acid (aa) differences of each gene compared to NtPIP1;5s.*

5.2 Methods

5.2.1. Generation of tobacco PIP OE lines and plant growth

PIP aquaporin plant over-expression constructs were generated via gateway cloning of commercially synthesised coding sequences of *NtPIP1;1t* (BK11393), *NtPIP1;3t* (BK011396), *NtPIP1;5t* (BK011398) and *NtPIP2;4s* (BK011406) or the GUS protein (non-AQP expression control) from pZEO entry vectors into pMDC32 destination vector (Curtis and Grossniklaus 2003), driven by the constitutive 2x35S CaMV promoter (Figure S5.1).

Wild-type tobacco plants (cv. Petit Havana) were germinated for transformation on Murashige and Skoog (MS) agar media (4.43g MS, 1L water, 6g agar, adjusted to *pH* 5.7 with 1M KOH, autoclaved), grown in a 25°C temperature controlled room. Seedlings were gently transferred after 1 week to sterile plastic

pots (clear/opaque container, 1L volume) with clear screw-top lids, containing MS agar media with sucrose (4.43g MS, 30g sucrose, 1L water, 6g agar, adjusted to *pH* 5.7 with 1M KOH, autoclaved), grown in a 25^{o}C temperature controlled room for 3-4 weeks.

Agrobacterium carrying each transformation construct was grown for 48 hours at 28^{o}C in 20mL of LB broth, with Rifampicin (50mg/mL) and Kanamycin (50µg/ml) antibiotic selection, shaking at 220 rpm. Cells were harvested by centrifugation of the 20mL culture (in 50mL tubes) at 330 rpm for 10mins. Supernatant was discarded, pellets resuspended in 40mL Resuspension media (see Table 5.1), and incubated at room temperature for 1 hour. Young leaves of tobacco plants grown in the sterile 1L pots were excised and cut into 5mm x 5mm squares (48 leaf squares required per transformation construct). Leaf squares were placed on co-cultivation plates (see Table 5.1) with the upper epidermis side of leaf in contact with the media. 5mL of the Agrobacterium-suspension solution were added to each co-cultivation plate and incubated at room temperature for 5-10 minutes. Leaf squares were then transferred to fresh co-cultivation plates (6 leaf squares per plate, totalling 8 plates per transformation construct). Plates were sealed with cling wrap tape and incubated in the dark at 28^{o}C for 2 days. Leaf squares were transferred to regeneration plates (see Table 5.1) and kept in the dark for 2 weeks at 25^{o}C. Plates were then exposed to light and kept at 25^{o}C for a week. Calli forming around the edges of the leaf squares were transferred to fresh regeneration plates, and kept at 25^{o}C until the formation of adventitious buds. Plantlets were transferred to rooting media (see Table 5.1) when a clear shoot had formed.

Table 5. 1. Recipes for Resuspension, Co-cultivation, Regeneration and Rooting medias required for tobacco transformation. MES is 2-(N-morpholino)ethanesulfonic acid; BAP is 6-benzylaminopurine; NAA is 1-naphthaleneacetic acid, DMSO is dimethyl sulfoxide.

	Media type			
	Resuspension	***Co-cultivation***	***Regeneration***	***Rooting***
MS basal medium	4.43g	4.43g	4.43g	4.43g
Sucrose	30g	30g	30g	30g
MES	3.9g	0.6g	0.6g	0.6g
	pH to 5.7 with 1M KOH			
De-ionised water	949 mL	949mL	998ml	998mL
Agar	-	6g	6g	6g
	Filter sterilise	*Autoclave*	*Autoclave*	*Autoclave*
Acetosyringone (diluted in DMSO)	0.04g in 50mL	0.04g in 50mL	-	-
BAP (1mg/mL in 0.1M HCl)	1mL	1mL	1mL	-
NAA (1mg/mL in 0.1M NaOH)	0.1mL	0.1mL	0.1mL	-
Hygromycin (50mg/ml)	-	-	1mL	1mL
Timentin (50mg/mL)	-	-	1mL	1mL

Once roots were established, plantlets (15 per transformation construct) were transferred to 0.5L pots filled with commercial potting mix (Seedling raising mix) and 3 g L^{-1} slow release fertiliser (Osmocote). Pots were placed in a tray and covered with a clear plastic cover (for 3-4 days) to maintain high humidity, ensuring high survivorship of the T_0 transformants and grown in a growth chamber with a 16 hour photo period and day/light temperatures set to 25°C/22°C for 1-2 weeks. Transformants were transferred to larger well-draining 5 L pots, filled with commercial potting mix (Wizard potting mix) and 3 g L^{-1} slow release fertiliser (Osmocote), and grown in a naturally lit glasshouse with day/light temperatures set to 25°C. Plants were watered daily. Seeds were harvested and used for experimental analyses.

5.2.2 Quantification of transgene expression: Selection of OE lines

Seeds from 5 independent transformation events were sown out on MS agar media with hygromycin selection (1mL of 50mg/mL Hygromycin in 1L of media) for each transformed PIP OE construct; sowing out 50 seeds per line. Seedlings were grown on plates for 10 days at 25^{o}C. Segregation counts were collected to test for T-DNA insertions by counting number of seedlings resistant to hygromycin antibiotic selection (with 70-80% survivorship indicating single locus T-DNA insertion).

RNA extractions were conducted to test for transgene expression. The Bioline Plant RNA extraction kit was used, eluting in 50µl of RNAse-Free water. 5 whole tobacco seedlings (10 days old) were pooled for each RNA extraction; 3 replicates of RNA extractions were conducted per line. RNA was diluted to 100ng/µL and subsequently DNAse treated using the Invitrogen DNAseI kit. DNAse treatment consisted of adding 1000ng RNA (10µL of 100ng/µL RNA) to 1.28µL Nuclease-free water, 1.3µL 10x DNAseI buffer and 1µL DNAseI enzyme in a 250µL PCR tube. PCR tube was then incubated at 22^{o}C for 15 mins. 1.35µL of 25mM EDTA was added and solution was gently mixed via pipetting. The tube was then incubated at 65^{o}C for 10 minutes.

cDNA was synthesised using the Bioline SensiFAST cDNA synthesis kit, each reaction containing 4µL of 5x TransAmp buffer, 1µL of Reverse Transcriptase and 15µL of DNAse-treated RNA (1000mg total). Reaction temperature cycle was: 10mins at 25^{o}C, 15mins at 42^{o}C, 15mins at 48^{o}C, 5mins at 85^{o}C and subsequently held at 4^{o}C.

The tobacco ubiquitin-conjugating enzyme 2 (Ntubc2) was selected as a housekeeping gene to compare transgene expression. Primers used for qRT-PCR amplification of house-keeping gene were the Ntubc2 Forward: AGCTGCTATACTGACTTCAATCCA and Reverse: TCTCACTGAACATGCGTGCT primers. Primers used for qRT-PCR amplification of PIP OE constructs consisted of gene-specific forward primers and a universal NOS terminator 3' UTR reverse primer: GAAATTCGAGCTCCACCGC. The following gene-specific forward primers were used; NtPIP1;1t forward: GAGCCATTCCATTCAAGAGCA, NtPIP1;3t forward: GAGCCATTCCATTCAAGAGCAA, NtPIP1;5t forward: TCAGAGCCATTCCATTCCACA,

and NtPIP2;4s forward: GAGCAATGCCTAATACCCAGC. Primer annealing sites are illustrated in Supplementary figure S5.1.

qRT-PCR reactions were set up using the Bioline SensiFast SYBR LoROX kit. Each reaction consisted of 5µL of 2x SensiFAST SYBR LoROX Mix; 3.2µL Nuclease-free water; 0.4µL of Forward primer (gene specific primers listed above); 0.4µL of NOS terminator 3'UTR reverse primer and 1µL of template cDNA (1/10 dilution of cDNA synthesis product). qPCR reaction cycle consisted of Hold stage: 2 min 95^{o}C; PCR stage: 40 cycles of 5 sec at 95^{o}C, 10 min at 60^{o}C, 10 sec at 72^{o}C and Melt curve stage: 15 sec at 95 ^{o}C, 1 min at 60^{o}C and 15 sec and 95 ^{o}C. The ΔΔCt method was used to derive Fold change expression differences between house keeping gene and NtPIP transgene.

5.2.3 Plant growth for Gas exchange measurements

Tobacco PIP (*PIP1;1t*, *PIP1;3t*, *PIP1;5t* and *PIP2;4s*) OE lines (T_1 plants) were germinated on Hygromycin selective MS plates, grown in a constant temperature room set to 25^{o}C for 10 days. Seedlings were then then transferred to soil in 0.5L pots using commercial potting mix (Seedling raising mix) and 3 g L^{-1} slow release fertiliser (Osmocote). Seedlings were covered with clear plastic cover to maintain high humidity, ensuring high seedling survivorship and grown in a temperature controlled growth cabinet for 1 week (clear plastic cover removed after 3 days) with a 16 hour photo period and day/light temperatures set to 25^{o}C/22^{o}C. Plantlets were then transferred to larger well-draining 5 L pots, filled with commercial potting mix (Wizard potting mix) and 3 g L^{-1} slow release fertiliser (Osmocote), and grown in a naturally lit glasshouse with full sunlight (November-March, summer months) and day/light temperatures set to 25^{o}C. Plants were watered daily. Gas exchange and leaf measurements were conducted on 6-week old plants.

Three experimental batches were conducted, allowing for collection of photosynthetic measurements in a 3-5 day period per experimental batch.
The following lines were measured in each experimental batch; Experiment 1: PIP1;1t line 3 and Gus control plants; Experiment 2: PIP1;3t lines 2 and 4, PIP1;5t line 1 and 3 and Gus control plants ; Experiment 3: PIP2;4s lines 3 and 14 and GUS

control plants. Comparison of photosynthetic parameters were made between the PIP OE line and the GUS control measured in each respective experiment.

5.2.4 CO_2 response curves

Two LI-6400 gas exchange systems were used for concurrent measurement of CO_2 response curves. Each LI-6400 was set up with flow rate 400 µmol s^{-1}, Leaf temperature was 25°C and irradiance was 1500 µmol PAR photons $m^{-2}s^{-1}$. The LI-6400s were set to an AutoProgram collecting gas exchange measurements at the following reference cell CO_2 concentrations: 400, 40, 70, 100, 200, 300, 400, 600, 800, and 1000 (µmol CO_2 mol^{-1}). The last fully-expanded leaf of a 6-week old plant was used for measurement.

5.2.5 Concurrent gas exchange and carbon isotope discrimination measurements

Plants were collected from the glasshouse and transferred to a temperature-controlled cabinet set to 25°C. Two LI-6400 gas exchange systems were set up in the temperature controlled growth cabinet, with a Tunable Diode Laser (TDL; TGA100, Campbell Scientific, INC Logan) sampling the reference and sample gases from each system. The LI-6400 gas exchange systems were set up as following: flow rate 200 µmol s^{-1}, sample cell CO_2 concentration 380 µmol CO_2 mol^{-1}, Leaf temperature 25 °C and irradiance was 1500 µmol PAR photons $m^{-2}s^{-1}$. Synthetic air containing 2% Oxygen was used to reduce the carbon isotope fractionation associated with photorespiration (Tazoe et al. 2011).

The same leaf used for the CO_2 response curve measurements was gently clipped into the measuring head of a LI-6400 gas exchange systems (6 cm^2 chamber) with a red-blue light emitting diode (LED) light source. The two leaves were measured in sequence, with each LI-6400 sampled by the TDL at 4 min intervals, together with zero and calibration gases (Tazoe et al. 2011).

The carbon isotope composition ($\delta^{13}C$) was calculated from the equation presented in Evans *et al.* (1986), as:

$$\delta^{13}C = \left(\frac{^{13}CO_2 \,/\, ^{12}CO_2}{R_{VPDB}} - 1 \right) \text{x } 1000 \qquad \text{(Equation 1)}$$

where R_{VPDB} is the $^{13}CO_2/^{12}CO_2$ of the standard Vienna Pee Dee Belemnite (VPDB), which is 0.0111797 (Griffis et al. 2004).

The observed Delta (Δ_o) values were calculated for PIP OE and GUS control plants from the equation presented in Evans *et al.* (1986), as:

$$\Delta_o = \frac{1000 \times \xi\ (\delta^{13}C_{sam} - \delta^{13}C_{ref})}{1000 + \delta^{13}C_{sam}\ \xi\ (\delta^{13}C_{sam} - \delta^{13}C_{ref})} \qquad \text{(Equation 2)}$$

Where $\delta^{13}C_{sam}$ and $\delta^{13}C_{ref}$ are the carbon isotope compositions of CO_2 in the leaf chamber and reference air sampled from LI-6400 respectively; $\xi = C_{ref} / (C_{sam} - C_{ref})$; and C_{ref} and C_{sam} are the CO_2 concentrations measured by the TDL of dry air entering and leaving the leaf chamber, respectively (Evans et al. 1986). ξ values varied between 4-12.

5.2.6 Calculation of mesophyll conductance

In C_3 plants, the theoretical relationship between $^{13}CO_2$ discrimination, the CO_2 partial pressures in ambient air and within the leaf and respiratory and photorespiratory fluxes is (Evans et al. 1986; Farquhar and Richards 1984):

$$\Delta = a\frac{C_a - C_i}{C_a} + a_i\frac{C_i - C_c}{C_a} + b\frac{C_c}{C_a} - \frac{eR_d/k + f\Gamma*}{C_a} \qquad \text{(Equation 3)}$$

Where a is the fractionation occurring due to diffusion in air (4.4‰), a_i is fractionation factor for hydration and diffusion through water (1.8‰); b is the net fractionation for Rubisco and phospho*enol*pyruvate carboxylase (PEPC) carboxylation, 29‰ was assumed in this study, (Roeske and O'Leary 1984); C_a, C_i and C_c are the CO_2 partial pressures in ambient air, intercellular airspaces and sites of carboxylation in the chloroplasts, respectively. e and f are fractionations with respect to average carbon composition associated with 'dark' respiration (R_d) and photorespiration, respectively. k is carboxylation efficiency (Farquhar et al. 1982)

and Γ* is the CO_2 compensation point in absence of R_d, (Γ* value used was 37.43 μbar) (von Caemmerer et al. 1994).

By assuming infinite g_m and negligible fractionation associated with respiration and photorespiration, the net fractionation equation can be simplified to:

$$\Delta_i = a + (b - a)C_i/C_a \qquad \text{(Equation 4)}$$

The relative significance of terms a and b varies with C_i; when C_i is zero, term a dominates, whereas when C_i increases relative to C_a, the discrimination is dominated by the biochemical fractionation, term b (Evans and Von Caemmerer 2013).

We calculate mesophyll conductance (g_m) using the following equation from Evans and von Caemmerer (2013):

$$g_m = \frac{\frac{1+t}{1-t}\left(b - a_i - \frac{eR_d}{(A+R_d)}\right)\frac{A}{C_a}}{\left(\frac{1}{(1-t)}\,a' + \frac{1}{(1-t)}\left((1+t)b - a'\right)\frac{c_i}{c_a}\right) - \Delta_o - \frac{1+t}{1-t}\left(\frac{eR_d}{(A+R_d)C_a}(C_i - \Gamma_*)\right) - \frac{1+t}{1-t}\left(f\frac{\Gamma_*}{C_a}\right)}$$

(Equation 5)

where $t = \frac{(1+a')E}{2g^t_{ac}}$, E denotes transpiration rate and g^t_{ac} the total conductance to CO_2 diffusion including boundary layer and stomatal conductance, a' denotes the combined fractionation factor through the leaf boundary layer and through stomata, $a' = \frac{a_b(C_a - C_s) + a(C_s - C_i)}{(C_a - C_i)}$, (von Caemmerer and Farquhar 1981; Farquhar and Cernusak 2012) and A is the net assimilation.

After TDL and LI-6400 measurements were obtained, the measured leaves were cut for fresh weight and leaf area measurements. They were then dried in an oven for 5 days (50°C) and dry weight measurements were collected.

5.2.6 Calculation of V_{cmax} and J

Photosynthetic parameters for the Rubisco carboxylation capacity (V_{cmax}) and potential chloroplast electron transport rate (J) were derived from the CO_2 response curves (A vs. chloroplastic CO_2 partial pressure, C_c).

V_{cmax} was obtained from the initial slope of the A:C_c curve, near the CO_2 compensation point, using the following equation (von Caemmerer and Farquhar 1981):

$$A = \frac{(C - \Gamma_*)V_{c_{max}}}{C + K_c\,(1 + O/K_o)} - R_d \qquad \text{(Equation 6)}$$

Where C and O are the partial pressures of CO_2 and O_2 inside the chloroplast and K_c and K_o are Michaelis-Menten constants for CO_2 and O_2, respectively, with K_c = 260 µbar and K_o = 179 mbar (von Caemmerer et al. 1994), Γ_* is the chloroplastic CO_2 partial pressure at which the rate of carboxylation equals the rate of photorespiratory CO_2 release and R_d is the rate of respiration.

J at 1500µmol photons m^{-2} s^{-1} (J_{1500}) was derived from the CO_2 response curves by rearranging the following equation (von Caemmerer and Farquhar 1981):

$$A = \frac{(C - \Gamma_*)\,J}{4(C + 2\Gamma_*)} - R_d \qquad \text{(Equation 7)}$$

Where J is the potential rate of linear electron transport5.3 Results

5.3.1. Fold-change expression of PIP OE transgenes

Five independent OE lines were tested for each constitutive transgene expression of NtPIP1;1t, NtPIP1;3t, NtPIP1;5t and NtPIP2;4s. OE lines which had the highest fold- change expression in comparison to the house keeping gene (HKG), Ntubc2, were selected. We paired fold-change expression with segregation counts for T_1 seedlings grown on hygromycin selective media, seeking lines which had around 75% survivorship (of which 50% would be heterozygous and 25% homozygous for transgene integration in the genome), indicating transgene

insertion in a single locus. Two independent lines were selected for each OE construct: Lines 2 and 3 for PIP1;1t; lines 2 and 4 for PIP1;3t; lines 1 and 3 for PIP1;5t and lines 3 and 14 for PIP2;4s (Figure 5.2).

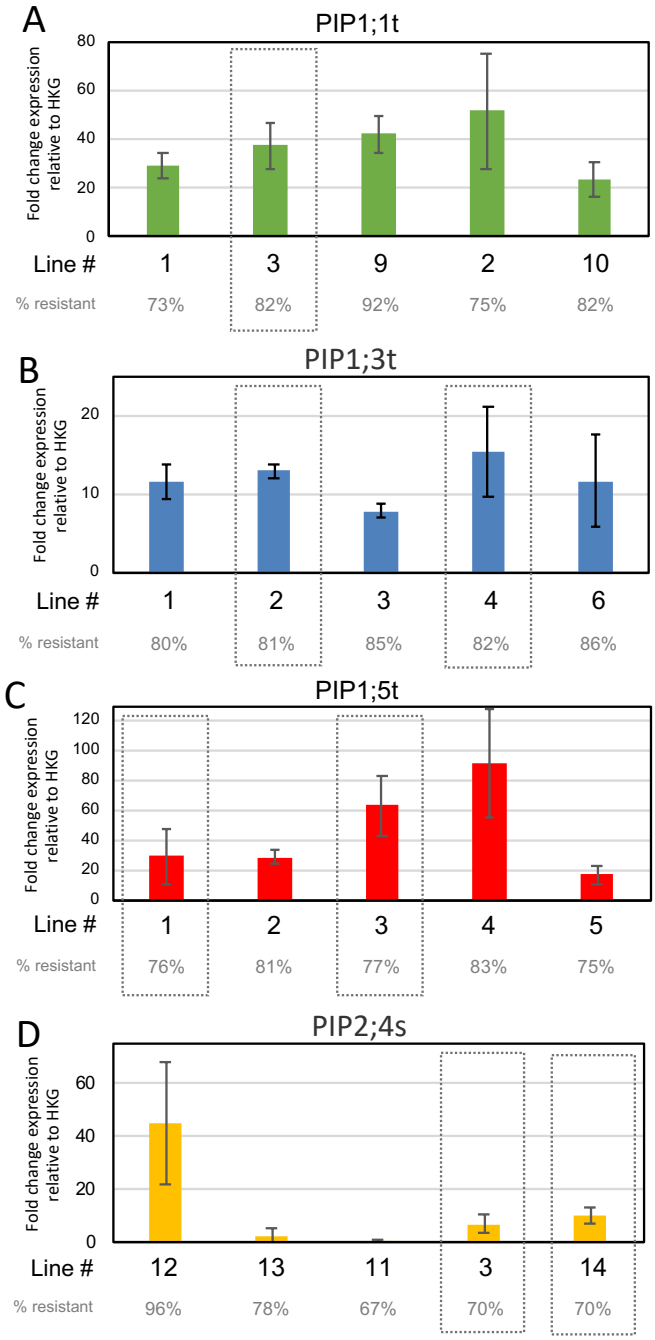


Figure 5. 2. Fold-change expression of PIP OE transgenes relative to Ntubc2 house-keeping gene (HKG). *Over expression of A. PIP1;1t, B. PIP1;3t, C. PIP1;5t and D. PIP2;4s was quantified via qRT-PCR, relative to expression of Ntubc2 housekeeping genes. Five independent lines were tested for each OE construct, Line numbers (#) are indicated below each histogram, along with the percentage survival in the presence of antibiotic (hygromycin) containing media. A survivorship around 70-80% is expected*

for a single locus T-DNA insertion in the OE line. Grey dashed boxes indicate OE line selected for further experiments.

5.3.2 A:C_i curves of PIP OE lines

The response of CO_2 assimilation rate (A) to increasing intercellular CO_2 concentrations (C_i) was investigated for *PIP1;1t*, *PIP1;3t*, *PIP1;5t* and *PIP2;4s* T_1 OE lines (Figure 5.3). We observed similar trends in assimilation rates for the two independent T_1 lines tested for each PIP OE construct.

We saw a difference in assimilation rate between the PIP OE lines and GUS control plants, where PIP1 lines (PIP1;1t, PIP1;3t and PIP1;5t) had higher assimilation rates for both the Rubisco-limited and the electron transport limited (low to high C_i respectively) regions of the A:C_i curves. By contrast, PIP2;4s OE lines had lower assimilation rates compared to the GUS control plants across the C_i range measured (Figure 5.3).

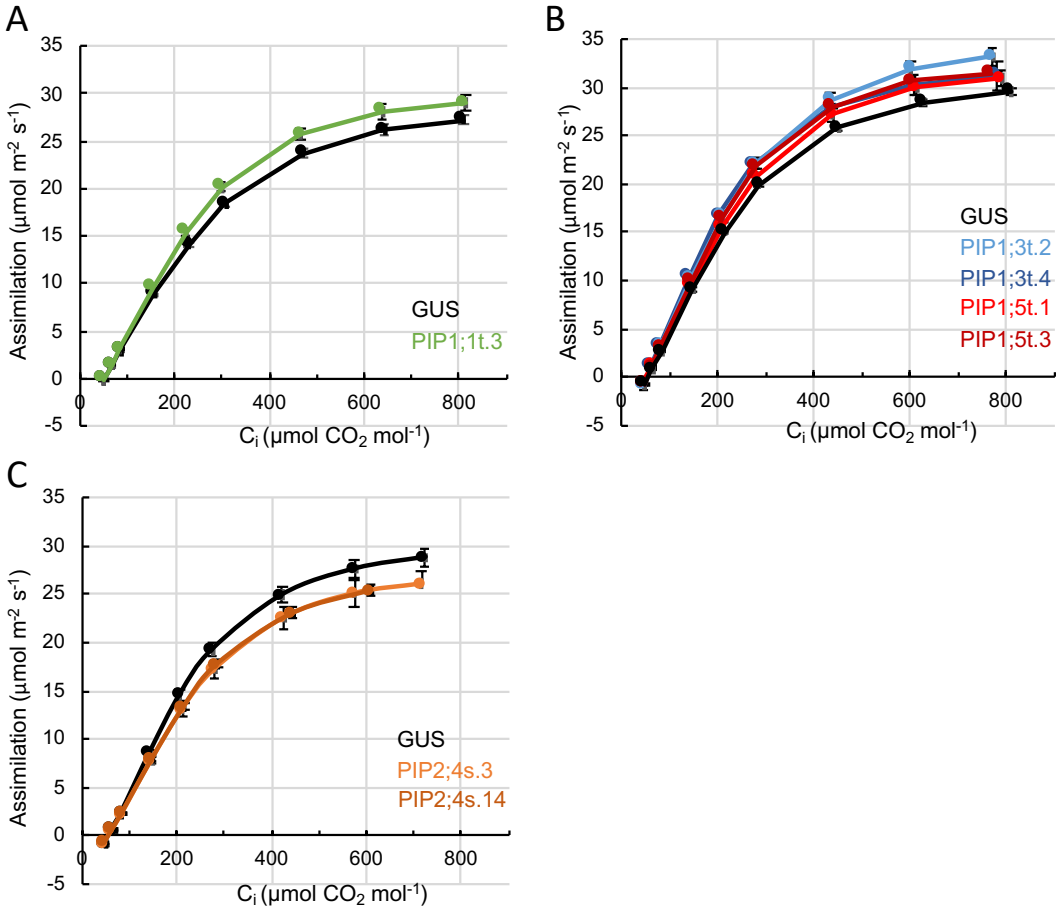


Figure 5. 3. CO_2 response (A:C_i) curves of tobacco PIP1;1t, PIP1;3t, PIP1;5t and PIP2;4s OE lines. *CO_2 assimilation rates were measured with LICOR-6400 gas exchange instruments over a range of ambient CO_2 concentrations. Three separate experiments were conducted measuring A. PIP1;1t Line 3 and GUS control line; B. PIP1;3t lines 2 and 4, PIP1;5t lines 1 and 3 and GUS control; and lastly C. PIP2;4s lines 3 and 4, and GUS control.*

5.3.3 Deriving mesophyll conductance from $^{13}CO_2$ discrimination measurements

Mesophyll conductance (g_m) and chloroplast CO_2 partial pressures were derived from concurrent measurements of carbon isotope discrimination (Δ) and leaf gas exchange (using the TDL and LICOR-6400s, respectively).

Representative Δ values are shown with respect to their C_i/C_a values for the tobacco PIP OE *PIP1;3t* (Lines 2 and 4), *PIP1;5t* (Lines 1 and 3) and GUS line control (Figure 5.4). The theoretical line for $^{13}CO_2$ discrimination in C_3 plants which assumes infinite g_m and no discrimination associated with respiration and photorespiration

(Δ_i = 4.4 +(29 - 4.4) C_i/C_a) is shown as the yellow line (Figure 5.4). The deviation of the measured discrimination from the theoretical line is related to the drawdown in CO_2 partial pressure within the mesophyll (C_i-C_c , indicated by green arrow in Figure 5.4). Points with a smaller deviation/difference from the theoretical line for $^{13}CO_2$ discrimination reflect that C_c is closer to C_i due to having smaller rates of CO_2 assimilation and/or of increased CO_2 diffusion into the chloroplast. Overall, the observed $^{13}CO_2$ discrimination values fall considerably below the line at their given C_i/C_a values, consistent with there being a significant drawdown in the partial pressure of CO_2 within the mesophyll. No significant differences were observed between the PIP OE lines and the GUS control (C_i-C_c values listed in Table 5.2).

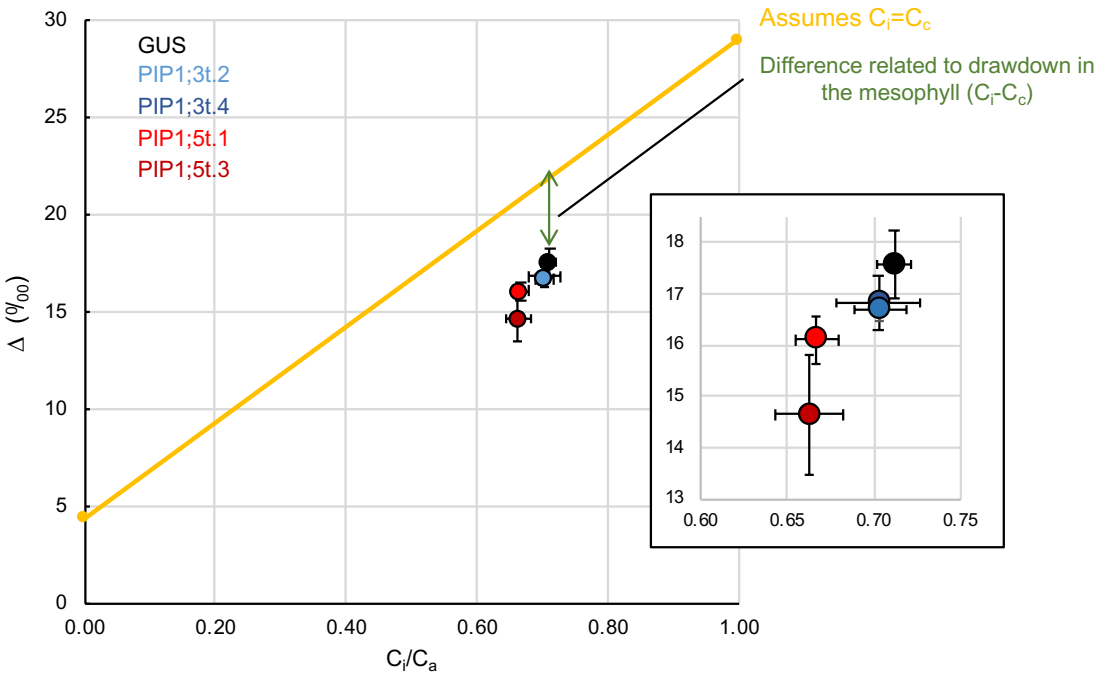


Figure 5. 4. Relationship between observed $^{13}CO_2$ discrimination (Δ_o) and the ratio of intercellular to ambient CO_2 partial pressure (C_i/C_a) for PIP1;3t and PIP1;5t OE lines. *The yellow line represents the simplified theoretical relationship for $^{13}CO_2$ discrimination in C_3 plants (Δ_i), which assumes $C_i = C_c$. Included are measurements for PIP1;3t lines 2 and 4 (light and dark blue respectively), PIP1;5t lines 1 and 3 (light and dark red respectively) and the GUS control plants (black). Green arrow represents the change in discrimination associated with the CO_2 drawdown in the mesophyll (C_i-C_c).*

5.3.4 Summary of measured photosynthetic parameters for NtPIP OE lines

The photosynthetic and leaf parameters for the PIP1 and PIP2 OE lines acquired during the measurement of $^{13}CO_2$ discrimination (TDL, 2% oxygen) and CO_2 response curves (LI-6400, 21% oxygen) are listed in Tables 5.2 and 5.3.

During TDL measurements (2% oxygen), no variation in assimilation rates between control plants and the PIP1 OE lines were observed in any of the three experiments. g_s was slightly lower in PIP1;1t and PIP1;5t OE lines, with OE PIP1;5t.3 having a statistically significant reduction compared to GUS control plants. g_m was slightly lower for all PIP1;1t, PIP1;3t and PIP1;5t OE lines, although reductions were not statistically significant ($p>0.05$). PIP2;4s OE lines appeared to have no

noticeable differences in A, g_s or g_m measurements obtained during the TDL measurements compared to GUS controls (Table 5.2). The drawdowns associated with the stomatal and mesophyll conductance, C_a-C_i and C_i-C_c, respectively, were similar between lines in each experiment (Table 5.2 and 5.3).

Gas exchange measurements in 21% oxygen showed that PIP1;1t, PIP1;3t and PIP1;5t OE lines had increased assimilation rates, with OE PIP1;3t.2 having a statistically significant increase relative to the GUS control plants. No differences in assimilation rates in 21% oxygen between GUS control and PIP2;4s OE lines were observed (Table 5.3). During the measurement of the A:Ci response curves, the CO_2 concentration in the reference air was controlled. This resulted in slightly lower C_a values for measured PIP1;3t.2 and PIP1;1t.3 OE lines due to their greater A compared to their GUS controls. These small differences would not have significantly impacted on g_s, but to make comparisons easier, the C_a - C_i drawdown in CO_2 partial pressure is presented. No differences in g_s and C_i were observed between any of the PIP1 and PIP2 OE lines and the GUS controls. The drawdowns associated with the stomata (C_a-C_i) for PIP1;1t and PIP2;4s OE lines were similar to the GUS controls. While PIP1;3t and PIP1;5t OE lines had slight increases in C_a-C_i compared to the GUS plants in their respective experiments, these differences were not statistically significant.

Increases in V_{cmax} were observed in all PIP1 OE lines compared to GUS plants in their respective experiments, with OE lines PIP1;1t.3, PIP1;3t.2 and PIP1;3t.4 having statistically significant increases (V_{cmax} of OE PIP1;5t.3 was not marked as statistically significant due to having a p-value of 0.052, Table 5.4). Similar to the increases in V_{cmax}, the PIP1 OE lines also had a higher J at an irradiance of 1500 µmol photons m^{-2} s^{-1} (J_{1500}) compared to the GUS plants in their respective experiments, with the PIP1;3t.2 OE having a statistically significant increase in J_{1500}. The J_{1500} values for the PIP2;4s OE lines were slightly lower than that of the measured GUS control plants, although such decreases were not statistically significant ($p<0.05$) (Table 5.4).

There were no differences in either Leaf Mass per Area (LMA, g m^{-2}) or Water content (percentage of water in leaf fresh mass, %) between GUS control plants and

PIP1;3t, PIP1;5t or PIP2;4s plants (Table 5.4). When PIP OE plants were compared against GUS controls at the time of measurement, there was no noticeable difference in plant size or rate of development (Figure 5.5).

Table 5. 2 Photosynthetic parameters obtained during TDL measurement in 2% oxygen for PIP1;1t, PIP1;3t, PIP1;5t and PIP2;4s OE lines. *Measurements obtained with LI6400s in association with the TDL included A (µmol m^{-2} s^{-1}), g_s (mol H_2O m^{-2} s^{-1}), g_m (mol CO_2 $m^{-2}s^{-1}bar^{-1}$), C_i, C_c, C_a-C_i and C_i-C_c (all µbar). Independent lines for each PIP are noted by the last number e.g. PIP1;1t.3 is line 3 for PIP1;3t. mean ± s.e., n=4-6. Asterisks * and ** indicate statistically significant difference to the GUS control in each experiment of p<0.05 and p<0.01 respectively.*

	2% Oxygen TDL						
Exp. 1	**A**	**g_s**	**g_m**	**C_i**	**C_c**	**C_a - C_i**	**C_i - C_c**
Gus control	24.6 ± 1.5	0.67 ± 0.04	0.49 ± 0.04	244.9 ± 10.1	191.9± 14.5	72.3 ± 3.6	53.0 ± 6.3
OE PIP1;1t.3	23.5 ± 1.4	0.57 ± 0.07	0.44 ± 0.04	238.0 ± 10.4	182.1 ± 10.2	86.0 ± 9.5	56.4 ± 5.2
Exp. 2							
Gus control	29.0 ± 1.0	0.51 ± 0.04	0.46 ± 0.03	253.6 ± 4.6	189.3 ± 7.9	106.3 ± 4.6	64.3 ± 5.25
OE PIP1;3t.2	27.8 ± 1.5	0.49 ± 0.04	0.42 ± 0.02	252.6 ± 8.3	185.4 ± 7.3	106.9 ± 8.8	67.2 ± 1.4
OE PIP1;3t.4	29.9 ± 1.1	0.54 ± 0.05	0.44 ± 0.02	253.1 ± 5.3	181.3 ± 6.0	106.8 ± 5.5	71.8 ± 5.2
OE PIP1;5t.1	27.3 ± 0.7	0.42 ± 0.03	0.43 ± 0.02	240.5 ± 4.2	174.8 ± 6.0	112.8 ± 6.2	66.3 ± 4.3
OE PIP1;5t.3	26.3 ± 1.95	**0.38* ± 0.02**	0.41 ± 0.05	235.5 ± 7.7	169.3 ± 6.0	124.1 ± 8.1	66.2 ± 3.6
Exp. 3							
Gus control	28.6 ± 1.1	0.47 ± 0.05	0.35 ± 0.02	248.0 ± 7.0	164.8 ± 6.3	116.5 ± 6.9	83.2 ± 3.0
OE PIP2;4s.3	27.7 ± 1.0	0.52 ± 0.04	0.35 ± 0.02	263.1 ± 3.4	181.8 ± 3.4	101.12 ± 3.2	81.9 ± 3.2
OE PIP2;4s.14	26.4 ± 1.5	0.48 ± 0.02	0.32 ± 0.02	262.2 ± 6.1	178.7 ± 5.3	102.4 ± 6.0	83.5 ± 2.3

Table 5. 3. Photosynthetic parameters obtained in 21% oxygen for PIP1;1t, PIP1;3t, PIP1;5t and PIP2;4s OE lines. *A (µmol m^{-2} s^{-1}), g_s (mol H_2O m^{-2} s^{-1}), C_a, C_i and C_a-C_i (µbar) values obtained from LI-6400 measurements at 400 µmol CO_2 mol^{-1} concentration. Independent lines for each PIP are noted by the last number e.g. PIP1;1t.3 is line 3 for PIP1;3t. mean ± s.e., n=4-6. Asterisks * and ** indicate statistically significant difference to the GUS control in each experiment of p<0.05 and p<0.01 respectively.*

	21% Oxygen LI-6400				
Exp. 1	**A**	**g_s**	**C_a**	**C_i**	**C_a-C_i**
Gus control	18.4 ± 0.4	0.56 ± 0.04	373.9 ± 1.0	306.4 ± 4.7	67.5 ± 6.7
OE PIP1;1t.3	19.5 ± 0.6	0.51 ± 0.03	**369.8* ± 1.3**	293.2 ± 4.8	76.5 ± 5.6
Exp. 2					
Gus control	20.0 ± 0.3	0.38 ± 0.01	368.1 ± 0.5	269.2 ± 1.7	98.9 ± 1.9
OE PIP1;3t.2	**21.7* ± 0.6**	0.44 ± 0.03	**365.2**± 1.0**	269.2 ± 3.9	95.9 ± 4.6
OE PIP1;3t.4	21.4 ± 0.5	0.40 ± 0.04	365.8 ± 1.0	262.4 ± 6.2	103.4 ± 7.0
OE PIP1;5t.1	20.5 ± 0.6	0.41 ± 0.05	367.4 ± 0.9	267.6 ± 7.6	99.8 ± 8.3
OE PIP1;5t.3	20.6 ± 0.4	0.39 ± 0.01	367.2 ± 0.7	267.9 ± 1.9	99.3 ± 1.9
Exp. 3					
Gus control	18.3 ± 0.8	0.32 ± 0.04	370.5 ± 1.4	258.4 ± 8.5	112.1 ± 9.7
OE PIP2;4s.3	18.0 ± 0.8	0.36 ± 0.02	370.6 ± 1.2	275.0 ± 4.8	95.6 ± 4.0
OE PIP2;4s.14	15.9 ± 1.0	0.27 ± 0.04	374.2 ± 1.7	259.6 ± 9.7	114.6 ± 11.0

Table 5. 4. Derived photosynthetic and leaf parameters obtained for PIP1;1t, PIP1;3t, PIP1;5t and PIP2;4s OE lines. *V_{cmax} (μmol CO_2 m^{-2} s^{-1}) and J_{1500} (μmol electrons m^{-2} s^{-1}) were derived from A:Cc curves assuming the g_m derived from the TDL measurements. Average Leaf Mass per Area (LMA, g m^{-2}) and Water content (percentage of water in leaf fresh mass, %) are also included for Experiments 2 and 3. Independent lines for each PIP are noted by the last number e.g. PIP1;1t.3 is line 3 for PIP1;3t. mean ± s.e., n=4-6. Asterisks * and ** indicate statistically significant difference to the GUS control in each experiment of p<0.05 and p<0.01, respectively.*

	Derived parameters			
Exp. 1	**V_{cmax}**	**J_{1500}**	**LMA**	**Water content**
Gus control	74.0 ± 2.1	131.8 ± 2.8	-	-
OE PIP1;1t.3	**86.8** ± 3.6**	140.4 ± 5.0	-	-
Exp. 2				
Gus control	95.8 ± 3.2	151.8 ± 3.2	23.3 ± 2.0	92.1 ± 0.6
OE PIP1;3t.2	**112.8*± 6.2**	**170.8** ± 4.3**	24.8 ± 2.3	91.8 ± 0.8
OE PIP1;3t.4	**113.7** ± 3.6**	159.6 ± 8.3	24.2 ± 1.6	91.6 ± 0.5
OE PIP1;5t.1	99.1 ± 3.4	156.8 ± 4.3	24.4 ± 1.4	91.8 ± 0.4
OE PIP1;5t.3	107.9 ± 3.8	160.4 ± 5.8	25.2 ± 2.0	91.8 ± 1.0
Exp. 3				
Gus control	102.2 ± 5.3	152.9 ± 5.5	22.9 ± 0.9	91.9 ± 0.6
OE PIP2;4s.3	87.8 ± 7.7	137.2 ± 8.7	24.3 ± 1.2	91.9 ± 0.8
OE PIP2;4s.14	88.2 ± 3.4	136.5 ± 4.5	21.8 ± 1.3	92.8 ± 0.4

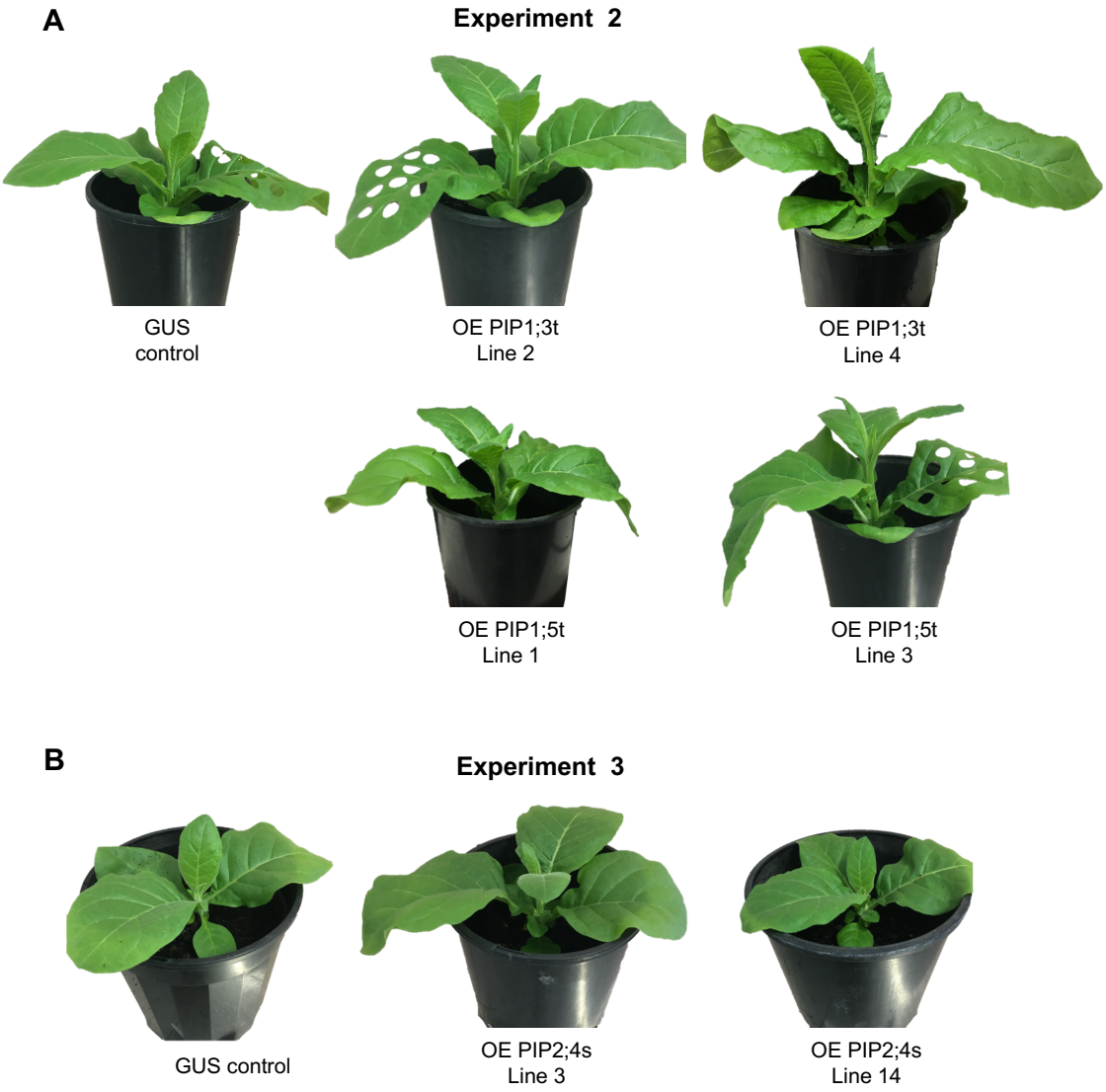


Figure 5. 5. Representative plants from PIP OE lines and the GUS control plants measured in experiments 2 and 3. ***A.*** *Plants measured in Experiment 2 were PIP1;3t OE lines 2 and 4, PIP1;5t OE lines 1 and 3 and GUS controls.* ***B.*** *Plants measured in Experiment 3 were PIP2;4s OE lines 3 and 14, and GUS controls. Diameter for each pot was 20cm.*

5.4 Discussion

5.4.1 Challenges in predicting CO_2-permeable AQP candidates based on primary sequence

Aquaporins in the PIP sub-family have been identified as genetic engineering targets with the potential to enhance crop productivity through increased photosynthetic efficiency (von Caemmerer and Evans 2010). As such, a better understanding of the factors determining PIP CO_2 permeability is crucial in order to efficiently engineer enhanced photosynthesis in target crop species (Groszmann et al. 2017).

CO_2 has been suggested to follow a different path to water through AQPs, potentially though the central pore of the AQP tetramer (Wang and Tajkhorshid 2007; Hub et al. 2009). Mori *et al.* 2014 described a residue in barley PIP2;3 (*HvPIP2;3*) which was deemed critical for enabling CO_2 permeability in a *Xenopus* oocyte expression system. They further highlighted the presence of this specific residue (Ile-254) in other PIPs which have been shown to also permeate CO_2 (Mori et al. 2014). Analysis of the same residue in tobacco PIPs, showed that all 29 PIPs have an 'Ile' at the characterised position (data not shown). It is unlikely that all 29 tobacco PIPs are permeable to CO_2, especially as studies have shown particular tobacco PIP2s to be impermeable to CO_2 in a yeast expression system (Otto et al. 2010). Therefore, homology at this site suggests that other factors are also required for CO_2 permeability.

We selected NtPIP genes (*NtPIP1;1t*, *NtPIP1;3t* and *NtPIP1;5t*) with high homology to an established CO_2 permeable AQP (*NtAQP1/NtPIP1;5s*) to functionally characterise *in planta*. This could have led us to identify any key amino acid residue differences that might occur between CO_2-permeable and non-permeable PIP1 genes. However, we were unable to dissect such residue variations due to the similarities we observed for the measured photosynthetic parameters in the PIP1;1t, PIP1;3t and PIP1;5t OE lines. We did see differences in photosynthetic parameters between the PIP1 and the PIP2;4s OE lines (discussed below in Section 5.4.5), indicating a PIP1 vs. PIP2-specific alteration to photosynthetic efficiency, for the genes tested.

The tobacco PIP 1 clade contains a further 4 genes (NtPIP1;2s, NtPIP1;2t, NtPIP1;7t and NtPIP1;8s) less similar to NtPIP1;5s/NtAQP1, occurring in a separate sub-clade within the PIP1s (Figure 5.1). As such, tobacco OE lines of these less-homologous PIP1s could be generated for future experiments in order to observe their photosynthetic phenotypes. Such sequence comparisons could be integrated into our 3D homology models to potentially improve current understanding of CO_2 permeation through AQPs.

5.4.2 OE of NtPIP1s increased V_{cmax} and J with no significant changes in mesophyll conductance

Over-expression of NtPIP1s with high homology to NtAQP1 in tobacco resulted in greater photosynthetic capacity, with the PIP1;1t, PIP1;3t and PIP1;5t OE lines having increased Rubisco carboxylation capacity (V_{cmax}) and potential electron transport rate (J_{1500}) compared to the GUS control plants. These results are concordant with findings from Flexas *et al*. (2006), where, over expression of NtPIP1;5s/NtAQP1 induced in tobacco leaves lead to higher assimilation rates, J and slight increases in V_{cmax}. Other studies have also reported increased photosynthetic rates upon OE of PIPs in plants (Hanba et al. 2004; Aharon et al. 2003; Xu et al. 2018; Sade et al. 2014; Heckwolf et al. 2011). In addition to increases in assimilation rates, these studies also reported an increase in mesophyll conductance (g_m). However, we did not observe an increase in g_m in any of our NtPIP1 OE lines.

The lack of change in g_m of the PIP1;1t, PIP1;3t and PIP1;5t OE lines was a puzzling result, as we would expect that the enhanced photosynthetic capacity measured in these plants would be accompanied by an increase in facilitated diffusion of CO_2 to the sites of fixation, through AQPs (Evans et al. 2009; Flexas et al. 2012). g_m has been observed to scale with photosynthetic capacity (Evans and Loreto 2000; Evans et al. 1994; von Caemmerer and Evans 1991). Modified intracellular CO_2 concentrations could potentially trigger differences in the development of leaf photosynthetic capacity (Flexas et al. 2006). Additionally, altered expression of PIP genes within plants could impact on gene expression of other PIP isoforms. As such, it is possible that the constitutive OE of PIP1s in tobacco

lead to pleiotropic effects that impacted g_m measurements and these differed from results obtained when using an inducible promoter.

5.4.3 Potential pleiotropic effects resulting from PIP1 OE: impacts on hydraulic conductivity and leaf anatomy

As well as enhancing CO_2 diffusion *in planta*, previous studies have also observed that reduced expression of NtPIP1;5s/NtAQP1 altered root hydraulic conductivity and the whole plant response to water stress (Siefritz et al. 2002). As our functional characterisation in yeast of NtPIP1;5s/NtAQP1 did not reveal any water transporting capabilities for this AQP monomer (Figure 4.2, Chapter 4), the altered expression of NtPIP1;5s/NtAQP1 *in planta* could be altering expression of other water-permeable AQPs if plant water relations have been altered.

Aharon et al. (2003) reported that constitutive OE of Arabidopsis PIP1b (homolog of NtAQP1) in tobacco resulted in transgenic plants having increased stomatal density and a greater rate of water consumption. This meant that under favourable conditions, AtPIP1b OE plants had greater rates of transpiration and photosynthesis (Aharon et al. 2003). However, under drought conditions, OE of AtPIP1b resulted in detrimental effects, such as faster wilting of tobacco plants (Aharon et al. 2003).

Our gas exchange measurements of the PIP1 OE lines did not show increases in stomatal conductance (indicative of increased plant water use). Rather, stomatal conductance was not significantly different from the GUS controls (with the exception of the NtPIP1;5t.3 OE line, which had a statistically significant reduction in g_s under 2% O_2, Table 5.2). It is possible that the NtPIP OE lines had altered leaf area to root water uptake capacity, or other anatomical differences that increased plant water consumption, perhaps arising early on in plant development.

Hanba *et al.* (2004) investigated leaves of rice transgenic lines over expressing the barley PIP2;1 gene. They identified anatomical differences between OE lines which had 1.7 and 4.2 fold OE of the PIP gene. The 4.2-fold OE lines had thicker mesophyll cell walls and less exposed chloroplast surface area. They hypothesised that leaf development might have been affected by water deficit due

to increased water loss, although the plants were grown in well-watered conditions. Other anatomical differences included thicker leaf cuticle, higher mesophyll cell density and more developed collenchyma, features associated with water stress (Esau 1965; Hanba et al. 2004). It was also noted that these leaf anatomical differences occurred in mature leaves, rather than in expanding leaves, suggestive of acclimation responses to physiological or water status of the plants (Hanba et al. 2004). A negative correlation has been observed between g_m and cell wall thickness, consistent with the cell wall resistance accounting for up to 50% of the total mesophyll resistance (Terashima et al. 2011; Evans et al. 2009). Therefore, the reported increase in mesophyll cell wall thickness in HvPIP2;1 transgenic lines could offset any AQP-induced increase in membrane permeability resulting in less change to mesophyll conductance (Hanba et al. 2004).

To investigate these ideas, it would be necessary to quantify leaf morphology of the PIP1 OE lines to observe whether there are differences in mesophyll cell wall thickness, surface area of chloroplasts exposed to intercellular airspaces or other anatomical changes. Additionally, gas exchange and leaf morphological characteristics of newly expanding and mature leaves could be compared.

5.4.4 Constitutive vs. inducible expression of PIP1s in tobacco

In Flexas *et al.* (2006), *NtPIP1;5s/NtAQP1* was over expressed with an inducible promoter, for which expression was initiated 2-3 days prior to measurement collection. By contrast, in this study the *NtPIP1;1t*, *NtPIP1;3t* and *NtPIP1;5t* genes were constitutively over expressed from seed germination (2x35S CaMV promoter). It is possible that the increases in g_m reported in Flexas *et al*. (2006) could be detected due to there being insufficient time for other pleiotropic changes to occur. Clearly, the creation and measurement of NtPIP1;5s/NtAQP1 OE lines with a constitutive promoter is needed to clarify these different results. It was unfortunate that the generation of the NtPIP1;5s/NtAQP1 OE lines had not been completed in time for this project, due to technical problems.

As such, even if PIP1;1t, PIP1;3t and PIP1;5t are CO_2 pores *in planta*, pleiotropic effects arising from constitutive over-expression could be masking our ability to detect enhancements in g_m.

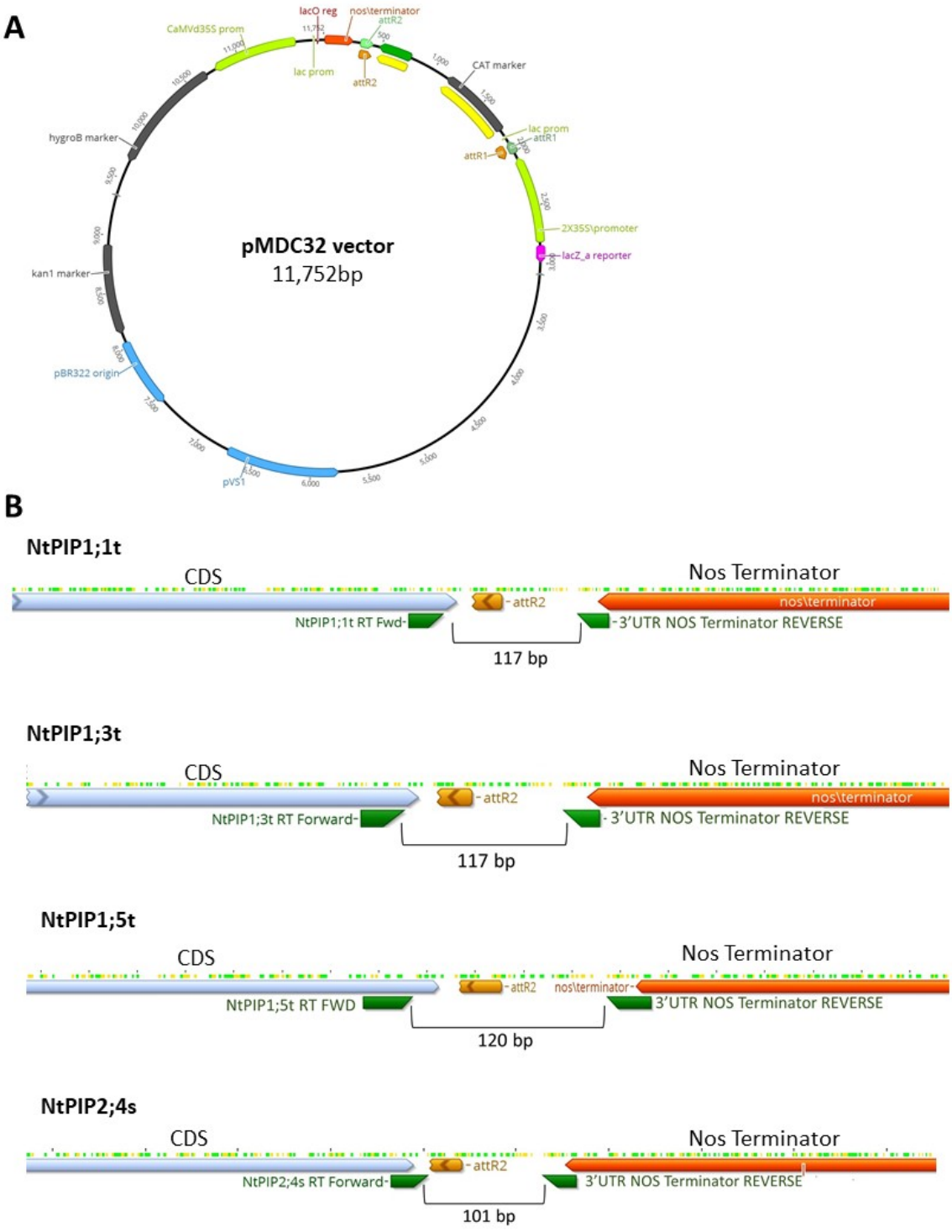
5.4.5 PIP1 vs PIP2-specific alteration to photosynthetic parameters

In contrast to the PIP1 OE lines, the over expression of NtPIP2;4s *in planta* resulted in no significant differences in measured photosynthetic parameters. Both PIP2;4s.3 and NtPIP2;4s.14 OE lines had slight reductions (although not statistically significant) in the derived parameters V_{cmax} and $J_{1500,}$ compared to the GUS control plants (Tables 5.2-5.4). These results suggest a difference in physiological function *in planta* between these two PIP sub-groups in tobacco, with OE of PIP1s resulting in an enhancement of photosynthetic capacity. The differences in observed photosynthetic parameters for the PIP1 and PIP2 OE lines could be the result of altered tetramer compositions, with different proportions of AQP monomers occurring within the assembled tetramers (Flexas et al. 2012). AQP tetramer composition has been shown to impact overall AQP functionality. Otto *et al.* (2010) suggested that 1 water-conducting PIP2 monomer is enough to facilitate water transport close to maximum level, whereas 3-4 CO_2 permeable monomers (PIP1s) are required to enable maximal CO_2 permeability (in a heterologous expression system). As such, it could be that in the PIP2;4s OE lines, the tetramer composition changed such that the increased integration of PIP2;4 in the plasma membrane reduced the diffusion of CO_2 across membranes.

5.5 Conclusions

Aquaporins in the PIP subfamily have been implicated in facilitating CO_2 diffusion across the plasma membrane and chloroplast envelope, making them targets for enhancing mesophyll conductance and photosynthetic rates. Our results showed that the constitutive OE of tobacco PIP1 isoforms highly homologous to NtAQP1/NtPIP1;5s (an established CO_2 pore) led to increases in assimilation rates, however we observed no alteration of mesophyll conductance. Constitutive OE of PIP isoforms could lead to pleiotropic effects, impacting our ability to detect enhancements to mesophyll conductance.

5.6 Supplementary Figures



Supplementary Figure S5. 1. Schematic representation of pMDC32 expression vector and primer annealing sites for the primers used for qRT-PCR analyses, testing for transgene expression. A. *Vector map for pMDC32 gateway expression vector, containing 2x35SS constitutive promoter (lime green), Nos terminator (red) and hygromycin antibiotic selection gene (grey).* ***B.*** *Over expression lines tested were NtPIP1;1t, NtPIP1;3t, NtPIP1;5t and NtPIP2;4s OE lines. PIP-specific Forward primers were*

designed at the 5' end of the CDS (cloned in the pMDC32 Gateway destination vector). Identical reverse primer was used, spanning the 3'UTR and Nos Terminator regions. Primers are indicated by green arrows. CDS indicated by light blue regtangles, NOS terminator indicated by red rectangles. Amplification fragment lengths are noted for each gene construct, ranging between 101-120bp in length.

Chapter 6: General Discussion and Conclusions

6.1 Overview

AQPs are a diverse protein family of growing interest to plant biologists due to their involvement in numerous physiological processes essential for plant growth. My PhD research aimed to contribute to current knowledge of AQP biology by establishing the AQP family in tobacco and further characterising a diverse set of tobacco AQP isoforms through functional studies in yeast. Additionally, candidate CO_2-permeating PIP genes were over expressed *in planta* to explore their effects on mesophyll conductance and photosynthetic rates of tobacco leaves.

6.1.1 Characterisation of the AQP family in tobacco

Tobacco is a popular study species in plant biology, being a model plant for photosynthesis research, suitable for a variety of biotechnology applications as well as well as being part of the horticulturally significant Solanaceae family (containing tomato and potato). Tobacco is also a relevant species for studying aquaporins, with the tobacco AQP1 gene (NtAQP1) being the first identified CO_2 permeable aquaporin in plants (Uehlein et al. 2003).

In order to study AQPs in tobacco and expand our knowledge of this diverse protein family, it was firstly essential to have a well-established AQP gene family. In Chapter 2, we characterised the tobacco AQP family, identifying 76 full-length genes and elucidating their gene structures, protein sequences, functionally relevant residues, sub-cellular localisations and gene expression tissue-specificity. We uncovered complexities in the tobacco AQPs family, with tobacco containing double the number of AQP gene isoforms due to its recent evolutionary emergence from a an allotetraploid hybridisation event between *N. sylvestris* and *N. tomentosiformis*. We structured the AQP genes in tobacco with reference to the already characterised tomato (Reuscher et al. 2013) and potato (Venkatesh et al. 2013) AQP genes. This allowed us to identify orthologous genes across these closely related species.

Our findings could be useful in exploring variation within the elucidated AQP ortholog groups within Solanaceae (i.e. tomato and potato orthologs and the tobacco

sister genes inherited from *N. sylvestris* and *N. tomentosiformis*), and comparing their gene expression tissue specificity and substrate specificities. For example, we could compare whether the Solanaceae ortholog groups, which have similarities in tissue-specific expression, also have homologous substrate specificities, linking homology in tissue specificity to potential functional roles across these species. Alternatively, ortholog groups which have variation in gene expression profiles (e.g. *PIP2;1*, *TIP2;3*, *NIP6;1* orthologs; Figure 2.7 Chapter 2) could also be screened for variation in permeating substrates. Such information contributes towards improving our understanding of which residues confer substrate specificity across AQPs which already have a high sequence identity. Above all, elucidation of substrate specificities of orthologous AQPs has the potential to facilitate translation of knowledge from tobacco into these horticulturally important crop species.

6.1.2 Establishment of high-throughput functional assays in yeast to test for multiple permeating substrates

Despite increasing information available on aquaporin functions, there is still insufficient information to predict with any reasonable certainty the possible permeating substrates of a given AQP based solely on amino acid primary sequence. Deciphering which substrates permeate through a particular AQP in isolation from other interacting AQP monomers or regulatory processes (e.g. gating), can be aided through the use of heterologous functional assays. In Chapter 3 we established AQP functional assays using yeast as a heterologous expression system. We developed methods based on growth kinetics to test for permeation of hydrogen peroxide or urea through AQPs expressed in yeast cells. These experiments, as well as others developed within our laboratory (testing for water and boron permeabilities), provide a novel and high-throughput methodology to accurately assay a variety of substrates for permeation through the plasma membrane facilitated by an AQP. Crucially, we first verified that each AQP localised to the plasma membrane before determining permeability characteristics. This ensured that a non-permeating result was not simply a reflection of failure of the AQP to successfully incorporate into the plasma membrane.

Our newly developed, functional assays could be expanded to test for permeation of substrates additional to the four presented here. For example, growth-based assays could be developed to test for ammonia permeability in the tobacco AQPs tested, and explore potential sequence variation and/or structural determinants of AQPs permeating urea and ammonia (Kirscht et al. 2016). Additionally, toxicity based assays could be developed for essential metalloid compounds required in plants, such as silicon, arsenous acid and antimonite which have also been shown to permeate through AQPs, predominantly through NIP isoforms (Bienert et al. 2008; Bárzana et al. 2014). Elucidation of these additional permeating substrates could add to our catalogue of annotated tobacco AQPs with known substrate specificities and enrich our understanding of AQP biology.

6.1.3 Functional characterisation of tobacco AQPs

In Chapter 4 we were able to utilise our newly established yeast functional assays to test substrate specificities (water, urea, boron and hydrogen peroxide) of a diverse set of tobacco AQPs. We selected PIP, TIP and NIP candidate genes, being the three of the larger AQP sub-families. 3D protein homology modelling analyses were conducted in order to link the size and chemistry of the newly identified permeating substrates to the specific AQP's pore diameter and physico-chemical properties of the pore lining-residues. For each gene candidate, we also confirmed *in planta* subcellular localisations, observing a mix of plasma membrane, tonoplast and ER membrane integrations. Although chloroplast envelope localisation was not investigated in my research, future expriments should explore AQP localisation in the chloroplast envelope in mesophyll cells, especially for the PIP1 isoforms hypothesised to be involved in enhancing photosynthetic parameters.

By combining functional and cellular localisation information with gene expression data (presented in Chapter 2), we speculated on *in planta* physiological functions for these PIP, TIP and NIP gene candidates.

Extending from our functional characterisation data, a more targeted exploration of certain AQP gene candidates could further tease apart spatial and temporal regulation of AQP expression. For example, *NtAQP* promoter::GUS fusion lines could be generated for genes which had highly targeted tissue specific

expression (e.g. to flowers, roots, leaves or stems). Previous studies have elucidated targeted spatial and temporal expression of certain AQP genes. For example, expression of NIP isoforms in Arabidopsis was shown to be spatially and temporally regulated during pollen tube development and pollination (Di Giorgio et al. 2016), with this differential expression localisation resulting in the assignment of distinct functional roles within the flower. Elucidating cell-specific expression of the tobacco AQPs we characterised, as well as identifying their permeating substrates (through our yeast-based assays), could enable us to predict more specific functional roles.

Overall, we hope that our broad functional characterisation approach (illustrated in Figure 6.1) will better direct scientists towards potential AQP gene candidates to use for subsequent plant experiments investigating physiological functions and manipulating plant traits of interest.

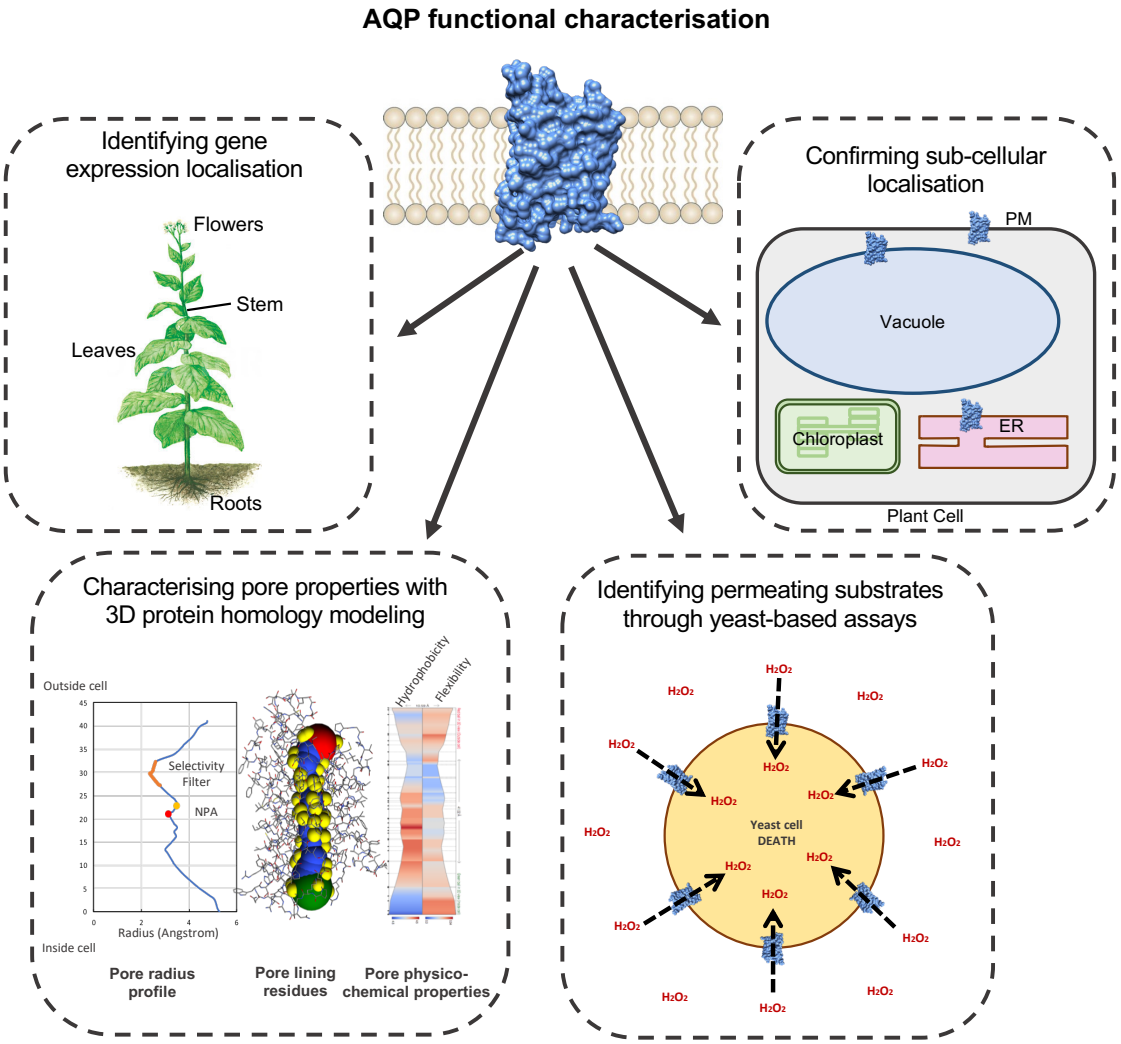


Figure 6. 1. Schematic representation of experimental components integrated within our tobacco AQP functional characterisation. *Experimental components consisted of confirming sub-cellular membrane integrations in planta through AQP::GFP translational fusion constructs, elucidating AQP substrate specificities through yeast based assays; utilising 3D protein homology modelling to obtain pore radius profile, pore lining residues and their physico-chemical properties; and lastly, incorporating RNA-seq gene expression results to identify tissue-specificity.*

6.1.4 In planta functional study of putative CO_2 pores in tobacco

In attempts to explore AQPs' role as facilitators of CO_2 diffusion across biological membranes, we identified a sub-set of tobacco PIP1s as putative CO_2 pores (based on homology to NtAQP1). We generated tobacco transgenic lines with constitutive OE of *PIP1;1t*, *PIP1;3t* and *PIP1;5t* genes to investigate their roles *in planta* (Chapter 5). Our aim was to test whether we could detect enhancement in

mesophyll conductance and plant photosynthetic efficiency, which we might expect if these AQPs enhanced CO_2 diffusion between intercellular airspaces and the mesophyll cells (Evans et al. 2009). Gas exchange measurements on PIP1 OE lines showed gene-specific increases in photosynthetic rates and modelled photosynthetic parameters (V_{cmax} and J_{1500}), concordant with other studies OE CO_2-permeable AQPs (Flexas et al. 2006; Heckwolf et al. 2011; Hanba et al. 2004). However, unlike in those studies, we observed no significant changes in mesophyll conductance. The undetected alteration of mesophyll conductance in the transgenic lines was an unexpected result, potentially due to pleiotropic effects arising from constitutive OE of PIP AQPs (addressed in Chapter 5, Discussion). A key next step towards resolving this result is the need to generate plants over-expressing *NtPIP1;5s/NtAQP1* to check the impact on mesophyll conductance. The generation of these plants had begun, but plants were not yet available to be included here.

Our *in planta* functional study highlights the complexities involved with altering AQP expression in plants, and the requirement of improved understanding of AQP regulation in order to engineer desired traits. The following sections address specific aspects of AQP biology and regulation that should be considered in future genetic engineering efforts.

6.2 Current challenges and future perspectives for engineering AQPs to increase crop yield and resilience.

Genetic engineering of AQPs in plants holds great potential in attempts to address food security challenges and future climatic conditions. AQPs have been implicated in enhancing plant photosynthetic efficiency (Flexas et al. 2006; Hanba et al. 2004; Heckwolf et al. 2011; Uehlein et al. 2008), improving plant resilience to drought stress (Sade et al. 2009; Cui et al. 2008; Zhou et al. 2012), improving tolerance to exposure to salinity (Liu et al. 2013) or toxic levels of micronutrients (Schnurbusch et al. 2010).

Although numerous studies have reported improved plant performance phenotypes upon altered expression of plant AQPs, a few have also reported some unexpected results. Over expression of certain AQP isoforms in plants increased

susceptibility to drought and salt stress. For example, OE of Arabidopsis AtPIP1;4 and AtPIP2;5 in Arabidopsis resulted in increased water loss (Jang et al. 2007). Likewise, the OE of Arabidopsis AtPIP1b in tobacco resulted in plants that were more susceptible to drought or salt-stress (Aharon et al. 2003). In rice, OE of barley HvPIP2;1 lead to enhanced CO_2 conductance and CO_2 assimilation, but there was also greater sensitivity to water and salt stress (Hanba et al. 2004).

These plant physiological responses upon altered expression of certain AQP isoforms highlights the need to expand our understanding of AQP biology and AQP regulation within plants, in order to better tailor the resulting plant traits to meet desired outcomes.

6.2.1 Accurately deciphering AQPs' multiple permeating substrates

Recent publications demonstrate that single AQP isoforms are capable of permeating multiple substrates (Maurel et al. 2015). Our yeast-based functional characterisation of tobacco AQPs have added to the information on the various permeating substrates of the isoforms tested, with tobacco PIP2s (NtPIP2;4s and NtPIP2;5t) permeable to water and hydrogen peroxide, TIPs (NtTIP1;1s and NtTIP5;1t) permeable to water, urea and boric acid, and NtNIP2;1s permeable to boric acid and urea (Chapter 4). In future functional characterisation studies, it is advisable to test AQP permeability to several substrates, rather than assaying one specific substrate (e.g. testing solely for H_2O_2 permeability).

The ability of AQPs to acts as multifunctional channels should also be taken into account in AQP engineering efforts, as the OE of an AQP isoform might result in the unintended enhanced diffusion of off-target solutes. This could lead to undesired plant phenotypes, where the enhanced diffusion of off-target solutes is detrimental to the transgenic plants. For example, if a NIP isoform permeable to urea and boron was over-expressed in plant roots, it might enhance uptake of urea from the soil, but it could also increase uptake of metalloids (e.g. boron), that could potentially reach toxic concentrations within the plant.

Alternatively, the capability of AQPs to enhance diffusion of multiple solutes throughout the plant could be advantageous in order to stack beneficial traits that might arise from over-expression of a particular AQP isoform *in planta*. For example,

over-expression of a TIP permeable to water and urea could enhance plant resilience to water stress as well as improving nitrogen equilibration within cells.

6.2.2 Identifying post-translational modifications regulating AQP function

Compared to other kingdoms, plants contain a high number of AQP isoforms which occur in a large variety of modified forms. The occurrence of isoform multiplicity in plants suggests that their regulation is mediated via intricate co- and post-translational mechanisms (Abascal et al. 2014; Maurel et al. 2008). Secondary regulatory mechanisms in plants ensure the tight regulation of AQPs in response to stresses, with the gating of monomeric pores in response to external stimuli ensuring a key control over AQP function (Luang and Hrmova 2017). Post-translational modifications include phosphorylation of certain key residues, which can result in conformational changes of the AQP structure and pore closure (Guenther et al. 2003; Törnroth-Horsefield et al. 2006; Johansson et al. 1998). Other gating stimuli include protonation, resulting from changes in pH (Alleva et al. 2006), and also cell signalling mechanisms, such as the inhibition of AQPs by ROS observed in Arabidopsis roots (Maurel et al. 2008).

Our AQP family characterisation (Chapter 2) included the identification of potential regulatory sites (i.e. phosphorylation and protonation sites), using protein sequence alignments to AQP monomers with resolved crystal structures (Spinach PIP2;1) and the use of predictive software (NetPhos) identifying phosphorylatable residues. Characterisation of regulatory sites could contribute to current knowledge of AQP post-translational modifications not only within the more widely studied PIP sub-family, but also across the TIP, NIP, SIP and XIP sub-groups.

Improved knowledge of post-translational modifications of AQP isoforms would be advantageous to future genetic engineering effort of AQPs. For example residues susceptible to gating could be engineered at known gating sites in AQP monomers, so that candidate AQPs are switched on or off upon onset of particular stresses. Alternatively, specific phosphorylated residues could be substituted with either phopho-deficient (e.g. alanine) or phosphor-mimetic (e.g. aspartic acid), to force specific protein states.

6.2.3 Identifying potential AQP-AQP protein interactions

AQP monomers form individual functional pores and assemble as tetramers in biological membranes. AQP oligomeric assemblies can be composed of either identical or differing monomeric units, resulting in homo- or hetero-tetramers, respectively (Jozefkowicz et al. 2017). The functional consequence of hetero-tetramerization in plant AQPs has been studied for isoforms in the PIP subfamily, in Xenopus oocytes, yeast cells and plant cell protoplasts, shedding light on intricate AQP-AQP interactions (Fetter et al. 2004; Otto et al. 2010; Zelazny et al. 2007). PIP1 and PIP2 isoforms can be co- or differentially regulated in plants and such PIP physical interactions could present an additional cooperative response to different physiological processes or even stresses. PIP hetero-tetramerization has been implicated in regulation of sub-cellular localisation of PIP1 isoforms. The physical interaction of PIP1 monomers with PIP2 monomers resulted in plasma membrane integration (Zelazny et al. 2007) and altered AQP tetramer transport profiles (Fetter et al. 2004). Otto *et al.* 2010 observed different transport profiles upon co-expression of a tobacco PIP2 and a PIP1 (NtAQP1/NtPIP1;5s) in yeast cells, with the shift in ratio of PIP2 to PIP1 tetramer composition changing from enhanced water to enhanced CO_2 permeabilities, respectively.

In plants, PIP hetero-tetramerization has been proposed as a way of diversification and control of net solute transport. Studies have shown that although PIP1 and PIP2 isoforms are co-expressed, the ratio between protein or mRNA amounts of each subgroup can vary considerably between plant tissues or cell types (Jozefkowicz et al. 2017).

In genetic engineering attempts, the effect that over-expression of a single PIP isoform might have on PIP tetramer interactions should be considered. The tight regulation of PIP isoform interactions within plants could be impacted, which could result in altered tetramer composition and solute permeability within cells. Additionally, the expression of other endogenous PIP isoforms might be altered, as seen in Arabidopsis *AtPIP1;4* and *AtPIP2;5* OE lines which had modulated transcript levels of several other PIP isoforms under osmotic stress (Jang et al. 2007).

In the tobacco PIP OE transgenic lines generated in my PhD research (presented in Chapter 5), we saw that OE of NtPIP2;4s resulted in slight reductions (although not statistically significant) in photosynthetic parameters. Perhaps the constitutive OE of this PIP2 isoform throughout the entire plant might have impacted the stoichiometry of PIP monomers available for tetramer assembly in various plant tissues, and/or changed the expression of other endogenous PIP isoforms in response to the transgene expression within the plant. Further experiments would be required to address these hypotheses.

Overall, a better understanding of AQP-AQP interactions could be advantageous to genetic engineering efforts. AQP multigene constructs could be engineered in plants to favour assembly of a particular tetrameric composition, or alternatively, to improve membrane integration of PIP1 isoforms. Further research could also uncover AQP-AQP interactions within other AQP sub-families.

6.2.4 Targeting AQP expression to specific tissues to avoid pleiotropic effects

AQP gene expression is tightly controlled within plants, with the relationship between AQP expression and their regulation being still unclear due to the large number of AQP isoforms and the number of possible regulatory events (Hachez et al. 2006). At a transcriptional level, AQP isoforms display diverse gene expression patterns within each sub-family (Hachez et al. 2006; Maurel et al. 2008). Our gene expression analysis of the tobacco AQPs (Chapter 2) highlighted the varied gene expression patterns of each gene, with some isoforms being highly targeted to a particular tissue and developmental stage (e.g. young flowers), whilst others more broadly expressed throughout several plant organs.

AQPs have been implicated in cell developmental processes, such as cell enlargement in Arabidopsis roots, hypocotyls, leaves and flower stems (Ludevid et al. 1992). Their expression has also been shown to follow diurnal patterns, such as that of NtAQP1/NtPIP1;5s coinciding with tobacco leaf unfolding throughout the day (Siefritz et al. 2004). In addition to associations with cell development, AQP gene expression has been found to be responsive to drought, low temperatures, salinity,

light, pathogens and plant hormones (Jang et al. 2007; Kawasaki et al. 2001; Maathuis et al. 2003).

Due to the highly targeted and regulated nature of AQP biology within plants, it is likely that constitutive OE of candidate genes might result in pleiotropic effects occurring in the transgenic lines. These could manifest as changes in plant morphology, such as the leaf developmental alterations reported in Hanba *et al.* (2004) upon OE of Barley PIP2;1 in rice. Alternatively, constitutive OE could result in impaired stress response upon exposure to drought of salinity (Aharon et al. 2003; Hanba et al. 2004). Future genetic engineering attempts could focus on localising OE of a candidate AQP isoform to a specific plant tissue. For example, a mesophyll cell-specific promoter could be used to drive expression of a CO_2 permeable PIP isoform in efforts to enhance mesophyll conductance in photosynthetically active tissue, avoiding potential pleiotropic effects that might arise from constitutive expression throughout the entire plant.

6.3 General conclusions

AQPs are central players in diverse plant physiological processes and their roles within plants are dynamic, with cellular AQP composition adjusting depending on the location within the plant, developmental stage and external stimuli. Due to the diversity of AQP isoforms and functional roles, genetic engineering of AQPs presents many challenges. A better understanding of multiple permeating substrates, regulatory processes and AQP-AQP tetramer interactions is necessary for successfully engineering a specific AQP to produce an improved plant. Potentially, AQPs could provide routes to increase crop yields (through enhanced photosynthetic efficiency) and improve resilience to environmental stresses

References

Abascal F, Irisarri I, Zardoya R (2014) Diversity and evolution of membrane intrinsic proteins. Biochimica et Biophysica Acta (BBA)-General Subjects 1840 (5):1468-1481

Afzal Z, Howton T, Sun Y, Mukhtar MS (2016) The roles of aquaporins in plant stress responses. Journal of developmental biology 4 (1):9

Aharon R, Shahak Y, Wininger S, Bendov R, Kapulnik Y, Galili G (2003) Overexpression of a plasma membrane aquaporin in transgenic tobacco improves plant vigor under favorable growth conditions but not under drought or salt stress. The Plant Cell 15 (2):439-447

Alberti S, Gitler AD, Lindquist S (2007) A suite of Gateway® cloning vectors for high-throughput genetic analysis in Saccharomyces cerevisiae. Yeast 24 (10):913-919

Alleva K, Niemietz CM, Sutka M, Maurel C, Parisi M, Tyerman SD, Amodeo G (2006) Plasma membrane of Beta vulgaris storage root shows high water channel activity regulated by cytoplasmic pH and a dual range of calcium concentrations. Journal of experimental botany 57 (3):609-621

Almasalmeh A, Krenc D, Wu B, Beitz E (2014) Structural determinants of the hydrogen peroxide permeability of aquaporins. The FEBS journal 281 (3):647-656

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

Baldocchi D, Valentini R (2004) Geographic and temporal variation of carbon exchange by ecosystems and their sensitivity to environmental perturbations. SCOPE-Scientific Committee On Problems of the Environment International Council of Scientific Unions 62:295-316

Bansal A, Sankararamakrishnan R (2007) Homology modeling of major intrinsic proteins in rice, maize and Arabidopsis: comparative analysis of transmembrane helix association and aromatic/arginine selectivity filters. BMC Structural Biology 7 (1):27

Bárzana G, Aroca R, Bienert GP, Chaumont F, Ruiz-Lozano JM (2014) New insights into the regulation of aquaporins by the arbuscular mycorrhizal symbiosis in maize plants under drought stress and possible implications for plant performance. Molecular Plant-Microbe Interactions 27 (4):349-363

Berny MC, Gilis D, Rooman M, Chaumont F (2016) Single mutations in the transmembrane domains of maize plasma membrane aquaporins affect the activity of monomers within a heterotetramer. Molecular plant 9 (7):986-1003

Bertl A, Anderson JA, Slayman CL, Gaber RF (1995) Use of Saccharomyces cerevisiae for patch-clamp analysis of heterologous membrane proteins: characterization of Kat1, an inward-rectifying K+ channel from Arabidopsis thaliana, and comparison with endogeneous yeast channels and carriers. Proceedings of the National Academy of Sciences 92 (7):2701-2705

Bi Z, Merl-Pham J, Uehlein N, Zimmer I, Mühlhans S, Aichler M, Walch AK, Kaldenhoff R, Palme K, Schnitzler J-P (2015) RNAi-mediated downregulation of poplar plasma membrane intrinsic proteins (PIPs) changes plasma membrane proteome composition and affects leaf physiology. Journal of proteomics 128:321-332

Biela A, Grote K, Otto B, Hoth S, Hedrich R, Kaldenhoff R (1999) The Nicotiana tabacum plasma membrane aquaporin NtAQP1 is mercury-insensitive and permeable for glycerol. The Plant Journal 18 (5):565-570

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 (2):306-317

Bienert GP, Cavez D, Besserer A, Berny MC, Gilis D, Rooman M, Chaumont F (2012) A conserved cysteine residue is involved in disulfide bond formation between plant plasma membrane aquaporin monomers. Biochemical journal 445 (1):101-111

Bienert GP, Chaumont F (2014) Aquaporin-facilitated transmembrane diffusion of hydrogen peroxide. Biochimica et Biophysica Acta (BBA)-General Subjects 1840 (5):1596-1604

Bienert GP, Desguin B, Chaumont F, Hols P (2013) Channel-mediated lactic acid transport: a novel function for aquaglyceroporins in bacteria. Biochemical Journal 454 (3):559-570

Bienert GP, Møller AL, Kristiansen KA, Schulz A, Møller IM, Schjoerring JK, Jahn TP (2007) Specific aquaporins facilitate the diffusion of hydrogen peroxide across membranes. Journal of Biological Chemistry 282 (2):1183-1192

Bienert GP, Schüssler MD, Jahn TP (2008) Metalloids: essential, beneficial or toxic? Major intrinsic proteins sort it out. Trends in biochemical sciences 33 (1):20-26

Blom N, Gammeltoft S, Brunak S (1999) Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. Journal of molecular biology 294 (5):1351-1362

Blomberg A (2011) Measuring growth rate in high-throughput growth phenotyping. Current opinion in biotechnology 22 (1):94-102

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30 (15):2114-2120

Bots M, Feron R, Uehlein N, Weterings K, Kaldenhoff R, Mariani T (2005) PIP1 and PIP2 aquaporins are differentially expressed during tobacco anther and stigma development. Journal of Experimental Botany 56 (409):113-121

Briesemeister S, Rahnenfï¿ ½hrer Jr, Kohlbacher O (2010) YLoc—an interpretable web server for predicting subcellular localization. Nucleic acids research 38 (suppl_2):W497-W502

Byrt CS, Zhao M, Kourghi M, Bose J, Henderson SW, Qiu J, Gilliham M, Schultz C, Schwarz M, Ramesh SA (2017) Non-selective cation channel activity of aquaporin AtPIP2; 1 regulated by Ca2+ and pH. Plant, cell & environment 40 (6):802-815

Chattopadhyay D, Tyagi AK, Sato S, Zamir D, Giuliano G, Consortium TG (2012) The tomato genome sequence provides insights into fleshy fruit evolution.

Chevalier AS, Chaumont F (2014) Trafficking of plant plasma membrane aquaporins: multiple regulation levels and complex sorting signals. Plant and Cell Physiology 56 (5):819-829

Choi W-G, Roberts DM (2007) Arabidopsis NIP2; 1, a major intrinsic protein transporter of lactic acid induced by anoxic stress. Journal of Biological Chemistry 282 (33):24209-24218

Chou K-C, Shen H-B (2010) Plant-mPLoc: a top-down strategy to augment the power for predicting plant protein subcellular localization. PloS one 5 (6):e11335

Chrispeels MJ, Crawford NM, Schroeder JI (1999) Proteins for transport of water and mineral nutrients across the membranes of plant cells. The Plant Cell 11 (4):661-675

Consortium PGS (2011) Genome sequence and analysis of the tuber crop potato. Nature 475 (7355):189

Cooper TG (1982) Nitrogen metabolism in Saccharomyces cerevisiae. The molecular biology of the yeast Saccharomyces: metabolism and gene expression 2:39-99

Cui X-H, Hao F-S, Chen H, Chen J, Wang X-C (2008) Expression of the Vicia faba VfPIP1 gene in Arabidopsis thaliana plants improves their drought resistance. Journal of plant research 121 (2):207-214

Curtis MD, Grossniklaus U (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant physiology 133 (2):462-469

Danielson JÅ, Johanson U (2008) Unexpected complexity of the aquaporin gene family in the moss Physcomitrella patens. BMC Plant Biology 8 (1):45

De Rosa A, Watson-Lazowski A, Evans JR, Groszmann M (2020) Genome-wide identification and characterisation of Aquaporins in Nicotiana tabacum and their relationships with other Solanaceae species. BMC Plant Biology 20 (1):1-29

Deshmukh RK, Sonah H, Bélanger RR (2016) Plant Aquaporins: genome-wide identification, transcriptomics, proteomics, and advanced analytical tools. Frontiers in plant science 7:1896

Di Giorgio JP, Bienert GP, Ayub N, Yaneff A, Barberini ML, Mecchia MA, Amodeo G, Soto G, MUSCHIETTI JP (2016) Pollen-specific aquaporins NIP4; 1 and NIP4; 2 are required for pollen development and pollination in Arabidopsis thaliana. The Plant Cell:tpc. 00776.02015

Diehn TA, Pommerrenig B, Bernhardt N, Hartmann A, Bienert GP (2015) Genome-wide identification of aquaporin encoding genes in Brassica oleracea and their phylogenetic sequence comparison to Brassica crops and Arabidopsis. Frontiers in plant science 6:166

Dreyer I, Horeau C, Lemaillet G, Zimmermann S, Bush DR, Rodríguez-Navarro A, Schachtman DP, Spalding EP, Sentenac H, Gaber RF (1999) Identification and characterization of plant transporters using heterologous expression systems. Journal of Experimental Botany:1073-1087

Drummond A, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, Field M, Heled J, Kearse M, Markowitz S (2011) Geneious. Biomatters Ltd

Dynowski M, Mayer M, Moran O, Ludewig U (2008) Molecular determinants of ammonia and urea conductance in plant aquaporin homologs. FEBS letters 582 (16):2458-2462

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research 32 (5):1792-1797

Edwards K, Fernandez-Pozo N, Drake-Stowe K, Humphry M, Evans A, Bombarely A, Allen F, Hurst R, White B, Kernodle S (2017) A reference genome for Nicotiana tabacum enables map-based cloning of homeologous loci implicated in nitrogen utilization efficiency. BMC genomics 18 (1):448

Engelman DM (2005) Membranes are more mosaic than fluid. Nature 438 (7068):578-580

Esau K (1965) Plant anatomy. Plant Anatomy (2nd Edition)

Evans J, Sharkey T, Berry J, Farquhar G (1986) Carbon isotope discrimination measured concurrently with gas exchange to investigate CO2 diffusion in leaves of higher plants. Functional Plant Biology 13 (2):281-292

Evans JR, Caemmerer S, Setchell BA, Hudson GS (1994) The relationship between CO2 transfer conductance and leaf anatomy in transgenic tobacco with a reduced content of Rubisco. Functional Plant Biology 21 (4):475-495

Evans JR, Kaldenhoff R, Genty B, Terashima I (2009) Resistances along the CO2 diffusion pathway inside leaves. Journal of Experimental Botany 60 (8):2235-2248

Evans JR, Loreto F (2000) Acquisition and diffusion of CO2 in higher plant leaves. In: Photosynthesis. Springer, pp 321-351

Evans JR, Von Caemmerer S (2013) Temperature response of carbon isotope discrimination and mesophyll conductance in tobacco. Plant, Cell & Environment 36 (4):745-756

FAO (2009) How to feed the world in 2050. Food and Agriculture Organisation, Rome.

Farquhar G, Richards R (1984) Isotopic composition of plant carbon correlates with water-use efficiency of wheat genotypes. Functional Plant Biology 11 (6):539-552

Farquhar GD, Cernusak LA (2012) Ternary effects on the gas exchange of isotopologues of carbon dioxide. Plant, Cell & Environment 35 (7):1221-1231

Farquhar GD, O'Leary MH, Berry JA (1982) On the relationship between carbon isotope discrimination and the intercellular carbon dioxide concentration in leaves. Functional Plant Biology 9 (2):121-137

Fernandez-Pozo N, Menda N, Edwards JD, Saha S, Tecle IY, Strickler SR, Bombarely A, Fisher-York T, Pujar A, Foerster H (2014) The Sol Genomics Network (SGN)—from genotype to phenotype to breeding. Nucleic acids research 43 (D1):D1036-D1041

Fetter K, Van Wilder V, Moshelion M, Chaumont F (2004) Interactions between plasma membrane aquaporins modulate their water channel activity. The Plant Cell 16 (1):215-228

Fischer R, Byerlee D, Edmeades G (2014) Crop yields and global food security. ACIAR: Canberra, ACT:8-11

Flexas J, Barbour MM, Brendel O, Cabrera HM, Carriquí M, Diaz-Espejo A, Douthe C, Dreyer E, Ferrio JP, Gago J (2012) Mesophyll diffusion conductance to CO2: an unappreciated central player in photosynthesis. Plant Science 193:70-84

Flexas J, Ribas-Carbó M, Hanson DT, Bota J, Otto B, Cifre J, McDowell N, Medrano H, Kaldenhoff R (2006) Tobacco aquaporin NtAQP1 is involved in mesophyll conductance to CO2in vivo. The Plant Journal 48 (3):427-439

Footitt S, Clewes R, Feeney M, Finch-Savage WE, Frigerio L (2019) Aquaporins influence seed dormancy and germination in response to stress. Plant, cell & environment

Fox AR, Maistriaux LC, Chaumont F (2017) Toward understanding of the high number of plant aquaporin isoforms and multiple regulation mechanisms. Plant Science 264:179-187

Frick A, Järvå M, Ekvall M, Uzdavinys P, Nyblom M, Törnroth-Horsefield S (2013) Mercury increases water permeability of a plant aquaporin through a non-cysteine-related mechanism. Biochemical Journal 454 (3):491-499

Froger A, Thomas D, Delamarche C, Tallur B (1998) Prediction of functional residues in water channels and related proteins. Protein Science 7 (6):1458-1468

Frommer WB, Hummel S, Riesmeier JW (1993) Expression cloning in yeast of a cDNA encoding a broad specificity amino acid permease from Arabidopsis thaliana. Proceedings of the National Academy of Sciences 90 (13):5944-5948

Gambetta GA, Knipfer T, Fricke W, McElrone AJ (2017) Aquaporins and root water uptake. In: Plant aquaporins. Springer, pp 133-153

Gao J, Yu H, Song Q, Li X (2005) Establishment of HEK293 cell line expressing green fluorescent protein–aquaporin-1 to determine osmotic water permeability. Analytical biochemistry 342 (1):53-58

Gao L, Lu Z, Ding L, Guo J, Wang M, Ling N, Guo S, Shen Q (2018) Role of aquaporins in determining carbon and nitrogen status in higher plants. International journal of molecular sciences 19 (1):35

Gebhardt C (2016) The historical role of species from the Solanaceae plant family in genetic research. Theoretical and Applied Genetics 129 (12):2281-2294

Georges F, Ray H (2017) Genome editing of crops: a renewed opportunity for food security. GM crops & food 8 (1):1-12

Godard P, Urrestarazu A, Vissers S, Kontos K, Bontempi G, van Helden J, André B (2007) Effect of 21 different nitrogen sources on global gene expression in the yeast Saccharomyces cerevisiae. Molecular and cellular biology 27 (8):3065-3086

Gomes D, Agasse A, Thiébaud P, Delrot S, Gerós H, Chaumont F (2009) Aquaporins are multifunctional water and solute transporters highly divergent in living organisms. Biochimica et Biophysica Acta (BBA)-Biomembranes 1788 (6):1213-1228

Griffis TJ, Baker J, Sargent S, Tanner B, Zhang J (2004) Measuring field-scale isotopic CO2 fluxes with tunable diode laser absorption spectroscopy and micrometeorological techniques. Agricultural and Forest Meteorology 124 (1-2):15-29

Groszmann M, Osborn HL, Evans JR (2017) Carbon dioxide and water transport through plant aquaporins. Plant, Cell & Environment 40 (6):938-961

Guenther JF, Chanmanivone N, Galetovic MP, Wallace IS, Cobb JA, Roberts DM (2003) Phosphorylation of soybean nodulin 26 on serine 262 enhances water permeability and is regulated developmentally and by osmotic signals. The Plant Cell 15 (4):981-991

Guex N, Peitsch MC, Schwede T (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis 30 (S1):S162-S173

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 9 (1):134

Hachez C, Heinen RB, Draye X, Chaumont F (2008) The expression pattern of plasma membrane aquaporins in maize leaf highlights their role in hydraulic regulation. Plant molecular biology 68 (4-5):337

Hachez C, Zelazny E, Chaumont F (2006) Modulating the expression of aquaporin genes in planta: a key to understand their physiological functions? Biochimica et Biophysica Acta (BBA)-Biomembranes 1758 (8):1142-1156

Hall BG, Acar H, Nandipati A, Barlow M (2013) Growth rates made easy. Molecular biology and evolution 31 (1):232-238

Halliwell B, Gutteridge JM (2015) Free radicals in biology and medicine. Oxford University Press, USA,

Hanaoka H, Uraguchi S, Takano J, Tanaka M, Fujiwara T (2014) OsNIP3; 1, a rice boric acid channel, regulates boron distribution and is essential for growth under boron-deficient conditions. The Plant Journal 78 (5):890-902

Hanba YT, Shibasaka M, Hayashi Y, Hayakawa T, Kasamo K, Terashima I, Katsuhara M (2004) Overexpression of the barley aquaporin HvPIP2; 1 increases internal CO2 conductance and CO2 assimilation in the leaves of transgenic rice plants. Plant and Cell Physiology 45 (5):521-529

Heckwolf M, Pater D, Hanson DT, Kaldenhoff R (2011) The Arabidopsis thaliana aquaporin AtPIP1; 2 is a physiologically relevant CO2 transport facilitator. The Plant Journal 67 (5):795-804

Hedrich R, Marten I (2006) 30-year progress of membrane transport in plants. Planta 224 (4):725-739

Heinen RB, Ye Q, Chaumont F (2009) Role of aquaporins in leaf physiology. Journal of experimental botany 60 (11):2971-2985

Hohmann S, Bill RM, Kayingo G, Prior BA (2000) Microbial MIP channels. Trends in microbiology 8 (1):33-38

Hooijmaijers C, Rhee JY, Kwak KJ, Chung GC, Horie T, Katsuhara M, Kang H (2012) Hydrogen peroxide permeability of plasma membrane aquaporins of Arabidopsis thaliana. Journal of plant research 125 (1):147-153

Horton P, Park K-J, Obayashi T, Fujita N, Harada H, Adams-Collier C, Nakai K (2007) WoLF PSORT: protein localization predictor. Nucleic acids research 35 (suppl_2):W585-W587

Hove RM, Bhave M (2011) Plant aquaporins with non-aqua functions: deciphering the signature sequences. Plant Molecular Biology 75 (4):413-430. doi:10.1007/s11103-011-9737-5

Hub JS, Grubmüller H, De Groot BL (2009) Dynamics and energetics of permeation through aquaporins. What do we learn from molecular dynamics simulations? In: Aquaporins. Springer, pp 57-76

Ishikawa F, Suga S, Uemura T, Sato MH, Maeshima M (2005) Novel type aquaporin SIPs are mainly localized to the ER membrane and show cell-specific expression in Arabidopsis thaliana. FEBS letters 579 (25):5814-5820

Jang JY, Lee SH, Rhee JY, Chung GC, Ahn SJ, Kang H (2007) Transgenic Arabidopsis and tobacco plants overexpressing an aquaporin respond differently to various abiotic stresses. Plant molecular biology 64 (6):621-632

Johanson U, Gustavsson S (2002) A new subfamily of major intrinsic proteins in plants. Molecular Biology and Evolution 19 (4):456-461

Johansson I, Karlsson M, Shukla VK, Chrispeels MJ, Larsson C, Kjellbom P (1998) Water transport activity of the plasma membrane aquaporin PM28A is regulated by phosphorylation. The Plant Cell 10 (3):451-459

Jozefkowicz C, Berny MC, Chaumont F, Alleva K (2017) Heteromerization of plant aquaporins. In: Plant Aquaporins. Springer, pp 29-46

Kaldenhoff R, Bertl A, Otto B, Moshelion M, Uehlein N (2007) Characterization of plant aquaporins. In: Methods in enzymology, vol 428. Elsevier, pp 505-531

Kaldenhoff R, Fischer M (2006) Aquaporins in plants. Acta Physiologica 187 (1-2):169-176

Kawasaki S, Borchert C, Deyholos M, Wang H, Brazille S, Kawai K, Galbraith D, Bohnert HJ (2001) Gene expression profiles during the initial phase of salt stress in rice. The Plant Cell 13 (4):889-905

Kawase M, Hanba YT, Katsuhara M (2013) The photosynthetic response of tobacco plants overexpressing ice plant aquaporin McMIPB to a soil water deficit and high vapor pressure deficit. Journal of plant research 126 (4):517-527

Khush GS (2001) Green revolution: the way forward. Nature reviews genetics 2 (10):815

Kirscht A, Kaptan SS, Bienert GP, Chaumont F, Nissen P, de Groot BL, Kjellbom P, Gourdon P, Johanson U (2016) Crystal structure of an ammonia-permeable aquaporin. PLoS biology 14 (3):e1002411

Kojima S, Bohner A, Gassert B, Yuan L, Wirén Nv (2007) AtDUR3 represents the major transporter for high-affinity urea transport across the plasma membrane of nitrogen-deficient Arabidopsis roots. The Plant Journal 52 (1):30-40

Kojima S, Bohner A, Von Wirén N (2006) Molecular mechanisms of urea transport in plants. The Journal of membrane biology 212 (2):83-91

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular biology and evolution 33 (7):1870-1874

Le Cahérec F, Bron P, Verbavatz J, Garret A, Morel G, Cavalier A, Bonnec G, Thomas D, Gouranton J, Hubert J (1996) Incorporation of proteins into (Xenopus) oocytes by proteoliposome microinjection: functional characterization of a novel aquaporin. Journal of cell science 109 (6):1285-1295

Lee J, Godon C, Lagniel G, Spector D, Garin J, Labarre J, Toledano MB (1999) Yap1 and Skn7 control two specialized oxidative stress response regulons in yeast. Journal of Biological Chemistry 274 (23):16040-16046

Leitão L, Prista C, Moura TF, Loureiro-Dias MC, Soveral G (2012) Grapevine aquaporins: gating of a tonoplast intrinsic protein (TIP2; 1) by cytosolic pH. PLoS One 7 (3):e33219

Leitch I, Bennett M (2004) Genome downsizing in polyploid plants. Biological journal of the Linnean society 82 (4):651-663

Leitch I, Hanson L, Lim K, Kovarik A, Chase M, Clarkson J, Leitch A (2008) The ups and downs of genome size evolution in polyploid species of Nicotiana (Solanaceae). Annals of Botany 101 (6):805-814

Li G-W, Peng Y-H, Yu X, Zhang M-H, Cai W-M, Sun W-N, Su W-A (2008a) Transport functions and expression analysis of vacuolar membrane aquaporins in response to various stresses in rice. Journal of plant physiology 165 (18):1879-1888

Li G-W, Zhang M-H, Cai W-M, Sun W-N, Su W-A (2008b) Characterization of OsPIP2; 7, a water channel protein in rice. Plant and cell physiology 49 (12):1851-1858

Li G, Santoni V, Maurel C (2014) Plant aquaporins: roles in plant physiology. Biochimica et Biophysica Acta (BBA)-General Subjects 1840 (5):1574-1582

Lian H-L, Yu X, Ye Q, Ding X-S, Kitagawa Y, Kwak S-S, Su W-A, Tang Z-C (2004) The role of aquaporin RWC3 in drought avoidance in rice. Plant and Cell Physiology 45 (4):481-489

Liu C, Fukumoto T, Matsumoto T, Gena P, Frascaria D, Kaneko T, Katsuhara M, Zhong S, Sun X, Zhu Y (2013) Aquaporin OsPIP1; 1 promotes rice salt resistance and seed germination. Plant Physiology and Biochemistry 63:151-158

Liu L-H, Ludewig U, Gassert B, Frommer WB, von Wirén N (2003) Urea transport by nitrogen-regulated tonoplast intrinsic proteins in Arabidopsis. Plant physiology 133 (3):1220-1228

Long SP, ZHU XG, Naidu SL, Ort DR (2006) Can improvement in photosynthesis increase crop yields? Plant, cell & environment 29 (3):315-330

Lopez F, Bousser A, Sissoëff I, Hoarau J, Mahé A (2004) Characterization in maize of ZmTIP2-3, a root-specific tonoplast intrinsic protein exhibiting aquaporin activity. Journal of experimental botany 55 (396):539-541

Loqué D, Ludewig U, Yuan L, von Wirén N (2005) Tonoplast intrinsic proteins AtTIP2; 1 and AtTIP2; 3 facilitate NH3 transport into the vacuole. Plant physiology 137 (2):671-680

Luang S, Hrmova M (2017) Structural basis of the permeation function of plant aquaporins. In: Plant Aquaporins. Springer, pp 1-28

Ludevid D, Höfte H, Himelblau E, Chrispeels MJ (1992) The expression pattern of the tonoplast intrinsic protein γ-TIP in Arabidopsis thaliana is correlated with cell enlargement. Plant physiology 100 (4):1633-1639

Luu DT, Maurel C (2013) Aquaporin trafficking in plant cells: an emerging membrane-protein model. Traffic 14 (6):629-635

Ma JKC, Drossard J, Lewis D, Altmann F, Boyle J, Christou P, Cole T, Dale P, van Dolleweerd CJ, Isitt V (2015) Regulatory approval and a first-in-human phase I clinical trial of a

monoclonal antibody produced in transgenic tobacco plants. Plant biotechnology journal 13 (8):1106-1120

Ma N, Xue J, Li Y, Liu X, Dai F, Jia W, Luo Y, Gao J (2008) Rh-PIP2; 1, a rose aquaporin gene, is involved in ethylene-regulated petal expansion. Plant physiology 148 (2):894-907

Maathuis FJ, Filatov V, Herzyk P, C. Krijger G, B. Axelsen K, Chen S, Green BJ, Li Y, Madagan KL, Sánchez-Fernández R (2003) Transcriptome analysis of root transporters reveals participation of multiple gene families in the response to cation stress. The Plant Journal 35 (6):675-692

Maeshima M, Ishikawa F (2008) ER membrane aquaporins in plants. Pflügers Archiv-European Journal of Physiology 456 (4):709-716

Mao Z, Sun W (2015) Arabidopsis seed-specific vacuolar aquaporins are involved in maintaining seed longevity under the control of ABSCISIC ACID INSENSITIVE 3. Journal of experimental botany 66 (15):4781-4794

Marschner H (2011) Marschner's mineral nutrition of higher plants. Academic press,

Martins CdPS, Pedrosa AM, Du D, Goncalves LP, Yu Q, Gmitter Jr FG, Costa MGC (2015) Genome-wide characterization and expression analysis of major intrinsic proteins during abiotic and biotic stresses in sweet orange (Citrus sinensis L. Osb.). PLoS one 10 (9):e0138786

Masood TB, Sandhya S, Chandra N, Natarajan V (2015) CHEXVIS: a tool for molecular channel extraction and visualization. BMC bioinformatics 16 (1):119

Maurel C, Boursiac Y, Luu D-T, Santoni V, Shahzad Z, Verdoucq L (2015) Aquaporins in plants. Physiological reviews 95 (4):1321-1358

Maurel C, Reizer J, Schroeder JI, Chrispeels MJ (1993) The vacuolar membrane protein gamma-TIP creates water specific channels in Xenopus oocytes. The EMBO Journal 12 (6):2241-2247

Maurel C, Verdoucq L, Luu D-T, Santoni V (2008) Plant aquaporins: membrane channels with multiple integrated functions. Annu Rev Plant Biol 59:595-624

Mitani-Ueno N, Yamaji N, Zhao F-J, 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 (12):4391-4398

Mitani N, Yamaji N, Ma JF (2008) Characterization of substrate specificity of a rice silicon transporter, Lsi1. Pflügers Archiv-European Journal of Physiology 456 (4):679-686

Mizutani M, Watanabe S, Nakagawa T, Maeshima M (2006) Aquaporin NIP2; 1 is mainly localized to the ER membrane and shows root-specific accumulation in Arabidopsis thaliana. Plant and cell physiology 47 (10):1420-1426

Mori IC, Rhee J, Shibasaka M, Sasano S, Kaneko T, Horie T, Katsuhara M (2014) CO2 transport by PIP2 aquaporins of barley. Plant and Cell Physiology 55 (2):251-257

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 (6804):599

Nelson BK, Cai X, Nebenführ A (2007) A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. The Plant Journal 51 (6):1126-1136

Newby ZE, O'Connell Iii J, Robles-Colmenares Y, Khademi S, Miercke LJ, Stroud RM (2008) Crystal structure of the aquaglyceroporin PfAQP from the malarial parasite Plasmodium falciparum. Nature structural & molecular biology 15 (6):619

Nyblom M, Frick A, Wang Y, Ekvall M, Hallgren K, Hedfalk K, Neutze R, Tajkhorshid E, Törnroth-Horsefield S (2009) Structural and functional analysis of SoPIP2; 1 mutants adds insight into plant aquaporin gating. Journal of molecular biology 387 (3):653-668

Otto B, Uehlein N, Sdorra S, Fischer M, Ayaz M, Belastegui-Macadam X, Heckwolf M, Lachnit M, Pede N, Priem N (2010) Aquaporin tetramer composition modifies the function of tobacco aquaporins. Journal of Biological Chemistry 285 (41):31253-31260

Parry MA, Hawkesford MJ (2010) Food security: increasing yield and improving resource use efficiency. Proceedings of the nutrition Society 69 (4):592-600

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C (2017) Salmon provides fast and bias-aware quantification of transcript expression. Nature methods 14 (4):417

Péret B, Li G, Zhao J, Band LR, Voß U, Postaire O, Luu D-T, Da Ines O, Casimiro I, Lucas M (2012) Auxin regulates aquaporin function to facilitate lateral root emergence. Nature cell biology 14 (10):991

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—a visualization system for exploratory research and analysis. Journal of computational chemistry 25 (13):1605-1612

Pettersson N, Hagström J, Bill RM, Hohmann S (2006) Expression of heterologous aquaporins for functional analysis in Saccharomyces cerevisiae. Current genetics 50 (4):247-255

Pommerrenig B, Diehn TA, Bienert GP (2015) Metalloido-porins: Essentiality of Nodulin 26-like intrinsic proteins in metalloid transport. Plant Science 238:212-227

Prak S, Hem S, Boudet J, Viennois G, Sommerer N, Rossignol M, Maurel C, Santoni V (2008) Multiple phosphorylations in the C-terminal tail of plant plasma membrane aquaporins: role in subcellular trafficking of AtPIP2; 1 in response to salt stress. Molecular & Cellular Proteomics 7 (6):1019-1030

Ray DK, Mueller ND, West PC, Foley JA (2013) Yield trends are insufficient to double global crop production by 2050. PloS one 8 (6):e66428

Ray DK, Ramankutty N, Mueller ND, West PC, Foley JA (2012) Recent patterns of crop yield growth and stagnation. Nature communications 3:1293

Reichel M, Liao Y, Rettel M, Ragan C, Evers M, Alleaume A-M, Horos R, Hentze MW, Preiss T, Millar AA (2016) In planta determination of the mRNA-binding proteome of Arabidopsis etiolated seedlings. The Plant Cell 28 (10):2435-2452

Reinhardt H, Hachez C, Bienert MD, Beebo A, Swarup K, Voß U, Bouhidel K, Frigerio L, Schjoerring JK, Bennett MJ (2016) Tonoplast aquaporins facilitate lateral root emergence. Plant physiology 170 (3):1640-1654

Reisen D, Leborgne-Castel N, Özalp C, Chaumont F, Marty F (2003) Expression of a cauliflower tonoplast aquaporin tagged with GFP in tobacco suspension cells correlates with an increase in cell size. Plant molecular biology 52 (2):387-400

Reuscher S, Akiyama M, Mori C, Aoki K, Shibata D, Shiratake K (2013) Genome-wide identification and expression analysis of aquaporins in tomato. PLoS One 8 (11):e79052

Riesmeier JW, Willmitzer L, Frommer WB (1992) Isolation and characterization of a sucrose carrier cDNA from spinach by functional expression in yeast. The EMBO Journal 11 (13):4705-4713

Roche JV, Törnroth-Horsefield S (2017) Aquaporin Protein-Protein Interactions. International journal of molecular sciences 18 (11):2255

Roeske C, O'Leary MH (1984) Carbon isotope effects on enzyme-catalyzed carboxylation of ribulose bisphosphate. Biochemistry 23 (25):6275-6284

Sade N, Gallé A, Flexas J, Lerner S, Peleg G, Yaaran A, Moshelion M (2014) Differential tissue-specific expression of NtAQP1 in Arabidopsis thaliana reveals a role for this protein in stomatal and mesophyll conductance of CO2 under standard and salt-stress conditions. Planta 239 (2):357-366

Sade N, Gebretsadik M, Seligmann R, Schwartz A, Wallach R, Moshelion M (2010) The role of tobacco Aquaporin1 in improving water use efficiency, hydraulic conductivity, and yield production under salt stress. Plant physiology 152 (1):245-254

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 (3):651-661

Santoni V (2017) Plant aquaporin posttranslational regulation. In: Plant Aquaporins. Springer, pp 83-105

Sato S, Tabata S, Hirakawa H, Asamizu E, Shirasawa K, Isobe S, Kaneko T, Nakamura Y, Shibata D, Egholm M (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485 (7400):635

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B (2012) Fiji: an open-source platform for biological-image analysis. Nature methods 9 (7):676

Schmidt M, Schaumberg JZ, Steen CM, Boyer MP (2010) Boric acid disturbs cell wall synthesis in Saccharomyces cerevisiae. International journal of microbiology 2010

Schnurbusch T, Hayes J, Hrmova M, Baumann U, Ramesh SA, Tyerman SD, Langridge P, Sutton T (2010) Boron toxicity tolerance in barley through reduced expression of the multifunctional aquaporin HvNIP2; 1. Plant Physiology 153 (4):1706-1715

Shah NA, Laws RJ, Wardman B, Zhao LP, Hartman JL (2007) Accurate, precise modeling of cell proliferation kinetics from time-lapse imaging and automated image analysis of agar yeast culture arrays. BMC systems biology 1 (1):3

Siefritz F, Otto B, Bienert GP, Van Der Krol A, Kaldenhoff R (2004) The plasma membrane aquaporin NtAQP1 is a key component of the leaf unfolding mechanism in tobacco. The Plant Journal 37 (2):147-155

Siefritz F, Tyree MT, Lovisolo C, Schubert A, Kaldenhoff R (2002) PIP1 plasma membrane aquaporins in tobacco from cellular effects to function in plants. The Plant Cell 14 (4):869-876

Sierro N, Battey JN, Ouadi S, Bakaher N, Bovet L, Willig A, Goepfert S, Peitsch MC, Ivanov NV (2014) The tobacco genome sequence and its comparison with those of tomato and potato. Nature communications 5

Sierro N, Battey JN, Ouadi S, Bovet L, Goepfert S, Bakaher N, Peitsch MC, Ivanov NV (2013) Reference genomes and transcriptomes of Nicotiana sylvestris and Nicotiana tomentosiformis. Genome Biology 14 (6):R60. doi:10.1186/gb-2013-14-6-r60

Sonah H, Deshmukh RK, Labbé C, Bélanger RR (2017) Analysis of aquaporins in Brassicaceae species reveals high-level of conservation and dynamic role against biotic and abiotic stress in canola. Scientific reports 7 (1):2771

Soto G, Fox R, Ayub N, Alleva K, Guaimas F, Erijman EJ, Mazzella A, Amodeo G, Muschietti J (2010) TIP5; 1 is an aquaporin specifically targeted to pollen mitochondria and is probably involved in nitrogen remobilization in Arabidopsis thaliana. The Plant Journal 64 (6):1038-1047

Sui H, Han B-G, Lee JK, Walian P, Jap BK (2001) Structural basis of water-specific transport through the AQP1 water channel. Nature 414 (6866):872

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. The Plant Cell 18 (6):1498-1509

Takano J, Yoshinari A, Luu D-T (2017) Plant aquaporin trafficking. In: Plant Aquaporins. Springer, pp 47-81

Tanaka M, Wallace IS, Takano J, Roberts DM, Fujiwara T (2008) NIP6; 1 is a boric acid channel for preferential transport of boron to growing shoot tissues in Arabidopsis. The Plant Cell 20 (10):2860-2875

Tanghe A, Van Dijck P, Dumortier F, Teunissen A, Hohmann S, Thevelein JM (2002) Aquaporin expression correlates with freeze tolerance in baker's yeast, and overexpression improves freeze tolerance in industrial strains. Appl Environ Microbiol 68 (12):5981-5989

Tazoe Y, Von Caemmerer S, Estavillo GM, Evans JR (2011) Using tunable diode laser spectroscopy to measure carbon isotope discrimination and mesophyll conductance to CO2 diffusion dynamically at different CO2 concentrations. Plant, Cell & Environment 34 (4):580-591

Terashima I, Hanba YT, Tholen D, Niinemets Ü (2011) Leaf functional anatomy in relation to photosynthesis. Plant Physiology 155 (1):108-116

Tholen D, Zhu X-G (2011) The mechanistic basis of internal conductance: a theoretical analysis of mesophyll cell photosynthesis and CO2 diffusion. Plant Physiology 156 (1):90-105

Tian S, Wang X, Li P, Wang H, Ji H, Xie J, Qiu Q, Shen D, Dong H (2016) Plant aquaporin AtPIP1; 4 links apoplastic H2O2 induction to disease immunity pathways. Plant physiology 171 (3):1635-1650

Törnroth-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 (7077):688

Tournaire-Roux C, Sutka M, Javot H, Gout E, Gerbeau P, Luu D-T, Bligny R, Maurel C (2003) Cytosolic pH regulates root water transport during anoxic stress through gating of aquaporins. Nature 425 (6956):393

Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A (2015) The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. Nucleic acids research:gkv485

Tusé D, Tu T, McDonald KA (2014) Manufacturing economics of plant-made biologics: case studies in therapeutic and industrial enzymes. BioMed research international 2014

Uehlein N, Lovisolo C, Siefritz F, Kaldenhoff R (2003) The tobacco aquaporin NtAQP1 is a membrane CO2 pore with physiological functions. Nature 425 (6959):734-737

Uehlein N, Otto B, Hanson DT, Fischer M, McDowell N, Kaldenhoff R (2008) Function of Nicotiana tabacum aquaporins as chloroplast gas pores challenges the concept of membrane CO2 permeability. The Plant Cell 20 (3):648-657

Uehlein N, Sperling H, Heckwolf M, Kaldenhoff R (2012) The Arabidopsis aquaporin PIP1; 2 rules cellular CO2 uptake. Plant, cell & environment 35 (6):1077-1083

Umenishi F, Verbavatz J-M, Verkman A (2000) cAMP regulated membrane diffusion of a green fluorescent protein-aquaporin 2 chimera. Biophysical journal 78 (2):1024-1035

Utsugi S, Shibasaka M, Maekawa M, Katsuhara M (2015) Control of the water transport activity of barley HvTIP3; 1 specifically expressed in seeds. Plant and Cell Physiology 56 (9):1831-1840

Vanhercke T, El Tahchy A, Liu Q, Zhou XR, Shrestha P, Divi UK, Ral JP, Mansour MP, Nichols PD, James CN (2014) Metabolic engineering of biomass for high energy density: oilseed-like triacylglycerol yields from plant leaves. Plant biotechnology journal 12 (2):231-239

Venkatesh J, Yu J-W, Gaston D, Park SW (2015) Molecular evolution and functional divergence of X-intrinsic protein genes in plants. Molecular Genetics and Genomics 290 (2):443-460

Venkatesh J, Yu J-W, Park SW (2013) Genome-wide analysis and expression profiling of the Solanum tuberosum aquaporins. Plant physiology and biochemistry 73:392-404

Verkman AS, Anderson MO, Papadopoulos MC (2014) Aquaporins: important but elusive drug targets. Nature reviews Drug discovery 13 (4):259

von Caemmerer S, Evans JR (1991) Determination of the average partial pressure of CO2 in chloroplasts from leaves of several C3 plants. Functional Plant Biology 18 (3):287-305

von Caemmerer S, Evans JR (2010) Enhancing C_3 photosynthesis. Plant Physiology 154 (2):589-592. doi:10.1104/pp.110.160952

von Caemmerer S, Evans JR, Hudson GS, Andrews TJ (1994) The kinetics of ribulose-1, 5-bisphosphate carboxylase/oxygenase in vivo inferred from measurements of photosynthesis in leaves of transgenic tobacco. Planta 195 (1):88-97

von Caemmerer Sv, Farquhar GD (1981) Some relationships between the biochemistry of photosynthesis and the gas exchange of leaves. Planta 153 (4):376-387

Wallace IS, Choi W-G, Roberts DM (2006) The structure, function and regulation of the nodulin 26-like intrinsic protein family of plant aquaglyceroporins. Biochimica et Biophysica Acta (BBA)-Biomembranes 1758 (8):1165-1175

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 (2):1059-1068

Wang M, Ding L, Gao L, Li Y, Shen Q, Guo S (2016) The interactions of aquaporins and mineral nutrients in higher plants. International journal of molecular sciences 17 (8):1229

Wang S, Mitani-Ueno N, Takano J (2018) Boron Uptake Assay in Xenopus laevis Oocytes. Bio-protocol 8 (5):e2755

Wang Y, Tajkhorshid E (2007) Molecular mechanisms of conduction and selectivity in aquaporin water channels. The Journal of nutrition 137 (6):1509S-1515S

Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L (2018) SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic acids research 46 (W1):W296-W303

Wei W, Alexandersson E, Golldack D, Miller AJ, Kjellbom PO, Fricke W (2007) HvPIP1; 6, a barley (Hordeum vulgare L.) plasma membrane water channel particularly expressed in growing compared with non-growing leaf tissues. Plant and Cell Physiology 48 (8):1132-1147

Wu B, Beitz E (2007) Aquaporins with selectivity for unconventional permeants. Cellular and molecular life sciences 64 (18):2413-2421

Xu F, Wang K, Yuan W, Xu W, Shuang L, Kronzucker HJ, Chen G, Miao R, Zhang M, Ding M (2018) Overexpression of rice aquaporin OsPIP1; 2 improves yield by enhancing

mesophyll CO2 conductance and phloem sucrose transport. Journal of experimental botany 70 (2):671-681

Xu X, Pan S, Cheng S, Zhang B, Mu D, Ni P, Zhang G, Yang S, Li R, Wang J (2011) Genome sequence and analysis of the tuber crop potato. Nature 475 (7355):189

Yoo S-D, Cho Y-H, Sheen J (2007) Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. Nature protocols 2 (7):1565

Yoo Y-J, Lee HK, Han W, Kim DH, Lee MH, Jeon J, Lee DW, Lee J, Lee Y, Lee J (2016) Interactions between transmembrane helices within monomers of the aquaporin AtPIP2; 1 play a crucial role in tetramer formation. Molecular plant 9 (7):1004-1017

Yuan D, Li W, Hua Y, King GJ, Xu F, Shi L (2017) Genome-wide identification and characterization of the aquaporin gene family and transcriptional responses to boron deficiency in Brassica napus. Frontiers in Plant Science 8:1336

Zelazny E, Borst JW, Muylaert M, Batoko H, Hemminga MA, Chaumont F (2007) FRET imaging in living maize cells reveals that plasma membrane aquaporins interact to regulate their subcellular localization. Proceedings of the National Academy of Sciences 104 (30):12359-12364

Zhao XQ, Mitani N, Yamaji N, Shen RF, Ma JF (2010) Involvement of silicon influx transporter OsNIP2; 1 in selenite uptake in rice. Plant Physiology 153 (4):1871-1877

Zhou S, Hu W, Deng X, Ma Z, Chen L, Huang C, Wang C, Wang J, He Y, Yang G (2012) Overexpression of the wheat aquaporin gene, TaAQP7, enhances drought tolerance in transgenic tobacco. PloS one 7 (12):e52439

Zwiazek JJ, Xu H, Tan X, Navarro-Ródenas A, Morte A (2017) Significance of oxygen transport through aquaporins. Scientific reports 7:40411