Investigating the Genetic Control of Boron Concentrations in *Brassica napus*

Helen Elizabeth Riordan

PhD

University of York

Biology

March 2019

Abstract

The global population is projected to increase dramatically in the coming decades. This, in conjunction with climate change and arable land degradation, means that food production needs to be increased significantly. One potential method to meet this demand is to augment our understanding in (and ultimately optimise) plant nutrition, with an aim to improve crop yields.

This research has focused on investigating the control of boron (B) concentrations in *Brassica napus.* Although B is an essential plant micronutrient, it is often deficient within the soil. Consequently, significant fertiliser inputs are required to maintain yields, which can have both negative economical and agronomical implications. One particularly pertinent example of this is the disorder 'Brown Heart', which affects *Brassica napus* var. *napobrassica* (the root vegetable swede). Brown Heart is specifically attributed to B-deficiency, however there is a lack of clear understanding around the basis of this disorder. This research designed a trial to explore if supplementing plants with B would reduce BH incidents. No significant reductions were observed, nor did B concentrations differ significantly between BH affected and clean roots; demonstrating that additional factors may be underlying this disorder.

To further understand specific genetic factors controlling B concentrations within *B. napus* leaves and seeds*,* an Associative Transcriptomics (AT) approach was adopted. AT looked to assess how natural variation in both gene sequence and expression affected these two traits. Ten leaf and six seed AT candidates were further analysed using *Arabidopsis thaliana* T-DNA insertions lines, with disruption in B concentrations relative to wild type controls identified in seven T-DNA lines. Notably decreased seed B concentrations in *Atnip2;1* (a member of a gene family which includes known B transporters) was identified. Markers identified from AT could be useful for marker-assisted selection strategies to improve fertiliser or B-use efficiency and ultimately positively impact crop yields.

Table of Contents

List of Tables

List of Figures

Acknowledgements

Thank you to the Biotechnology and Biological Sciences Research Council (BBSRC) and Elsoms Seeds for funding this project (award reference: 1576285).

I would like to thank my supervisor Prof. Ian Bancroft. Thank you for giving me the opportunity to work in such a supportive environment and for your help and guidance always; you really did deserve that ‘supervisor of the year’ award! Secondly, thanks you to my secondary supervisor Dr Andrea Harper, your assistance in everything was invaluable.

Thank you to my industrial supervisor Dr Sue Kennedy for your help with all things Swede, your questions were always insightful and helped shape my research. Additional thanks Dr Richard Tudor and all other colleagues at Elsoms, especially to those who helped look after my plants and kept the aphids at bay.

Thank you to my thesis advisory panel, Prof. Frans Maathuis and Prof. Seth Davis; your advice in the thesis advisory meeting was always incredible helpful and I gained so much knowledge because of you both.

Thank you to all members, past and present, of the Bancroft lab. You have all been amazing people to work with. Special thanks to Dr Lenka Havlickova and Dr Zhesi He for help with all things AT. Thanks to Lihong Cheng and Alison Fellgett for being the best lab techs ever. Thanks to Dr Natalia Stawniak for always giving out great advice. Thanks to Varanya Kittipol, Dr Harjeevan Kaur and Dr Aoife Sweeney for being the best fellow PhD students and for always being great fun, both inside and outside of the lab. Special Thanks to Aoife for being my plant nutrition buddy; all those hours of digestions and ICP-MS lamenting were much more fun because of you.

Thanks to Prof. Martin Broadley and his research group at the University of Nottingham for help with the ionomic work. A special thanks to Lolita Wilson for help in the setup of our own digestion protocol and assistance in running all of the many ICP-MS samples.

Lastly thanks to all of my family and friends. Thank you to Brian for your love and support over the past years. Special thanks for helping me with the swede sampling and for reading this whole thing. Thanks to Paul for being a pretty great brother; you've always been so supportive of everything I've done and I'm forever grateful to you for your help and advice. And finally, thanks to Mum and Dad for everything; none of this would have been possible without the both of you.

Declaration

‘I, Helen Elizabeth Riordan, declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.’

Chapter 1: Introduction and Review of Literature

1.1 Introduction to thesis content

This thesis comprises research conducted to further understand the genetic control of boron (B) concentration within *Brassica napus* (*B. napus*). Initially, the aims and scope of this research are detailed, with the subsequent literature review providing current knowledge from the discourse. Next, the materials and methods utilised throughout this research are outlined. Following this, the first and second results chapters detail research relating to the identification of loci controlling leaf and seed B concentrations respectively. These chapters begin with an introduction to leaf and seed B traits. The results show leaf and seed B concentrations within the *B. napus* diversity panel and correlations with other essential elements; Associative Transcriptomic (AT) outputs; trait predictions; identification of markers selected for further characterisation in *Arabidopsis thaliana* (*A. thaliana*) and the subsequent ionomic profile of the *A. thaliana* T-DNA mutants. Within the discussion, further information on candidates found to vary significantly in B concentrations has been provided, along with how these candidates should be further analysed and a critical review of utilised techniques. The discussion also provides information on additional candidates not selected for further characterisation in *A. thaliana*. Finally, an overall summary of the research presented in each chapter has been provided. The third results chapter details research relating to the swede B-deficiency disorder Brown Heart (BH). This starts with a review of literature relating to this disorder and subsequently details chapter specific methods. Within the results, the presence of BH within swede roots and the effect of B supplementation on the presence and severity of the disorder are given, as are B concentrations within root and leaf material and BH correlations with additional elements. The discussion summarises results obtained and notable trial limitations. Finally, the general discussion summarises how the research aims of the thesis were met and the key hypotheses tested; notable research outcomes; critical analysis of the methodologies utilised and potential directions for further investigation.

1.2 Aims and Scope

This thesis' main research aim was to further understand the control of B concentrations within *B. napus,* an oilseed, root (including swede) and leafy vegetable crop of significant economic importance. Firstly, a quantitative genomics technique, 'Associative Transcriptomics' (AT), was extensively employed to identify loci controlling leaf and seed B concentrations in a large diversity panel of *B. napus* accessions, grown under nutrient sufficiency.

Through conducting AT analyses, one key hypothesis was tested:

> *Variation in boron concentrations of B. napus leaves and seeds, grown under nutrient sufficiency, is a result of underlying natural genetic variation*

Through AT analysis, variation in both gene sequence and gene expression across the *B. napus* diversity panel identified regions associated with leaf and seed B concentrations. These associated regions were then further analysed to identify candidate genes potentially causative of the observed associations (and therefore playing a role within the control of B concentrations within different tissues). This technique elucidated two types are candidates; 1) genes previously identified as affecting nutrient concentrations and 2) novel candidates. Orthologues of novel candidate genes were then further characterised in *Arabidopsis thaliana* (*A. thaliana*); the model plant species and member of the *Brassicaceae* family. Extensive genome collinearity exists between *A. thaliana* and *B. napus,* enabling the identification of orthologous genes between species. As such, this research exploited *A. thaliana* T-DNA insertion mutants for quick validation of candidate genes. For this, *A. thaliana* orthologues of *B. napus* AT candidates were tested for disruption in B concentrations within T-DNA lines (relative to wild type controls). Any perturbation within the mutant validated the candidate gene.

Ultimately increased understanding in the genetic basis for variation in B concentrations would enable marker-assisted breeding strategies to maximise B-use efficiency in *B. napus*. This is particularly relevant given that B is required in high concentrations in crops (relative to other micronutrients), particularly during reproductive growth. Furthermore, *B. napus* is highly susceptible to B-deficiencies, with B-deficient soils prevalent globally. Such research could reduce

the reliance upon fertiliser input, enhance cultivation within B-deficient soils and help to stabilise crop yields.

The secondary aim of this research was to specifically study B concentrations in a selection of swede (*B. napus* var. *napobrassica*) accessions. A disorder known as 'Brown Heart' (BH) is currently attributed to B deficiency in swedes. Resultantly, growers are required to input high levels of fertilisers to attempt to mitigate BH symptoms. However, this is a costly and often ineffective process as instances of BH are still reported frequently at crop harvest. Certain swede accessions appear to have a heightened susceptibility to BH, indicating a genetic link behind the disorder. However incidents are still reported in accessions thought to be more BH resistant. Consequently a trial was designed in order to assess two questions:

1. *Does swede accession affect the presence and severity of BH symptoms within the root?*
2. *Does supplementation with B affect the presence and severity of BH symptoms within the root?*

Through analysing the effect of B supplementation on the presence and severity of BH across sixteen swede accessions, it was hoped further insight into the basis of BH would be attained. Ultimately, such research may allow for future recommendations for BH control.

1.3 Literature Review

1.3.1 Boron as an essential plant micronutrient

1.3.1.1 Plant nutrition basics

The concept of mineral fertilisation has been known for millennia, with the use of livestock manure for the enhancement of crop yields demonstrated up to 8,000 years ago (Bogaard et al., 2013). However, the essentiality of certain nutrients (i.e. being fundamentally necessary for life cycle completion) took longer to elucidate, with nickel only being explicated in the 1980s (Marschner, 1995).

Plant mineral elements can be split into two broad categories; macronutrients and micronutrients, with fourteen having been identified as essential requirements for plant growth (Marschner, 1995). The ability to distinguish between macro and micronutrients is based on the plant's requirements for the nutrient, with macronutrients required in relatively large quantities (> 1000 mg kg^{-1} DM), and micronutrients in relatively small (~ 5-100 mg kg^{-1} DM) quantities. Six elements have been defined as macronutrients; nitrogen (N), phosphorus (P), sulphur (S), potassium (K), calcium (Ca) and magnesium (Mg), whereas there have been eight micronutrients currently identified; boron (B), iron (Fe), zinc (Zn), copper (Cu), chlorine (Cl), nickel (Ni), manganese (Mn) and molybdenum (Mo) (Marschner, 1995). The availability of these nutrients is essential for plant health and for maximising crop yields. However, it is possible for toxic concentrations of some elements, both essential nutrients and non-essential elements, to be accumulated. This in itself can be detrimental to plant health (White and Brown, 2010), with B toxicity, for example, having the potential to negatively affect crop yields (Nable, Bañuelos and Paull, 1997). Further details on B toxicity will be discussed within Section 1.3.1.8.

1.3.1.2 B chemistry

Elemental B is a metalloid and, like other members of this group, such as the beneficial silicon (Si) and the toxic Arsenic (As), exhibits intermediate properties between metals and non-metals. In conditions of near-neutral physiological pH (found within most biological fluids), B primarily is found as uncharged boric acid (H_3BO_3); with up to 96% being present in this form (Bolaños et al., 2004). Boric acid is classified as a weak Lewis acid, with a pKa of 9.2. However, when B forms covalent bonds with biological ligands, and pKa is reduced to ≤ 6, the negatively charged borate anion [$B(OH)_4^-$] is formed. These monomeric forms (boric acid and borate), are almost solely found at typical physiological B concentrations (6.0 x 10^{-7} mol/L - 9.0 x 10^{-3} mol/L) (Hunt, 2003), with polymeric species unlikely to be found except if B is at toxic levels (Marschner, 1995). The distribution of B is widespread, with major sites of B ore found in California and Turkey. Nevertheless, B availability varies significantly, with average concentrations found within oceans of 4.5 mg/L (Lemarchand et al., 2000) and 10 mg/kg in the

earth's crust. However, this latter figure can vary significantly dependent upon rock type (Shorrocks, 1997).

1.3.1.3 B as an essential element

B was first identified as an essential nutrient in the 1920s (Warington, 1923), with studies conducted in Broad Bean (*Vicia faba*) concluding that a continual supply of boric acid was required for healthy plant growth. Subsequent studies designated B as a micronutrient, although it is required in high concentrations relative to other micronutrients (Marschner, 1995). However, the role B plays within the plant is still not fully understood, with a wide array of symptoms associated with B deficiency (see Section 1.3.1.7). Consequently this had led researchers to hypothesise a diverse array of functions (Bolaños et al., 2004). There has, for example, been no evidence to show that B is involved in enzymatic activity, nor that it forms part of any enzyme constituents. Whilst potential functions have been discussed, such as roles in respiration, sugar transport and the metabolism of varied compounds, the molecular basis of their roles are yet to be defined (Marschner, 1995). To date, the only function fully elucidated is the role B plays within the formation of the cell wall. A discussion of this is reviewed in Section 1.3.1.4, however other postulated roles will be discussed in Section 1.3.1.5.

The World Health Organisation (WHO), although deeming B beneficial, does not currently designate B as an essential nutrient for human health (Uluisik, Karakaya and Koc, 2018). Irrespective of this, nutritional studies in other animal models have shown B to be essential for various developmental stages; from embryonic development (Fort et al., 1998; Lanoue et al., 1998) to bone formation (Chapin et al., 1997; Armstrong et al., 2002). At the microorganism scale, studies in algae (McLachlan, 1977), diatoms (Lewis and Chen, 1976) and cyanobacteria (Bonilla, Garciagonzalez and Mateo, 1990) have demonstrated B to be essential.

1.3.1.4 B in cell wall formation

Although known to be essential for plant growth (Warington, 1923), elucidating the main function of B was a significant challenge. A major breakthrough however

was the discovery of B integration into the cell wall, with a primary study demonstrating the correlation between the pectic content of the cell wall and a plants' requirement for B (Matoh, Kawaguchi and Kobayashi, 1996). Further details of this are explored below.

A principal component of plant cell walls are cellulose microfibrils; crystalline insoluble structures composed of hydrogen-bonded glucose molecules (Somerville, 2006). Additionally, there are matrix polysaccharides, grouped into two categories, which are also vitally important within the structure and function of the cell wall. These include 1) *pectic polysaccharides,* which consist of homogalacturonan, rhamnogalacturonan I and rhamnogalacturonan II (Harholt, Suttangkakul and Vibe Scheller, 2010), and 2) *hemicellulosic polysaccharides* including xyloglucans, xylans, and mixed-linkage glucans (Scheller and Ulvskov, 2010). The important component listed here, in as far as B function is concerned, is the pectic polysaccharide rhamnogalacturonan II (RG-II).

The RG-II backbone is formed from a minimum of seven 1,4-linked α-D-galactopyranosyluronic acid (α-D-GalpA) residues (Puvanesarajah, Darvill and Albersheim, 1991) and holds four (A-D) oligosaccharide side chains. However, the specific organisation of these sidechains is not known with certainty, with the glycosyl sequences often differing between species (Bar-Peled, Urbanowicz and O'Neill, 2012) and within populations. For example, Pabst et al., (2013) characterised the RG-II sidechains in *A. thaliana*; showing that up to 45% of wild type plants had a substituted L–galactose rather than L–fucose within the A sidechain. In total, they describe 18 RG-II structural variants (including methylation and chain length variations) in *A. thaliana,* of which the majority were found within the A sidechain. It is the A sidechain in which the role of B is crucial, whereby the crosslinking of two RG-II molecules occurs between apiosyl residues on these chains (Ishii et al., 1999) by a borate diester (Kobayashi, Matoh and Azuma, 1996; O'Neill et al., 1996) (Figure 1.1).

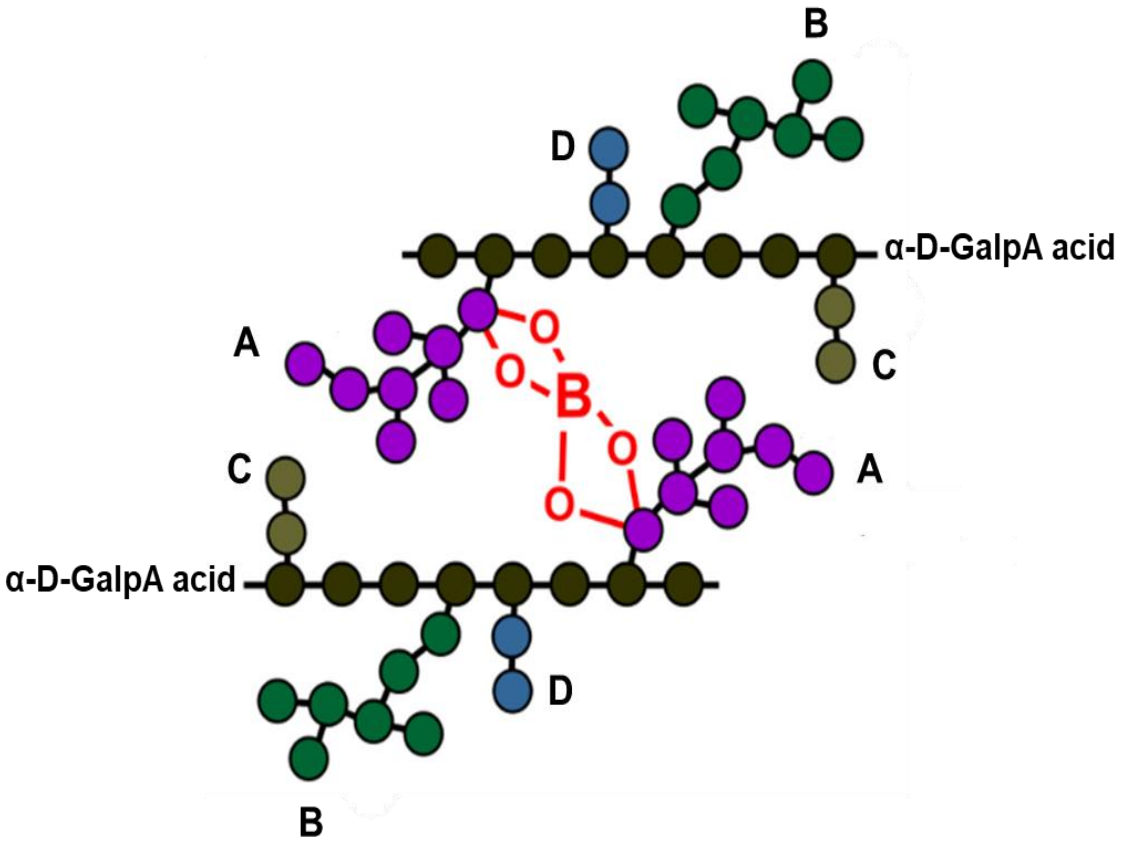


Figure 1.1 Representation of rhamnogalacturonan II cross-linkage in plant cell walls

Rhamnogalacturonan II (RG-II) has a backbone formed from 1,4-linked α-D-GalpA acid residues. Four (A-D) oligosaccharide side chains are attached to this backbone, shown here with different colours differentiating between the four side chains. The specific organisation and formation of these side chains can differ between species. A borate diester bond (as seen in red) is formed between the apiosyl residues found on the A side chain on two separate RG-II molecules, with this dimerization an essential constituent of the plant cell wall.

Adapted from Bar-Peled, Urbanowicz and O'Neill., (2012).

RG-II-borate dimerization is essential for the 3D pectic network required for the formation of the primary cell wall, thus enabling normal growth and development (O'Neill et al., 2004). Cell wall porosity has also been linked to this dimerization; with *Chenopodium album* (Pigweed) cells grown in borate-lacking media exhibiting atypical pore size, with phenotype restored once re-supplied with B. This suggests that molecule transport across the cell wall may be affected by RG-II-borate crosslinking (Fleischer, O'Neill and Ehwald, 1999). Furthermore, tensile strength of the cell wall is affected by cross linkage, with mutational studies in *A. thaliana* showing reduced organ strength when dimerization is reduced (Ryden

et al., 2003). Several cell wall mutants have been identified which exhibit altered RG-II dimerization. A key example of this is the fucose deficient mutant *mur1* (Reiter, Chapple and Somerville, 1993; Reuhs et al., 2004), which exhibits A sidechain truncation (Pabst et al., 2013). The dwarfed phenotype of *mur1* is ultimately restored with boric acid treatment (O'Neill et al., 2001). When grown in sufficient B supply, at least 90% of RG-II is cross-linked with borate (O'Neill et al., 1996), with linkage significantly reduced in borate-lacking medium (Fleischer, O'Neill and Ehwald, 1999). In terms of total cellular B present, up to 95% is found within the cell wall fraction (dry weight) (Blevins and Lukaszewski, 1998).

1.3.1.5 Other postulated roles of B

The wide array of symptoms reportedly affected by B deficiency (Brown et al., 2002), in addition to the discovery that B exhibits phloem mobility (see Section 1.3.1.6) (Marschner, 1995; Brown et al., 1999), highlights the potential for additional roles *in planta*, other than cell wall formation. B availability has been found to significantly affect carbohydrate metabolism in a variety of model species including *Nicotiana tabacum* (*N. tabacum*; tobacco) (Camacho-Cristóbal and González-Fontes, 1999), *A. thaliana* (Chen et al., 2014) and others (Han et al., 2009; Ruuhola et al., 2011; Xiuwen et al., 2018). The permeability of the plasma membrane (PM) to ions and various other solutes also appears altered by B availability (Wang et al., 1999), however no definitive mechanism by which B affects the structure of the PM has been identified. Due to this, one hypothesis is that membrane hydroxylated ligands (such as glycoproteins or glycolipids) are possible candidates forming the basis of B-linked complexes (Goldbach and Wimmer, 2007); however the relevant mechanisms are yet to be identified. A potential role in auxin metabolism has also been reviewed (Marschner, 1995) with B deficient plants exhibiting a reduction in both auxin *and* cytokinin concentrations (Wang et al., 2006). In the cyanobacterium *Anabaena*, B was found essential for N fixation (Blevins and Lukaszewski, 1998), whereas in *N. tabacum* roots, the expression of a PM H^+-ATPase was repressed under B deficiency, resulting in reduced nitrate intake (Camacho-Cristóbal and González-Fontes, 2007). Various other postulated roles of B have been reviewed, such as nucleic acid metabolism, pollen tube growth and germination (Marschner, 1995). Goldbach and Wimmer

(2007) have also suggested that B may act as a cellular signal (similar to Ca) which is capable of interacting with various transcription factors. They argue this may explain how B is able to affect such a wide array of physiological processes, although relevant mechanisms for this have not been determined.

As all these postulated processes have primarily been identified under B-deficiency stresses, it is difficult to discern whether these form primary functions or indirect effects (of deficiency). As such, the only definitive role of B is its well elucidated role in cell wall formation. However, with improved technologies, it may be possible to identify further, albeit undoubtedly minor, functions.

1.3.1.6 B Transport

For many years B uptake at the root surface and transportation throughout the plant was thought only to occur via passive diffusion (Hu and Brown, 1997). This was primarily due to the fact that boric acid (the principal form of B in the soil), as an uncharged molecule, is highly permeable across lipid bilayers (Brown and Shelp, 1997). However, the demonstration that B can be transported against concentration gradients (Dannel, Pfeffer and Romheld, 2000) suggested that active transport mechanisms exist, at least in conditions of B deficiency when passive transport cannot sufficiently supply the plant.

B transport from the root surface towards the xylem requires two transmembrane events; 1) import into the epidermis cells through the cortex to the endodermis cells, and 2) export from the endodermis cells into the stellar apoplasm (xylem loading). Under B deficiency there are two key facilitators of these processes in *A. thaliana*; the boric acid channel *At*NIP5;1 (Takano et al., 2006) and the transporter *At*BOR1 (Takano et al., 2002) (Figure 1.2).

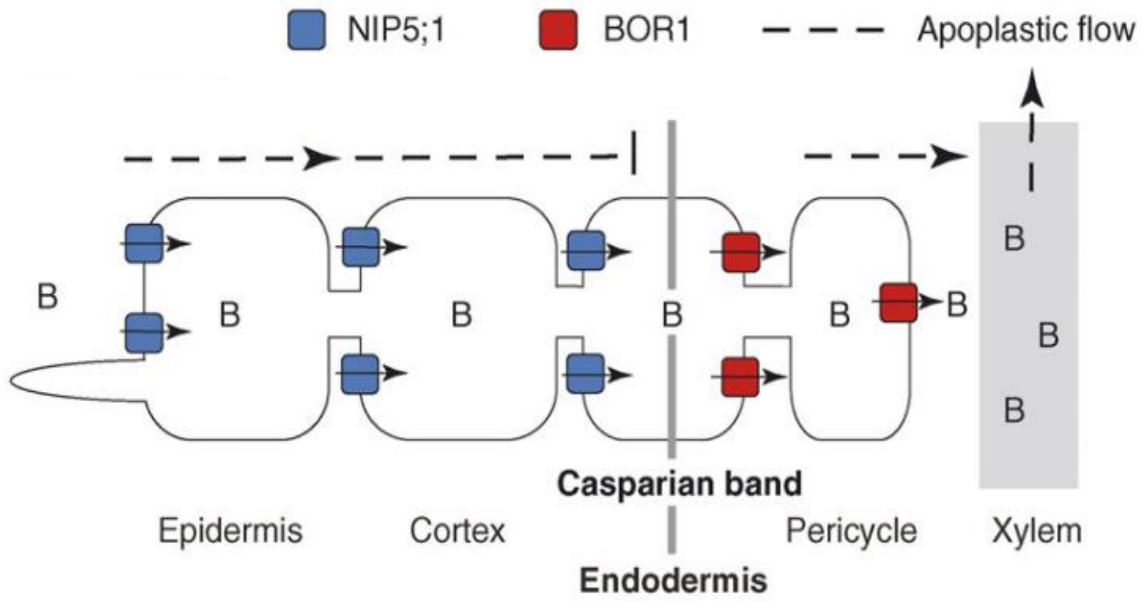


Figure 1.2 Boron transport in A. thaliana under B-deficient conditions

In low B conditions, the boric acid channel AtNIP5;1 imports B into the epidermis cells through to the endodermis cell layer. The AtBOR1 transporter subsequently facilitates xylem loading. Adapted from Takano, Miwa and Fujikara., (2008).

Aquaporins

Aquaporins (AQPs) are members of the major intrinsic protein (MIPs) family (Maurel and Chrispeels, 2001), of which several key members have been characterised in higher plants. These include; NOD26-like intrinsic proteins (NIPs), plasma membrane intrinsic proteins (PIPs), tonoplast intrinsic proteins (TIPs) and small basic intrinsic proteins (SIPs) (Li, Santoni and Maurel, 2014). Originally, AQPs were solely thought to be involved in water-transport, however their roles are actually wide-spread, facilitating gas-transfer, solute distribution and nutrient uptake (Tyerman, Niemietz and Bramley, 2002).

Microarray and RT-PCR analyses initially identified *AtNIP5;1* (one member of the NOD26-like intrinsic proteins) upregulation in both the root elongation and root hair zones under B deficient conditions, with a 10-fold increase in transcript levels (Takano et al., 2006). As a result, *At*NIP5;1 was confirmed as a key facilitator for B uptake at the root surface and transport across the PM under low B conditions. Furthermore, *AtNIP5;1*-facilitated transport of boric acid across the PM has been shown to be a limiting step for the sufficient supply of B required for RG-II

dimerization (Takano et al., 2006). Under B-sufficiency, *AtNIP5;1* mRNA degradation is essential for acclimation to high-B conditions (Tanaka et al., 2011). *At*NIP6;1, shares the highest level of homology to *At*NIP5;1, with 66.4% conserved amino acid sequence (Wallace, 2004), and is involved in the preferential distribution of B towards developing tissues. Furthermore, *AtNIP6;1* is likely involved in the xylem to phloem transfer of B at the nodal regions (Tanaka et al., 2008). Additionally, *AtNIP7;1* has been identified as a boric acid channel, primarily expressed in the anthers (Li et al., 2011; Routray et al., 2018). However, NIPs do not always exhibit nutrient-specificity; the ability to transport Arsenite (AsIII) has been demonstrated for each of the three aforementioned *At*NIPs (Bienert et al., 2008; Isayenkov and Maathuis, 2008). *At*NIP5;1, *At*NIP6;1 and *At*NIP7;1 all belong to the NIP-II subgroup, characterised by a divergent aromatic arginine (ar/R) tetrad, with an amino acid substitution from alanine to tryptophan distinguishing these from the NIP-I subgroup (Wallace, 2004). Furthermore, *NIP7;1* proteins are only found in dicots and show a greater divergence from both *NIP5;1* and *NIP6;1*, indicating their more recent evolution (Routray et al., 2018).

In *B. napus,* 121 full length AQPs have been identified, including 32 NIPs (Yuan et al., 2017), with up-regulation of some of these genes under B-deficiency. Furthermore, certain PIPs exhibit upregulation under both B deficiency and toxicity (Yuan et al., 2017). A summary of these AQPs is provided in Table 1.1 (*BnaNIP7* isoforms have not been included as upregulation under B deficiency was not reported in studies from which data has been summarised). Interestingly, recent research into the function of *Bna*NIPs by Diehn et al., (2019) has demonstrated that *BnaNIP2*, *BnaNIP3* and *BnaNIP4* isoforms exhibit functional B transport activity when expressed in *Xenopus* oocytes. This provides the first experimental evidence to demonstrate the B transport capabilities of these NIPs, to the author's knowledge, in any plant species. However, further research is required to assess their functional capacity for B transport *in planta*.

Both the divergent roles of AQPs, and their non-specific functions, compound the complexity of the control of B accumulation in different tissues and under differing nutritional statuses. This is particularly relevant in species exhibiting polyploidy, where homoeologues (corresponding genes found in each genome) do not necessarily exhibit equal gene expression or function (Adams et al., 2003).

Differential expression of *BnaNIP*s in different tissues has also been observed. For example *BnaC05.NIP3;1a,* is upregulated during B-deficiency in both roots and closed flowers, whereas *BnaC05.NIP3;1b* (a gene in tandem duplication with *BnaC05.NIP3;1a)* is down-regulated in the rachis (stem bearing flower stalks) (Diehn et al., 2019).

Table 1.1 Aquaporins upregulated under boron deficiency

*The gene name and gene model (as identified from the pseudomolecule database) for the AQP gene induced under B-deficiency is given, as is the specific tissue where upregulation occurs; juvenile leaf (JV), old leaf (OV), 10-day seedling leaves (SL), closed flowers (CF). *Expression under B toxicity also reported. Data summarised from Hua et al., (2017); Yuan et al., (2017) and Diehn et al., (2019).*

Gene name	Gene model	Tissue where upregulation occurs under B deficiency	Reference
BnaC05.NIP3;1a	Bo5g091760.1	Root; CF	Diehn et al., (2019)
BnaC04.NIP4;1b	Bo4g143920.1	Root	
BnaC06.NIP4;2a	Bo4g143930.1	Root	
BnaA02.NIP5;1a	Cab044883.1	Root; JV; OL	Yuan et al., (2017)
BnaA03.NIP5;1b	Cab003613.1	Root; JV; OL	
BnaC02.NIP5;1a	Bo2g110350.1	Root; JV; OL	
BnaC03.NIP5;1b	Bo3g044690.1	Root; JV; OL	
BnaA02.NIP6;1b	Cab014832.1	SL	Hua et al., (2017)
BnaA09.PIP2;5	Cab000644.1	Roots	Yuan et al., (2017)
BnaC08.PIP2;5b	Bo8g086930.1	Roots	
*BnaA04.PIP1;2b**	Cab045734.2	SL	Hua et al., (2017)
*BnaA05.PIP1;2c**	Cab025376.1	SL	
*BnaC04.PIP1;2a**	Bo4g025110.1	SL	

Gene name	Gene model	Tissue where upregulation occurs under B deficiency	Reference
*BnaC04.PIP 1;2b**	Bo4g195940.1	SL	Hua et al., (2017)
*BnaA05.PIP2;2c**	Cab025641.1	SL	
*BnaC04.PIP2;2a**	Bo4g030810.1	SL	

BOR Transporters

The *AtBOR1* transporter was first shown to encode a membrane protein with significant homology to mammalian bicarbonate transporters (Takano et al., 2002), such as *NaBC1* (Park et al., 2004). *AtBOR1* is essential for xylem loading, a key step for B accumulation in shoots under deficient conditions. *Atbor1-1* mutants are defective in this process (Takano et al., 2002), whilst sufficient B supply reverses the mutant phenotype (Takano et al., 2005). Furthermore, over-expression of *AtBOR1* improves seed-set under conditions of B-deficiency, in which WT plants would fail to set seed (Miwa, Takano and Fujiwara, 2006). Under B-sufficiency, the regulation of *AtBOR1* is controlled by the vacuolar degradation of the protein, following the ubiquitination of the C-terminal Lysine-590 (Kasai et al., 2011). This regulation is essential for the control of B under conditions of both B deficiency and sufficiency. Various *AtBOR1* paralogues have also been identified; *AtBOR2* shares the highest sequence identity and promotes RG-II crosslinking and root elongation under low B supply (Miwa et al., 2013), thus improving tolerance to B deficiency (Takada et al., 2014).

In *B. napus*, the molecular mechanisms for B transport are complex and six *AtBOR1* orthologues have been identified (Table 1.2). These genes exhibit 74.0 - 96.8% amino acid sequence homology with *AtBOR1* (Sun et al., 2012), and all six *BnaBOR1* copies exhibit functional B efflux capacity when expressed in yeast (Diehn et al., 2019). Whilst each of the homoelogous gene pairs (i.e. the A and C genome *BOR1* copies) have the same gene and protein lengths, variation is observed between the other gene pairs. For example *BnaA03.BOR1;3a* and *BnaC03.BOR1;3c* are 2112 bp (703 amino acids) whilst *BnaA04.BOR1;1a* and

BnaC04.BOR1;1c are 2016 bp (701 amino acids) (Sun et al., 2012). *BnaC04.BOR1;1c*, is reportedly expressed in both roots and shoots, and has been shown to control the preferential distribution of B towards growing tissues under B deficiency (Zhang et al., 2017; Chen et al., 2018) which typically have higher nutrient demands. Conversely, Diehn et al., (2019) found that only *BnaC04.BOR1;2c* and *BnaA04.BOR1;1a* were significantly upregulated under B deficiency (and not *BnaC04.BOR1;1c*). They concluded that this discrepancy with the results reported by Chen et al., (2018) was due to cultivar-specific expression patterns of *BnaBOR1*s, which may be a result of adaptation to different environmental demands (i.e. B availability in different regions or soil types).

Table 1.2 Overview of six BnaBOR1s and AtBOR1

The gene name and gene model (as identified from the pseudomolecule database) for the six BnaBOR1s are given, as is the A. thaliana orthologue, AtBOR1. The coding DNA sequence (CDS) base pair length, protein length and number of introns and exons are also provided, with data summarised from Sun et al., (2012). Additionally if upregulation under B deficiency has been reported this is stated, as is the specific tissue where upregulation occurs and the reference from which this was identified; leaves at the peduncle (LP). Data for expression profiles is summarised from Chen et al., (2018); Diehn et al., (2019) and Takano et al., (2002).

Gene Name	Gene Model	CDS (base pairs)	Protein Length (amino acids)	Intron Number	Exon Number	Significant Upregulation Under B Deficiency Documented?	Tissue	Reference
BnaA03.BOR1;3a	Cab003325.2	2112	703	11	12	N	-	
BnaA04.BOR1;1a	Cab047047.1	2106	701	9	10	Y	Roots; LP	Diehn et al., (2019)
BnaA05.BOR1;2a	Cab036553.1	2115	704	9	10	N	-	
BnaC03.BOR1;3c	Bo3g039540.1	2112	703	11	12	N	-	
BnaC04.BOR1;2c	BnaC04g00370D	2115	704	9	10	Y	Roots; LP	Diehn et al., (2019)
BnaC04.BOR1;1c	Bo4g198480.1	2106	701	9	10	Y	Roots; Shoots	Chen et al., (2018)
AtBOR1	AT2G47160	2115	704	11	12	Y	Roots	Takano et al., (2002)

Xylem and Phloem Transport

After xylem loading, irrespective of whether this occurs via passive transfer or facilitated loading by boric acid channels and transporters, long distance transport from root to shoot primarily occurs via the xylem vessels. Transport is driven by gradients in both hydrostatic pressure and water potential (White, 2011) and is essential for water and solute distribution to the leaves. A significant relationship between transpiration rate and B leaf accumulation has been reported (Ben-Gal and Shani, 2002).

The transfer of solutes via the xylem is driven by transpiration, rather than nutrient demand (Stangoulis et al., 2001). However *phloem transport* allows for nutrient remobilisation under periods of deficiency or senescence (Maillard et al., 2015); an essential process for the optimal usage of nutrients acquired over an entire growing period. The phloem mobility of B has been determined as '*intermediate'* (Marschner, 1995) and is seemingly species dependent. B-polyol complexes allow for ready translocation (Shelp et al., 1995), with mobility particularly pronounced in species that translocate sorbitol as their primary assimilate, such as almond (*Prunus amygdalus*) and apple (*Malus domestica*) (Brown and Hu, 1996). Transgenic experiments in *N. tabacum* have shown increased phloem transport of B when sorbitol synthesis is increased (Brown et al., 1999). In *Brassicaceae,* B phloem mobility has been reported (Shelp et al., 1995). Furthermore, in *B. napus* (where sucrose is the primary assimilate) Stangoulis et al., (2010) have identified a mechanism for B phloem transport, through the observation of sucrose-borate complexes within the phloem. This work provided the first quantitative evidence for B mobility in *B. napus*, with this potentially an essential mechanism to limit deficiency symptoms in developing tissues. A study in white lupin (*Lupinus albus*) provides a particularly salient example of this, in which Huang, Bell and Dell., (2008) were able to distinguish between 'old' and 'new' B within different tissues using stable B isotopes. Plants were initially grown in sufficient concentrations (20 µM) of ^{11}B, whilst during early flowering (48 days post sowing) plants were transferred to either 0.2 µM or 20 µM ^{10}B. Following this, translocation of the previously acquired ^{11}B towards growing tissues (shoots and inflorescence) was identified through analysis of both xylem sap and phloem exudates. Resultantly, phloem B transport enabled the net increase of the ^{11}B within the floral organs, without a supply of this particular isotope in the nutrient

solution. Reproductive growth is more susceptible to B deficiency than vegetative (Dell and Huang, 1997), with the inflorescence acting as a B sink (as observed in *B. napus* by Pommerrenig et al., (2018)). Therefore, phloem mobility/ translocation is essential for B supply to be maintained at such critical growth stages.

1.3.1.7 B Deficiency

In higher plants, B deficiency affects a multitude of cellular and physiological processes, with growth inhibited in both vegetative and reproductive stages (Camacho-Cristóbal, Rexach and González-Fontes, 2008). The expression of a wide array of genes, not limited to the major role of B (cell wall formation) is reportedly affected by B deficiency (Peng et al., 2012), along with both a modulation of phytohormone responses (Zhou et al., 2016) and a reduction in nitrate assimilation (Camacho-Cristóbal and González-Fontes, 2007). During vegetative growth, a major response to B deficiency is the inhibition or cessation of root elongation, with affects observed within hours after the nutrient is withdrawn (Dell and Huang, 1997). A signalling pathway involving phytohormones (including ethylene and auxin) and reactive oxygen species (ROS), initiated through impaired cell wall integrity, has been found to affect this reduction in elongation (Camacho-Cristobal et al., 2015). Additionally, leaf expansion is limited in newly initiated leaves, with symptoms first appearing at the shoot apex. In highly deficient plants these leaves will eventually become necrotic unless B supply is increased (Dell and Huang, 1997). As a result, photosynthesis is reportedly indirectly affected by B deficiency via decreased photosynthetic capacities. This effect is attributed to both a reduction in leaf area (Dell and Huang, 1997) and chlorophyll content (Han et al., 2007). Comparatively, the effects of B deficiency observed during reproductive growth are often more pronounced than vegetative growth, with flowering and seed yield particularly sensitive. This is particularly problematic when looking at overall crop yields, where significant reductions are observed, often with no previous B deficiency symptoms present (Dell and Huang, 1997).

B deficiency symptoms have been reported in at least 132 agricultural crops in over eighty countries worldwide, including the United Kingdom, which exhibit low

soil B concentrations (< 0.25 mg/kg^{-1}) (Shorrocks, 1997). The availability of B decreases with increasing soil pH (Goldberg, 1997), under drought conditions (Marschner, 1995) and conversely in areas with high rainfall (Shorrocks, 1997). Furthermore, soil structure plays an important role in B availability, with sandy soils often exacerbating the risk of deficiency (Goldberg, 1997). Generally dicots, which have higher percentages of pectin within their cell walls and consequently require higher concentrations of B, resultantly show an increased sensitivity to deficiency compared to monocots (Hu, Brown and Labavitch, 1996). *B. napus,* a dicot, shows a particular susceptibility to B deficiency (Yang et al., 1993) with a hot-water soluble B (HWSB) concentration of approximately 0.5 mg/kg^{-1} reported as a minimum requirement for normal growth (Zhao et al., 2008). As previously discussed, sorbitol concentration can affect tolerances to B deficiency (Brown et al., 1999), as can treating plants with sorbitol (Will et al., 2012), with the promotion of B phloem mobility increasing translocation in deficiency conditions.

Furthermore, specific B deficiency disorders have been characterised. In oilseed rape, 'flowering without setting seed' was identified in the 1970s (Wang et al., 1976). Another disorder characterised in *B. napus* is Brown Heart in the root vegetable swede (*Brassica napus var. napobrassica*). Typified as the internal brown discolouration of the marketable root (Sanderson, Sanderson and Gupta, 2002), the presence of this disorder is not visible unless the root centre is inspected upon crop harvesting (Figure 1.3). Similar reported disorders include 'hollow heart' in Soybean (*Glycine max*) and peanut (*Arachis hypogaea*) (Rerkasem et al., 1993), and 'hollow stem' in broccoli (*Brassica oleracea var. italic*) (Hussain et al., 2012).

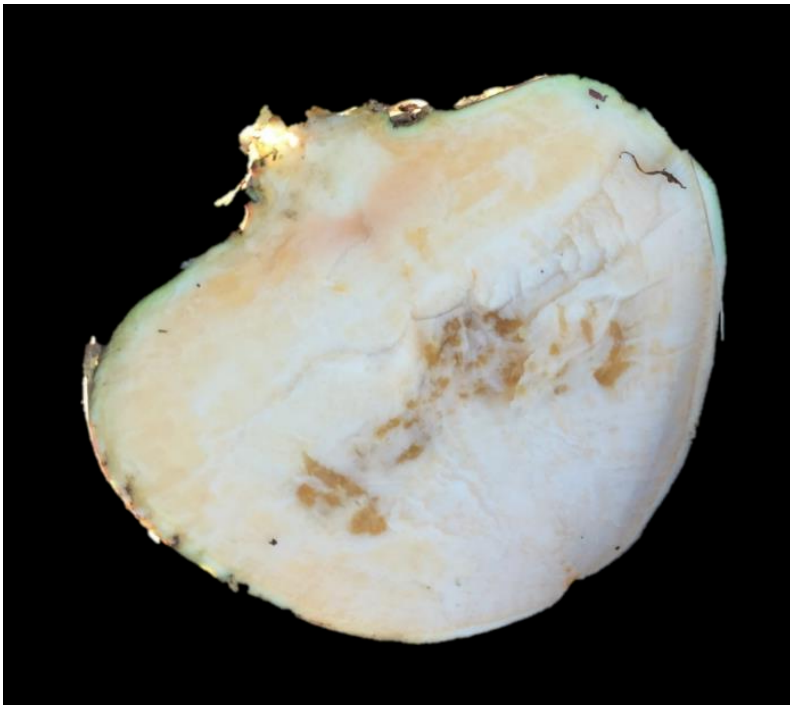


Figure 1.3 Severe Brown Heart symptoms within a swede root

As discussed in Section 1.3.1.6, numerous characterised genes have been found to aid or enhance B transport under deficiency conditions. These include *AtBOR1* (Takano et al., 2002), its paralogues (Miwa et al., 2013) and various AQP family members (Takano et al., 2006; Li et al., 2011) (See Table 1.1 and Table 1.2) Orthologues of these genes have been characterised in multiple species (Reid, 2014), including *B. napus,* and are able to promote B deficiency tolerance (Sun et al., 2012; Yuan et al., 2017). Furthermore *AtWRKY6*, a member of the WRKY transcription factor family (Phukan, Jeena and Shukla, 2016), is involved in the B deficiency response (Kasajima et al., 2010). *AtWRKY6* does not affect *AtNIP5;1* expression, indicating a role other than B uptake and translocation under deficiency. However, *AtWRKY6 mediated* regulation of variety root tip genes has been identified, implicating this gene is involved in tolerance to B deficiency through a modulation of root processes, such as elongation.

A multitude of quantitative trait loci (QTL) studies into B deficiency responses has highlighted various mechanisms for B-usage and efficiency. A focus on studies conducted in *B. napus* is reviewed in Section 1.3.1.9.

1.3.1.8 B Toxicity

Although not as prevalent as B deficiency in terms of areas affected, B toxicity still poses significant risks for agricultural yields in a multitude of regions worldwide. Toxic B concentrations are prevalent in semi-arid areas (such as California or Turkey) (Nable, Bañuelos and Paull, 1997), where low rainfall limits soil leaching (McDonald, Eglinton and Barr, 2010). Furthermore excess-fertilisation can result in an accumulation of toxic B concentrations (Marschner, 1995). The range between B toxicity and deficiency is narrow, more so than for any other nutrient (Goldberg, 1997). During rotation cycles this causes further issues, whereby it is possible for a crop to be grown in B supply that was adequate for the preceding crop, but then becomes in excess for subsequent rotations (Pommerrenig et al., 2018). Undoubtedly B fertilisation is imperative for maintaining yields and limiting deficiency, but it is vital that agricultural management is tightly controlled.

A reduction in both the quality and yield of crops grown in regions of B toxicity is typical (Nable, Bañuelos and Paull, 1997). The inhibition of root growth is an early symptom of toxicity (Reid et al., 2004), with more visual symptoms including leaf chlorosis and necrosis (Marschner, 1995). Both chlorosis and necrosis generally occur in older tissues or leaf tips, where B is more likely to be accumulated (Nable, Bañuelos and Paull, 1997). Although not definitive, three actions for B toxicity have been proposed: i) alterations to cell wall structure; ii) metabolic disruption of ATP, NADP and NADPH; and iii) disruption of cell division and development (Reid et al., 2004). However the specific molecular mechanism behind toxicity remains uncertain (Aquea et al., 2012) with more research required to elucidate the modes of action B toxicity poses.

It has been suggested that the main control mechanisms for B toxicity are applied within the roots via a modulation of B concentrations (Reid, 2007). As such, recent research has highlighted key regulatory genes controlling toxicity tolerance in a variety of species. In Barley (*Hordeum vulgare*) a *AtBOR1* orthologue, *HvBot1*, was found to limit B entry at the root surface (Sutton et al., 2007) with reduced expression of *HvNIP2;1* (a member of the NIP AQP family) limiting B uptake (Schnurbusch et al., 2010) and therefore enhancing B toxicity tolerance. Additional AQP family members are also thought to play a role in B

toxicity tolerance, such as various rice (*Oryza sativa*) (Kumar et al., 2014; Mosa et al., 2016) and *A. thaliana* (Macho-Rivero et al., 2018) plasma membrane intrinsic proteins (PIPs). Furthermore, in *A. thaliana* the overexpression of a *AtBOR1* paralogue, *AtBOR4*, was found to reduce B concentrations in plants grown under B toxicity (Miwa et al., 2007), with endogenous *AtBOR4* expression induced under high B being an essential mechanism for high-B tolerance (Miwa, Aibara and Fujiwara, 2014).

1.3.1.9 Quantitative trait loci studies for B traits

Mapping quantitative trait loci can be a useful tool to identify genomic regions controlling complex traits, based upon the association between molecular markers and phenotypic variation (Collard et al., 2005). Many of the QTL studies mapping the control of B in *B. napus* have focused on traits related to B-efficiency, perhaps unsurprisingly given the importance of B within agriculture and that oilseed rape exhibits heightened susceptibilities to B deficiencies (relative to other species) (Yang et al., 1993). In an early study by Xu, Wang and Meng, (2001), AFLP (amplified fragment length polymorphism) and RFLP (restriction fragment length polymorphism) markers were used to map B efficiency genes in an F2 population, formed from a cross between the B-efficient Qingyou10 and the B-inefficient Bakow. A number of agronomic traits were initially scored to assess the phenotypic differences between these parents, including maturity date, 1000-grain seed weight and 'B-efficiency coefficient' (seed yield ratio when plants were grown in soil with B = 0.25 mg/kg as compared to B = 1.00 mg/kg). The final seed yield within the F2 individuals was used to score for B-efficiency, whereby an estimation of the plants' ability for B uptake and utilisation when grown in low B conditions was assessed. This study identified one major QTL (*be9*) as controlling 64% of the phenotypic variation in B efficiency between individuals. From this, a gene for high B-efficiency, designated *BE1,* was determined to be closely located to the interval for the *be9* QTL. Subsequently, *A. thaliana* cDNA clones were used to map the gene; one RFLP marker, detected within a particular cDNA clone (pa28), was tightly linked to the *BE1* locus, and was located on *A. thaliana* chromosome one. However, the specific causative gene underlying *BE1* was not identified at this time. A further three minor loci

were also identified within this study, controlling 32.5% - 39.2% of the phenotypic variation (Xu, Wang and Meng, 2001). Interestingly, a study in *A. thaliana* by Zeng et al., (2007) found that a QTL for B-efficiency, *AtBE1-2*, overlapped with the *BE1* region on chromosome one in *A. thaliana* (as identified by Xu, Wang and Meng, (2001)). One candidate within this region included an AQP family member, *AtTIP3;1* (Zeng et al., 2007). Subsequent comparative mapping in *B. napus* by Zhao et al., (2012) found that two QTLs; *BEC-A2* and *SYLB-A2b* (B efficiency coefficient and seed yield in low B respectively) mapped to the same region as *AtBE1-2*, highlighting the importance of this region on chromosome A2 as a major locus for B-efficiency. Based on the speculated region of these QTLs in *B. napus*, putative candidate genes were screened from the equivalent region within chromosome two in *B. rapa*. Of these genes, the orthologues in *A. thaliana* included AT5G27350 and AT5G27360, members of a superfamily of monosaccharide transporters (Quirino, Reiter and Amasino, 2001); AT5G28470 a protein involved in pollen tube growth (Qin et al., 2009) and, perhaps most strikingly, *AtNIP5*;1 (AT4G10380) the B channel protein (Takano et al., 2006).

Further mapping studies in *B. napus* have also identified additional orthologues of *AtNIP5;1*. Zhang et al., (2014a) characterised a major QTL for 'B-efficiency coefficient' (defined as the ratio of the plant dry weight in low B to that in high B conditions), *qBEC-A3*, as controlling 30.79% of the phenotypic variation between individuals. Subsequent fine mapping by Hua et al., (2016a) proposed that *BnaA03.NIP5;1* (an orthologue of *AtNIP5;1*) was the underlying gene regulating B efficiency. Digital gene expression (DGE) profiling analyses demonstrated that the expression of *BnaA03.NIP5;1* was higher in the roots of the B-efficient cultivars (relative to B-inefficient). Furthermore, observed allelic polymorphisms in the 5′ untranslated region (UTR) were found to be conserved between two inefficient (Westar10 and Bakow) and two efficient (Qingyou10 and Zhongshuang 11) cultivars, with the authors suggesting that such variation may play a role in the regulation of *BnaA03.NIP5;1* expression, and ultimately variation in B-usage efficiency between individuals. Further orthologues of *AtNIP5;1* have been identified via QTL analyses, such as *BnaC02.NIP5;1* underlying *qBEC-C2a*; a QTL for B-efficiency (Hua et al., 2016b). A further QTL identified within this study was *qBEC-C2b*, of which the ATP-BINDING CASSETTE (ABC) G21 transporter, *BnaC02.ABCG21*, was a putative candidate gene. Although not known to be

involved in B transport, ABC transporters have been previously found to facilitate the transport of a variety of compounds (Hwang et al., 2016). RT-qPCR analysis demonstrated that *BnaC02.ABCG21* was upregulated under B-deficiency (both long and short term stress), with mRNA degradation when B was resupplied (Hua et al., 2016b).

Rather than focusing on specific B-efficiency traits (which itself can be quite abstract given that the specific definition changes based upon the research), Shi et al., (2012) conducted a study attempting to identify QTL related to root and shoot growth in *B. napus* seedlings grown under both B sufficiency and deficiency. They identified eighteen QTL (eleven at normal B and seven at low B), with orthologous genes in *A. thaliana* including AT5G25430.1 (a *AtBOR1* paralogue). A summary of a selection of QTLs and putative candidate genes identified in *B. napus* for B-traits, for both the examples reviewed here and additional studies, is provided in Table 1.3.

Table 1.3 Quantitative trait loci for traits related to B usage in B. napus

Initially the traits scored in the QTL study has been stated, as is the specific method by which the trait was measured. Next the name of the QTL, as given by the authors, is stated, as is the A. thaliana orthologue of the putative candidate gene(s) underlying the QTL. If available, the A. thaliana annotation name, as detailed on TAIR (Lamesch et al., 2012), is given. The putative link to the B traits under observation has also been provided, as detailed in the corresponding publication. References from which data within this table was summarised from is also provided for each QTL study.

Traits Scored	**Method for trait scoring**	**QTL**	***A. thaliana* orthologue of putative candidate gene**	***A. thaliana* annotation name**	**Gene function/ speculated link to specific B trait under investigation**	**Reference**
B-efficiency	g of seed yield	*be9*	Unknown: mapped to chromosome 1			Xu et al., 2001
		be3	Unknown			
		be8	Unknown			
		be14	Unknown			
B-efficiency	B utilisation efficiency (dry seedling material / B accumulation)	*be2*	Unknown			Zhao et al., 2008
B-efficiency coefficient	Ratio of the plant dry weight in low B to that in high B conditions	*qBEC-A3*	AT4G10380.1	*NIP5;1*	AQP boric acid channel	Zhang et al., 2014; Hua et al., 2016a
		qBEC- C2a	AT4G10380.1	*NIP5;1*	boric acid channel	Hua et al., 2016b
		qBEC-C2b	AT3G25620.1	*ABCG21*	transmembrane transport	

Traits Scored	**Method for trait scoring**	**QTL**	***A. thaliana* orthologue of putative candidate gene**	***A. thaliana* annotation name**	**Gene function/ speculated link to specific B trait under investigation**	**Reference**
Seed yield under low B *and* B-efficiency coefficient	Seed yield under low B; ratio of the yield under low B condition to that under normal B conditions	*SYLB-A2* and *BEC-A2* (QTL region co-localised)	AT5G27360 .1	*SFP2*	Monosaccharide transporter family	Zhao et al., 2012
			AT5G27350.1	*SFP1*	Monosaccharide transporter family	
			AT5G28470.1	No annotation	pollen tube growth	
			AT4G10380.1	*NIP5;1*	boric acid channel	
Seed weight	seed weight under normal B	*SWNB-A9*	AT1G31885.1	*NIP3;1*	AQP water channel transport family	
			AT1G35720.1	*ANNEXIN*	abiotic stress related genes induced by low B	
			AT4G18910.1	*NIP1;2*	AQP water channel transport family	
		SWNB-A10	AT5G59290.2	*ATUXS3*	response to low B stress	
			At5g17920.1	*ATCIMS*	response to low B stress	
Seed yield under low B *and* B efficiency coefficient	Seed yield low B; ratio of the yield under low B condition to that under normal B conditions	*SYLB-C3* and *BEC-C3* (QTL region co-localised)	AT2G34390.1	*NIP2;1*	water channel transport family	
			AT2G34630.1	*GPPS/GPS1*	plant height/seed number	
			AT2G34710.1	*PHB*	branch number	

Traits Scored	**Method for trait scoring**	**QTL**	***A. thaliana* orthologue of putative candidate gene**	***A. thaliana* annotation name**	**Gene function/ speculated link to specific B trait under investigation**	**Reference**
Seed yield under low B *and* B efficiency coefficient	Seed yield low B; ratio of the yield under low B condition to that under normal B conditions	*SYLB-C3* and *BEC-C3* (QTL region co-localised)	AT2G36880.1	*MAT3*	biosynthetic process gene induced by low B	Zhao et al., 2012
B efficiency coefficient	Ratio of the yield under low B condition to that under normal B conditions	*BEC-A7a*	AT2G18390.1	*TTN5/ARL2*	seed weight	
			AT2G19450.1	*TAG1*	seed weight	
			AT2G20120.1	*COV1*	branch number	
			AT2G20180.1	*PIL5/PIF1*	plant height	
			AT2G20190.1	*CLASP*	plant height	
Plant height	Plant height under normal B	*PHNB-C4*	AT2G38050.1	*DET2/DWF6*	plant height	
			AT2G36880.1	*MAT3*	biosynthetic process genes induced by low B	
Plant height	Plant height under normal B	*PHNB-C4*	AT2G36830.1	*GAMMA-TIP1*	plant height	
			AT2G36800.1	*DOGT1*	plant height	
			AT1G62300.1	*WRKY6*	cellular response to B starvation	

Traits Scored	**Method for trait scoring**	**QTL**	***A. thaliana* orthologue of putative candidate gene**	***A. thaliana* annotation name**	**Gene function/ speculated link to specific B trait under investigation**	**Reference**
Plant height	Plant height under normal B	*PHNB-C4*	AT1G62360.1	*STM*	branch number/flowering time/plant height	Zhao et al., 2012
			AT1G63030.1	*DDF2*	plant height	
Shoot B	Shoot B under low B	*L_SBA2-A1*	AT5G25430.1	*BOR6*	BOR1 homologue; Anion exchange activity	Shi et al., 2012
Root dry weight	Root dry weight in normal B	*N_RDW-A2*	AT5G25430.1	*BOR6*	BOR1 homologue; Anion exchange activity	
			AT5G25760.2	*PEX4*	Ubiquitin-conjugating enzyme	
Root dry weight	Root dry weight in normal B	*N_RDW-A2*	AT1G78340.1	*GSTU22*	Glutathione transferase activity	
			AT1G78360.1	*GSTU21*	Glutathione transferase activity	
B Transfer coefficient	ratio of root dry weight to shoot dry weight	*L_R/S2-C6*	AT1G73190.1	*TIP3;1*	AQP water channel transport family	
			AT1G74810.1	*BOR5*	Anion exchange activity	
Shoot dry weight	Shoot dry weight in normal B	*N_SDW-C9*	AT5G35630.2	*GS2*	Glutamate–ammonia ligase activity	

1.3.2 Ionomics

1.3.2.1 Studying the plant ionome

The term *ionome* encompasses all of an organism's essential and nonessential mineral nutrients and trace elements, in a particular tissue and time-point of interest (Lahner et al., 2003). This is an extension of the previously named *metallome* (Outten and O'Halloran, 2001) to include relevant non-metals, such as N and S (Salt, 2004). *Ionomics* therefore is the study of the ionome, whereby the elemental composition of an organism is quantitatively measured, with a specific focus on physiological, developmental and genetic stimuli impacting on its composition (Salt, Baxter and Lahner, 2008). The first large scale ionomic study was conducted by Lahner et al., (2003), who utilised a panel of ~ 6000 fast neutron-mutagenized *A. thaliana* plants. Elements were quantified through a combination of ICP-MS and ICP-OES (see below for information regarding these techniques). This screen analysed approximately 12% of the *A. thaliana* genome and identified 51 mutants with an altered elemental profile This included one mutant containing a deletion of FERRIC REDUCTASE DEFECTIVE 3 (*AtFRD3*); a gene known to regulate the Fe-deficiency response in *A. thaliana*. Previous characterisation of *AtFRD3* demonstrated that mutants accumulate two-fold excess Fe and Zn and four-fold excess Mn in their shoots (Rogers and Guerinot, 2002). Subsequently, the first complete *genome-wide* scan for ionome regulatory genes was conducted in yeast (*Saccharomyces cerevisiae*) by Eide et al., (2005), who utilised a population of 4,385 mutagenized strains; 212 of which were found to exhibit an altered ionomic profile. This study, like Lahner et al., (2003), identified the prevalence of 'multi-element phenotypes', highlighting the complexity and interdependence of elemental interactions which form the basis of the ionome. Various factors are attributed to this (Baxter, 2015). Similar transport mechanisms can have an impact upon multiple elements; the NIP AQPs, for example, facilitate both As and B transport (Bienert et al., 2008; Isayenkov and Maathuis, 2008). Similar chemical properties can also impact upon an organisms ionomic phenotype, as observed for Mg^{2+} and Ca^{2+} (Broadley et al., 2008). Additionally, alterations to specific physiological properties (such as the Casparian strip or roots hairs) may result in the perturbation of multiple elements (Baxter, 2015). For example, *Atesb1* (ENHANCED SUBERIN1)

mutants exhibit increased root suberin, changes to water transpiration (Baxter et al., 2009) and a disordered Casparian strip (Hosmani et al., 2013). The ionome of these mutants is resultantly altered, with decreased accumulation of Ca, Mn and Zn and an increased accumulation of Na, S, K, As, Se, and Mo within the shoot reported (Baxter et al., 2009). Another example of how alterations to the Casparian strip can perturb multiple elements is shown with *Atmyb36.* Again, multiple changes within the ionome are observed, including increased concentrations of Na, Mg and Zn and decreased concentrations of Ca, Mg and Fe (Kamiya et al., 2015). Alterations to phloem structure can also have multi-element effects. For example the metal binding mutant *nakr1* (SODIUM POTASSIUM ROOT DEFECTIVE1), which is specifically expressed within phloem companion cells, exhibits Na, K and Rb (rubidium) accumulation (Tian et al., 2010).

Ionomic studies have also demonstrated the prevalence of correlations between elements (Baxter, 2015), with various studies identifying multiple significant interactions between pairs or groups of elements (*c.f.* Broadley et al., 2008; Bus et al., 2014; Pinson et al., 2015). However, these elemental correlations can vary, and are often dependent upon numerous factors; including tissues, environments and populations under observation (Baxter, 2015). Within the renewable industrial products from rapeseed (RIPR; BB/L002124/1) diversity panel seed B concentrations was negatively correlated with seed S (Thomas et al., 2016), whereas Bus et al., (2014) describe a strong positive correlation between B and Mn concentrations during *B. napus* seedling development. In another study in *B. napus,* Liu et al., (2009) report that B was significantly correlated with Ca, Cu, Mg, P and Zn. Furthermore, under B deficiency conditions, concentrations of Fe, Mg, P and Zn are significantly increased. Within this same study by Liu et al., (2009), QTLs for B/Cu and B/P shoot concentrations co-localised; on chromosomes A4 and A1 respectively. Orthologues of genes known to control P accumulation in response to P deficiency, *AtPHR1* and *AtPHO2* (Delhaize and Randall, 1995; Rubio et al., 2001), were putative candidate genes underlying the QTL for P accumulation, controlling ~ 17% of the phenotypic variation. However, whether B concentrations were affected by these (or indeed any other) shared genes within this QTL region was not reported (Liu et al., 2009). Baxter et al., (2008) also describe the so-called ‘P model’, from which they were able to classify

A. thaliana plants as either responding to P deficiency or sufficiency, based on the shoot concentrations of B, P, Fe, Co, Cu, Zn, and As specifically. Under P deficiency, alterations in concentrations of these elements was observed; with B concentrations increasing significantly. Whilst an explanation for some elemental interactions was provided (e.g. As increase under P deficiency may be a result of increased expression of high affinity phosphate transporters; Samuel, Noble and Parkway, (2004)), mechanisms underlying the interaction between P concentrations and the accumulation of B were not clear. However, removing any one of these elements from the model decreased its effectiveness by < 5%, highlighting its stringency (Baxter et al., 2008). B availability has also been shown to affect nitrate concentrations, with decreased assimilation under B deficiency (Camacho-Cristóbal and González-Fontes, 2007).

As previously mentioned, the plant ionome is not static. Notably it differs in individuals grown in equivalent environments (Broadley et al., 2004; Watanabe et al., 2016), but also between different cellular and extracellular structures (Neugebauer et al., 2016). When grown in comparable environments, differences quantified are invariably due to variation in gene expression, evolved through both mutation and phenotypic adaptation (Neugebauer et al., 2016). Consequently, we can exploit natural populations (rather than forward genetic screening using mutagenized populations) to assess how phenotypic adaptions are optimised to withstand diverse environments. This is exemplified in *A. thaliana* with the Na transporter *AtHKT1,* which has been characterised through DNA microarray bulk segregant analysis (BSA) (Rus et al., 2006) and genome wide association study (GWAS) (Baxter et al., 2010). *A. thaliana* populations growing in more saline conditions (such as coastal habitats) exhibit reduced *AtHKT1* gene function, which is correlated with an increased salinity tolerance (through a modulation of shoot Na accumulation) (Baxter et al., 2010).

When conducting ionomic studies, extensive pools of natural genetic variation have the potential for significant exploitation. Researchers can look to improve nutrient acquisition and utilisation in crops (Veneklaas et al., 2012; White, 2013), with an aim to improve yields in adverse or suboptimal environments. Similarly, when enhancing agricultural systems, nutrient fortification of plants grown for human consumption or animal feeds (*c.f.* Watson et al., 2012) can be improved

by looking for, and exploiting, phylogenetic ionome variation. Therefore, studying natural variation in diverse populations has been an important tool in the functional genomic catalogue and has further advanced the field of ionomics. Further detail on how using such diverse populations can aid the field shall be further addressed within section 1.3.4.

1.3.2.2 Quantifying elemental concentrations

A wide array of analytical techniques has been developed for elemental analysis. Those using inductively coupled plasma (ICP) are particularly prevalent, such as inductively coupled plasma-mass spectrometry (ICP-MS) and optical emission spectrometry (ICP-OES). Both are relatively inexpensive and permit high throughput, allowing for the high sample numbers typical of ionomic studies. A prerequisite of each is a homogenised sample, which is subsequently digested in acid and analysed in solution. Laser ablation (LA-ICP-MS) however, does not require pre-digestion and can allow for two-dimensional analytics, effective when studying specific regions, such as trichromes or root tips. Numerous additional analytical techniques for ionomics exist, for instance, x-ray fluorescence and neutron activation analysis. For a comprehensive review, see Salt, Baxter and Lahner., (2008).

All elemental analyses presented for this thesis were conducted by ICP-MS. ICP-MS was developed in the 1980s and has since become a highly versatile technique for multi-element analysis, with over 50 measurable elements (Pröfrock and Prange, 2012). ICP-MS has a higher sensitivity than ICP-OES and therefore smaller sample weights can be used. For this study ~ 100 mg of material was utilised, although > 20 mg was permissible. This sensitivity is beneficial for researchers, however it is imperative for material to be accurately weighed before digestion (Salt, Baxter and Lahner, 2008). A useful tool utilised throughout this project was the Purdue Ionomics Information Management System (PiiMS; ionomicshub.org/home/PiiMS; Baxter et al. 2007), which details quantified elemental concentrations for mutagenized *A. thaliana* material characterised via ICP-MS. This material includes *A. thaliana* T-DNA insertion lines (Krysan, Young and Sussman, 1999); an extensive collection of mutant lines widely available for order from, for example, the Nottingham *Arabidopsis* Stock Centre

(arabidopsis.info; NASC, University of Nottingham, UK). T-DNA insertions are large (~ 5-25 kb) and generally produce a significant enough effect to knockout individual genes lines (Krysan, Young and Sussman, 1999). This therefore allows researchers to equate whether a mutation has an effect upon a particular phenotype. This technique has been extensively employed for a wide array of studies, including the initial screening of the B transporters *AtBOR1* and *AtNIP5;1* (Takano et al., 2002, 2006). Alcock et al., (2017) similarly utilised T-DNA insertion lines to elucidate the function of *A. thaliana* genes putatively involved in Ca and Mg accumulation. This research was part of the RIPR collaborative research project (Thomas et al., 2016), of which greater detail has been provided within section 1.3.4.

1.3.3 Brassica napus as a model species

1.3.3.1 Origins and speciation

The 'Triangle of U' depicts the evolutionary relationship between six members of the *Brassicaceae* family; three diploid parental species (*B. rapa, B oleracea, B. nigra*) and three hybrid polyploids (*B. napus*, *B juncea, B. carinata*) (Figure 1.4). *B. napus* is an allotetraploid, formed within the last 10,000 years following the hybridisation of the progenitor species *B. rapa* and *B. oleracea* (U.N, 1935), which provided the A and C genomes respectively. Both *B. rapa* (AA) and *B. oleracea* (CC) are closely related, sharing a common ancestor around 3 million years ago (mya) (Inaba and Nishio, 2002). The model species *A. thaliana* is a member of the *Brassicaceae* family and diverged from the *Brassica* genus approximately *20 mya* (Koch, Haubold and Mitchell-Olds, 2001). Extensive genome collinearity between *B. napus* and *A. thaliana* is observed (Cheung et al., 2009).

The genome of *B. napus* is complex. This is compounded by the fact that each of its parental genomes exhibit extensive triplication, indicating their derivation from an original paleohexaploid ancestor (Cheung et al., 2009). Furthermore, a gene found in the A genome is likely to have a corresponding gene within the C genome (defined as a homoeologue) which shares significant sequence similarity (~ 3% nucleotide divergence) (Trick et al., 2009b). In total, it is feasible to find up to six orthologues of each *A. thaliana* gene in *B. napus* (Chalhoub et al., 2014).

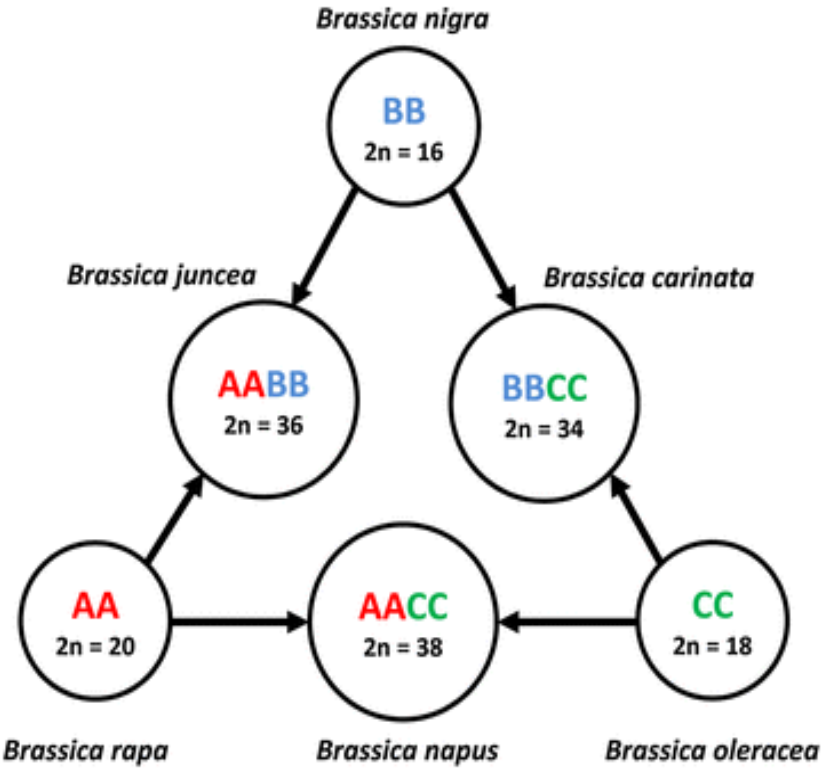


Figure 1.4 Brassica species within the Triangle of U

The Triangle of U (U.N, 1935) depicts the evolutionary relationship between six members of the Brassicaceae family. The three parental species *(B. rapa, B. nigra, B oleracea) are shown, which represent the AA, BB and CC genomes. The hybrid combinations of the parental special are also provided (B. napus, B juncea, B. carinata). B. napus (AACC), was formed following the hybridisation between B. rapa and B. oleracea. The diploid chromosome number (2n) is also provided.*

Koh et al., (2017) adapted from U.N. (1935).

The selection of *B. napus* as an agricultural crop potentially occurred ~ 400 years ago (Morrell, Buckler and Ross-ibarra, 2011). Subsequently, *B. napus* was then grown extensively in northern Europe throughout the 19th century due to its high levels of erucic acid, which were utilised in industrial applications such as the lubrication of steam engines (Wang et al., 2014). Currently OSR is used within bio-fuel industry (Zhang and Malhi, 2010) and as protein for animal-feed (Nesi et al., 2008). However oilseed rape is primarily utilised as an oilseed producer. In terms of global vegetable oil supply, OSR is the third largest supplier behind soybean and palm, whilst in Europe it is the primary vegetable oil source (European Commission, 2018). Leafy vegetables, such as Japanese or Siberian

Kale, and the root vegetable swede, are additional *B. napus* crop types of economic significance (Allender and King, 2010).

1.3.3.2 Breeding programs and the impact on genetic diversity

Over the past decades there have been extensive *B. napus* breeding programmes, which has affected many phenotypic traits. As discussed in Section 4.3.2.2, breeding to lower seed GSL content has been conducted (Allender and King, 2010) which may have perturbed additional seed traits, such as seed B. Improvements to seed oil content, fatty acid composition and overall crop yields (Abbadi and Leckband, 2011) has also been facilitated by extensive OSR breeding programs. However, such strategies have impacted upon allelic diversity (Allender and King, 2010), with Wang et al., (2014) questioning whether there is sufficient allelic variation left to meet increasingly intensive agricultural demands. As such, various studies have looked to assess diversity within the *B. napus* germplasm. A diversity panel containing 509 *B. napus* accessions was collated by Bus et al., (2011), who looked to assess various genotypic factors, including population structure and genetic diversity between crop types. Population structure was determined by STRUCTURE (Pritchard, Stephens and Donnelly, 2000; see Section 1.3.4.2), revealing that winter types, spring types, and swedes could be assigned to three major sub-population clusters (Bus et al., 2011). The clear separation of swedes from winter and spring types was thought to be due to the original speciation of swedes, following a separate interspecific hybridisation event (than that of winter and spring OSR). Allender and King, (2010), for example, discusses the potentiallity of polyphyletic origins of *B. napus*. Furthermore, genetic differences between spring and winter could in part be explained by adaptation to different environments; spring types, for example, do not require vernalisation in order to flower and have low winter hardiness. Resultantly, spring OSR is primarily grown in regions with short summers (e.g. Canada or Scandinavia) or areas lacking sufficient low temperatures to induce flowering (e.g. Australia). Conversely winter and semi-winter OSR are adapted to Western Europe and Asia, respectively. Genetic diversity as reported by the overall fixation index, $F_{ST,}$ was lower for winter OSR, whilst accessions developed more recently, associated with breeding strategies to lower erucic acid and GSL

content, also exhibited reduced genetic diversity (Bus et al., 2011). Other studies have shown lower genetic diversity in low oil erucic acid and low seed GSL (termed double-low/ 00) accession (*c.f.* Cao et al., 2010). Conversely a study by Wang et al., (2014) found that genetic diversity for recently developed accessions had not been reduced in recent years (as was reported by Bus et al., (2011)). Such discrepancies could be due to the specific accessions utilised or genotyping methods. However Wang et al., (2014) did find that winter type OSR, specifically, semi-winter accessions originating from China, exhibited the lowest genetic diversity. They surmised this could be a result of genetic bottleneck of Chinese OSR accessions, given the original germplasm was introduced from Europe (and other regions such as Canada) in the 20th century (Chen et al., 2008). Furthermore, Wang et al., (2014) report two sub-populations in their *B. napus* panel (comprised of 472 accessions; no swede types), with numerous genomic regions showing high F_{ST}. They compared these regions in high F_{ST} with past QTL studies reported for important OSR traits. For example a region on chromosome C2 (with a F_{ST} of 0.241) overlapped with a locus identified by Harper et al., (2012) for seed GSL content (see Section 1.3.4.2) (Wang et al., 2014).

Further understanding into *B. napus* genetic diversity and the impact of breeding on populations utilised for breeding strategies is highly advantageous. Such resources are vital for crop improvement, particularly in the dissection of complex traits (Morrell, Buckler and Ross-ibarra, 2011), and may allow for accelerated breeding strategies and resultantly improved germplasm management (Bus et al., 2011) within the coming decades.

1.3.4 Advanced analytics for plant nutrition studies

1.3.4.1 Genome Wide Association Study

Although QTL studies are important tools in the analysis of complex traits, association mapping studies have been increasingly utilised as an alternative approach. Genome wide association studies (GWAS) can be employed by researchers wishing to scan entire genomes to aid the identification of genetic loci associating with trait(s) of interest. There are various advantages of utilising GWAS rather than QTL analysis (Korte and Farlow, 2013). GWAS, for example,

utilises existing germplasm negating the necessity to produce mapping populations which are prerequisites for QTL studies. Furthermore, GWAS looks to exploit historical recombination between loci and molecular markers in association with a particular trait. Populations undergoing high levels of recombination see a reduction in non-random allele association; a term called linkage disequilibrium (LD) (Slatkin, 2008). As a result, the genome is composed of regions containing alleles in high LD and as such markers associating highly with a trait of interest are likely to be located close to the causative allele. This enables a more focused search in comparison to QTL studies; in plants, linkage analysis typically localises QTLs to 10-20 centimorgan (cM) (Zhu et al., 2008).

The unit of genetic variation studied for GWAS is the single nucleotide polymorphism (SNP) - a single base-pair change observed between individuals at a specific location within the DNA sequence. With modern sequencing technologies, such as Illumina (Goodwin, McPherson and McCombie, 2016), low cost and high throughput sequencing for SNP discovery has aided the advancement of GWAS-based studies. However, for polymorphisms to be identified, a reference sequence is required in order to align and score SNP markers. Unfortunately, the complexity of polyploidy often results in reference sequences being unavailable. In order to address this issue, a novel transcriptome-based method was first employed in *B. napus* to aid SNP-discovery. Trick et al., (2009b) used a set of 94,558 *Brassica* unigenes assembled from publicly available expressed sequence tags (ESTs) (Trick et al., 2009a) as a reference from which to score SNPs against. 41,593 SNPs were identified via the software package MAQ (Li, Ruan and Durbin, 2008), including both inter-homoeologue polymorphisms (IHPs) and hemi-SNPs (Figure 1.5). IHPs represent variation between homoeologues (corresponding genes found between subgenomes of a polyploid species) and as such do not represent true allelic variation. The term hemi-SNP is applied when a polymorphism observed between accessions originates from a single genome, with a majority of SNPs (91.2%) found to be this form. It is also possible to identify simple SNPs in polyploid species, whereby variation is derived from a polymorphism at a single locus (Trick et al., 2009b). This technique proved a suitable approach to detect SNPs in species lacking genomic references, and allowed for the identification of

sequence polymorphisms between two OSR cultivars (Tapidor and Ningyou 7; European winter OSR and Chinese semi-winter OSR respectively).

Following this analysis of sequence variation, Bancroft et al., (2011) were able to construct SNP linkage maps in order to align the *B. napus* genome with *A. thaliana* (based on the similarity of the *Brassica* unigenes and the *Arabidopsis* Genome Initiative (AGI) gene models) and for genome sequence assemblies of its progenitor species (*B. rapa* and *B. oleracea*). This study also looked at transcript abundance from transcriptome data, from which Higgins et al., (2012) were able to build upon; detecting differential expression between homoelogous gene pairs in *B. napus*.

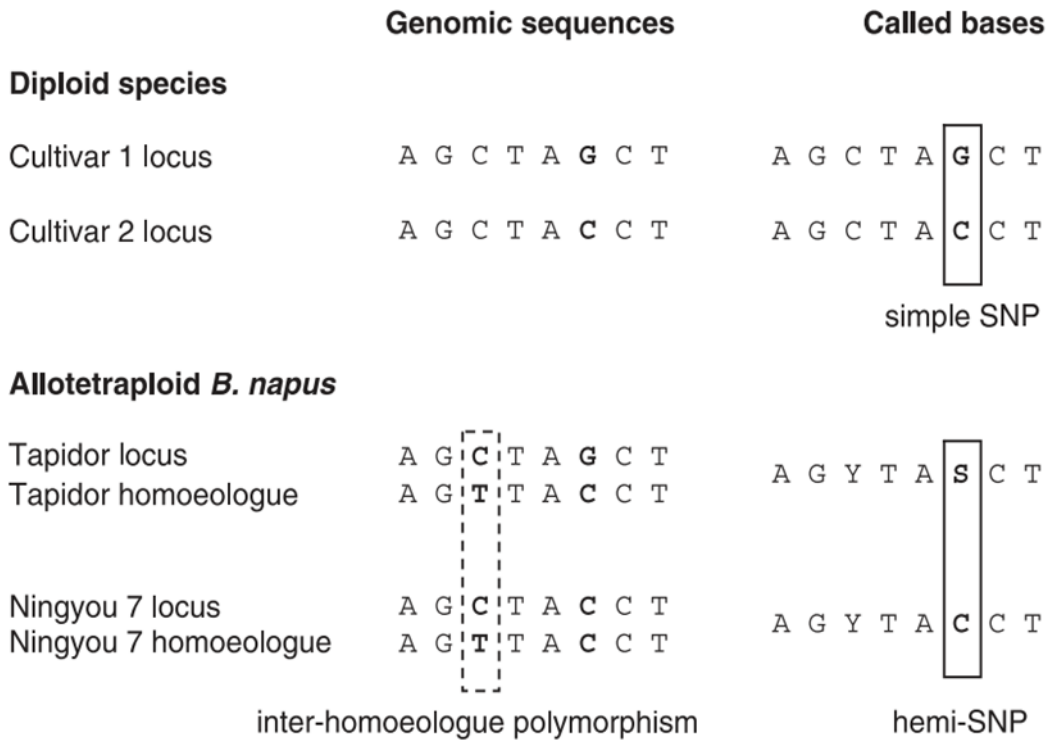


Figure 1.5 Representation of different forms of sequence polymorphisms identified in B. napus

Solid boxes represent allelic SNPs, whereas the dashed boxed show non-allelic SNPs. In polyploid species, such as B. napus, hemi-SNPs are the most frequent form of polymorphisms and occur when a polymorphism is observed between accessions originating from a single genome. Inter-homoeologues polymorphisms are sequence polymorphism between homoelogous gene pairs and do not represent true allelic variation. Simple SNPs can also be identified, where variation is derived from a polymorphism at a single locus. International Union of Biochemistry ambiguity codes: Y=C or T, S= C or G.
Trick et al., (2009b).

1.3.4.2 Associative Transcriptomics

Following work on quantifying sequence variation (Trick et al., 2009b) and transcript abundance (Bancroft et al., 2011; Higgins et al., 2012) in complex polyploids, the novel transcriptome-based GWAS termed Associative Transcriptomics (AT) was developed (Harper et al., 2012). AT utilises mRNA-seq data, following previously described methods (Trick et al., 2009b), and enables researchers to equate whether variation in gene sequence (SNPs) *and* transcript abundance (gene expression markers; GEMs) are associated with a trait of

interest. GEMs are quantified by sequence read depth (as reads per kb per million aligned reads; RPKM) (Harper et al., 2012).

Like GWAS, AT exploits historical recombination and LD to identify loci associated with a trait of interest. Because of this, a large and genetically diverse panel of plants is required. However, as with all forms of GWAS there is potential of identifying spurious associations due to underlying relatedness (kinship; K) and population structure (Q) (Price et al., 2010). To account for this, a number of techniques have been developed. Inference of genetic ancestry through Structured Association (SA) or Principal Component Analysis (PCA) based approaches can be employed to assess population stratification. SA is a model based approach in which samples are assigned to subpopulation clusters (Price et al., 2010) using programmes such as STRUCTURE (Pritchard, Stephens and Donnelly, 2000) or ADMIXTURE (Alexander, Novembre and Lange, 2009). Both are desirable as they return a Q-matrix; whereby for each accession within a diversity panel the proportion of its genotype coming from one of the founder populations is estimated (where $K \geq 2$ populations) (Popescu et al., 2014). However, such programmes are computationally intense with long run times (Price et al., 2010) and therefore non-model PCA approaches, such as EIGENSTRAT (Price et al., 2006), are advantageous. PCA approaches use the top principle components as covariates to correct for population stratification, however a Q-matrix is not produced (Popescu et al., 2014). Furthermore, neither SA or PCA approaches correct for relatedness (Price et al., 2006). Due to these limitations in the aforementioned techniques, a mixture of approaches was adopted within this research. Firstly, PSIKO (Population Structure Inference Using Kernel-PCA and Optimization; Popescu et al., 2014) was used to return the principle components of the dataset (like the PCA approach EIGENSTRAT), however, in addition, this programme generates a Q-matrix. PSIKO is also much (up to 30 times) faster than comparable methods; highly advantageous for large datasets. This is then implemented in a compressed Mixed Linear Modelling (cMLM; Zhang et al., 2010)) approach using GAPIT (Genome Association and Prediction Integrated tool; Lipka et al., (2012)), from which a K-matrix is generated automatically. cMLM is an advantageous approach as compared to older Mixed Linear Modelling (MLM) methods, which can be computationally challenging when conducting association analyses using large datasets.

Conversely cMLM is less computationally involved, but can still account for both population structure and relatedness within the model (Zhang et al., 2010). Additionally, Genomic Controls were used within GEM analyses (Devlin and Roeder, 1999); to correct for spurious associations, p-value adjustments were applied where the genomic inflation factor observed was greater than one (Havlickova et al., 2018; see Section 2.2.2)

In a proof of concept study Harper et al., (2012) employed AT to identify the genetic basis of seed glucosinolate content using a panel of 84 *B. napus* accessions. Unlike the cMLM approach adopted for research within this thesis, SNP associations were performed through MLM in TASSEL (Bradbury et al., 2007), using glucosinolate trait data and a Q-matrix derived by STRUCTURE (Pritchard, Stephens and Donnelly, 2000). GEM analysis utilised the linear regression between RPKM (dependent variable) and glucosinolate trait data (independent variable) (Harper et al., 2012). The existing SNP linkage map (Bancroft et al., 2011) was refined to generate the A and C genome sequence scaffolds, from which 'pseudomolecules' (which represent the 19 chromosomes in *B. napus*) were constructed (Harper et al., 2012). This association analyses identified regions of genome deletion containing orthologues of *AtHAG1*, a known regulator of aliphatic glucosinolate biosynthesis in *A. thaliana* (Hirai et al., 2007), to be causative of the low glucosinolate phenotype observed in certain *B. napus* accessions (Harper et al., 2012).

Recently, a diversity panel of 383 *B. napus* accessions, composed of 127 spring oilseed rape (OSR), 7 semi-winter OSR, 160 winter OSR, 35 swede, 15 winter fodder and 39 unspecific habit (Thomas et al., 2016), was assembled as part of the Renewable Industrial Products from Rapeseed (RIPR; BB/L002124/1) consortium. Leaf transcriptome sequencing (Illumina HiSeq 2000) produced 1.92 x 10^{12} bases of sequence data, which were subsequently aligned to the coding DNA sequence (CDS) gene model 'pan-transcriptome' (He et al., 2015). The use of the AC pan-transcriptome is a recent improvement to the AT pipeline. Rather than using unigenes (Trick et al., 2009b), published sequences of *B. rapa* and *B. oleracea* have been interpolated with any *B. napus* specific CDS models not already represented by orthologues from the progenitor species (resulting in 52,790 and 63,308 CDS gene models ordered in the 'AC pan-transcriptome'

respectively) (He et al., 2015). A total of 355,536 SNPs were detected within the leaf transcriptome, of which 256,397 had a second allele frequency (frequency at which the second most common allele occurs; saf) > 0.01. Transcript abundance for each of the 116,098 CDS gene models (which represent the AC pan-transcriptome) was also quantified, with significant expression (> 0.4 RPKM) detected for 53,889 CDS gene models (Havlickova et al., 2018). Following the generation of functional genotypes, both SNP and GEM associations may be performed, as discussed above (for greater detail regarding implementation and analysis of AT outputs, see Section 2.2).

To summarise, SNP AT utilises the SNP functional genotypes, the inference of population structured by Q-matrix as produced by PSIKO (Popescu et al., 2014) and the trait data of interest. This is then implemented in a cMLM (Zhang et al. 2010) approach using GAPIT, from which a K-matrix is automatically generated (Lipka et al., 2012). The estimation of population structure within the RIPR panel is shown in Figure 1.6 (as determined by PSIKO; Popescu et al., (2014)), where the highest likelihood is a subpopulation of K = 2 (Havlickova et al., 2018). Comparatively, GEM analyses performs fixed-effect linear modelling, with RPKM values and the Q matrix inferred by PSIKO as the explanatory variables (with the use of Q-matrix for GEM analysis an addition from the early AT methodology outlined by Harper et al., (2012)), and with trait data as the response variable. The generation of Manhattan Plots for both SNP and GEM associations then allow for a visualisation of these associations (Havlickova et al., 2018) (see Section 2.2.3).

The increased size of the RIPR diversity panel (383 accessions) is more typical of those used for association studies (Spencer et al., 2009) and has therefore allowed for a diverse array of complex traits to be studied. To date association analyses for seed erucic acid and tocopherol isoform variation (Havlickova et al., 2018), cell wall composition (Wood et al., 2017) and leaf macronutrient accumulation (Alcock et al., 2017, 2018) have been published as part of the RIPR collaboration. The ionomic variation of seed and leaf material within this diversity set has also been published (see 2.2.1 and Thomas et al., (2016)), with data of B concentrations from this forming the foundation of this thesis.

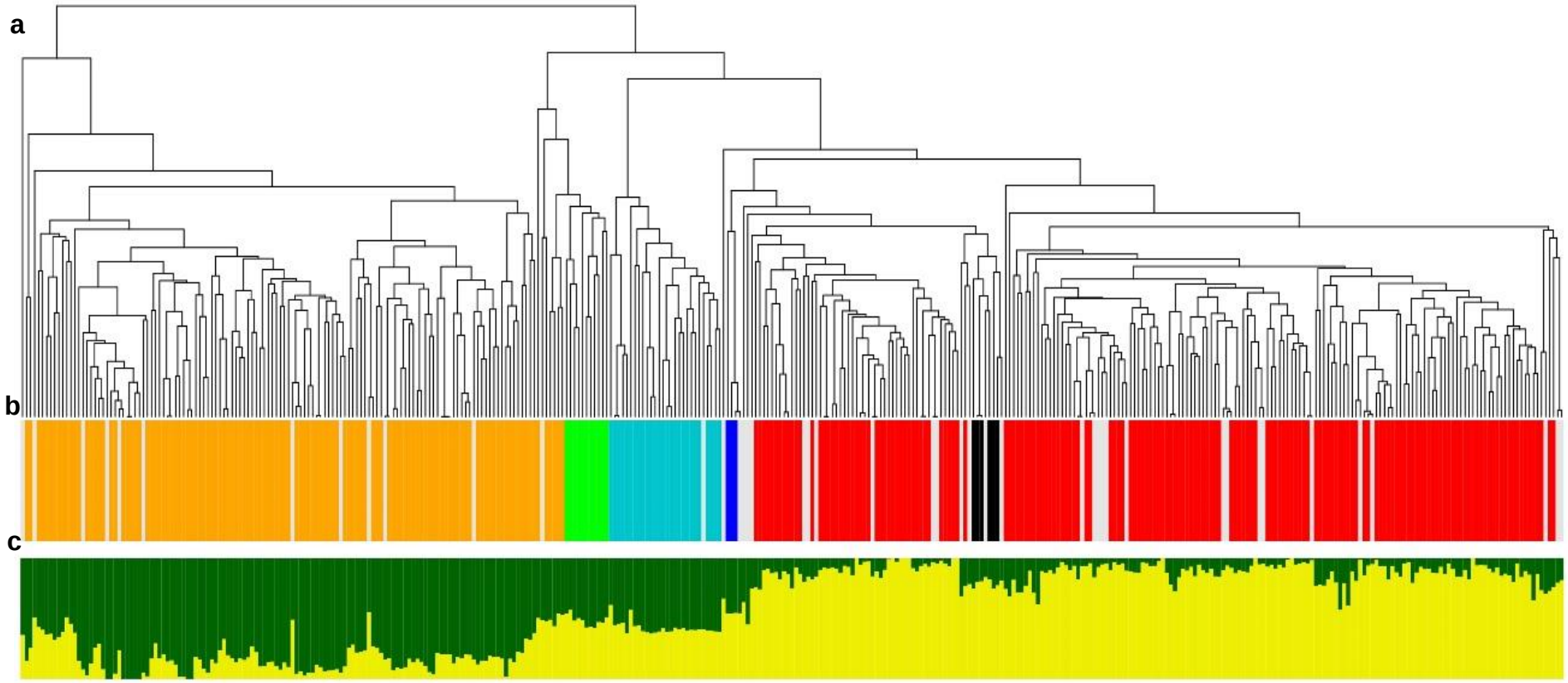


Figure 1.6 Population structure of the Renewable Industrial Products from Rapeseed (RIPR) panel

a. *Dendrogram depicting the relatedness of all 355,536 scored single-nucleotide polymorphisms within the RIPR panel.* ***b.*** *Assigned crop types in the panel: orange for spring oilseed rape, green for semi-winter oilseed rape, light blue for swede, dark blue for kale, black for fodder, red for winter oilseed rape and grey for crop type not assigned.* ***c.*** *Population structure as determined using PSIKO, (Popescu et al., 2014) where the highest likelihood is a subpopulation of K=2 Adapted from Havlickova et al., (2018)*.

Chapter 2: General Methods

2.1 Generation of data sets for Associative Transcriptomic analysis

All datasets utilised for Associative Transcriptomic analyses, including; seed and leaf ionomic analysis, transcriptome sequencing for the functional genotype (SNPs and GEMs) generation and the generation of the Q-matrix to account for population structure, were all collated before this thesis' research began (as outlined in Thomas et al., (2016); Havlickova et al., (2018)). A review of the AT technique, as implemented within this research, can be found in Section 1.3.4.2. Methods in brief are outlined below.

2.1.1 Brassica napus diversity panel

The RIPR *B. napus* diversity panel was utilised for this project; comprising 127 spring OSR, 7 semi-winter OSR, 160 winter OSR, 35 swede, 15 winter fodder and 39 unspecific habit (unassigned crop type), as described by Thomas et al., (2016).

Sowing took place in late October 2013, with seeds initially sown directly into propagation trays containing fine-grade (< 3 mm particle size) compost-based media (Levington Seed & Modular + Sand F2S, Everris Ltd., Ipswich, UK). Plants were grown in a glasshouse vented at 15°C, with supplementary lighting (Philips Master GreenPower SON-T 400 W bulbs controlled by Grasslin Uni 45 timer) maintaining 12 hour day lengths. Following approximately 2 months growth (in January 2014) five plants from each accession were transplanted into individual 5 L pots containing Levington C2 compost (Scotts Professional, Ipswich, UK) and placed into two unheated polytunnels (Visqueen Luminance Skin, Northern Polytunnels, Colne, UK). No additional heating was provided. Pots were arranged in a randomised block design of five replicate blocks, with three replicates in one polytunnel, and two in the second. Plants were watered three times a day using an automatic irrigator controlled by a Hunter Irrigation Controller (Hunter Industries, San Marcos, CA, USA, provided by Hortech Systems Ltd., Holbeach,

UK). Following sampling for leaf ionome analysis (in early March 2014; see below), plants were supplied with Kristalon Red NPK fertiliser (Yara, Grimsby, UK) applied via a direct feed injector (Dosatron D3GL-2, Tresses, France). Plants were bagged before flowering to prevent cross pollination and harvested for seed in July 2014.

2.1.2 Transcriptome Sequencing

RNA extraction occurred at the 2nd true leaf stage, with Illumina transcriptome sequencing carried out according to previously outlined methods (He et al., 2016). RNA-sequencing data was mapped onto the *Brassica* A and C pan-transcriptome (He et al., 2015) by previously described mapping methods (Bancroft et al., 2011; Higgins et al., 2012). A total of 355,536 SNPs were detected, of which 256,397 had a saf >0.01. Transcript abundance (GEMs) for each of the 116,098 CDS gene models was quantified as reads per kb per million aligned reads (RPKM), with significant expression (> 0.4 RPKM) detected for 53,889 CDS models (Havlickova et al., 2018).

2.1.3 Quantification of tissue mineral concentration through Inductively Coupled Plasma- Mass Spectrometry

ICP-MS allows for the quantification of over fifty elements. For this a thoroughly acid-digested sample is ionised with inductively coupled plasma (usually a noble gas such as Argon). The ions are then separated and quantified based on the mass to charge ratio by mass spectrometry (Pröfrock and Prange, 2012). See Section 1.3.2.2 for further detail regarding ICP-MS analysis.

Leaves were sampled at rosette stage (6-8 true leaf stage) and freeze-dried for 48 hours. 200 mg of homogenised dry-weight material was digested in a Multiwave 3000 platform (Anton Paar GmbH, Graz, Austria), with 2 mL Trace Analysis Grade 70% nitric acid (HNO_3), 1 mL Milli-Q H_2O and 1 mL H_2O_2 (hydrogen peroxide) [power = 1400 W, temperature = 140°C, Pressure = 2 MPa, time = 45 minutes]. Operational blanks and certified reference material (CRM; Tomato SRM 1573a, NIST) were included in the digestion runs. Samples were diluted with 11 mL Milli-Q H_2O to a final volume of 15 mL for storage. For

elemental analysis, samples were diluted 1:5 using Milli-Q H_2O and analysed via ICP-MS (Thermo Fisher Scientific iCAPQ).

Dried seeds (3-4 per digestion) were allowed to pre-digest overnight in 1.16 mL Trace Analysis Grade 70% HNO_3, indium internal standard and 1.2 mL H_2O_2. Samples were digested in a dry heat-block at 115°C for 4 hours then diluted to 11.5 mL with Milli-Q H_2O. Elemental analysis was performed via ICP-MS (Perkin Elmer NexION 300D).

In order to calculate specific elemental concentrations for each data-point, an average of the operational blanks was subtracted. This value was then multiplied by the sample volume, divided by the initial dry mass of digested material and converted to mg element per kg^{-1} dry weight. Percentage recovery of CRM was calculated from an average of the CRM processed concentrations. This was divided by the CRM consensus value and multiplied by 100 (Boron = 88.7% CRM recovery). Additionally, element-specific limits of detection (LOD) were calculated as three standard deviations (SD) of the operational blank concentrations (B; leaf = 1.87 mg kg^{-1}, seed = 28 mg kg^{-1}).

Further details of RIPR plant growth and sampling conditions; acid digestion protocols; elemental analysis via ICP-MS and data processing is described in Thomas et al., (2016).

2.2 Associative Transcriptomics

2.2.1 SNP association analysis

SNP association analysis was performed according to previously described methods (Harper et al., 2012) with modification allowing for the increased size of this data set. Inference of population structure by Q-matrix was produced using PSIKO (population structure inference using kernel-PCA and optimisation; Popescu et al., (2014). Analysis was performed in R, using the package GAPIT (Lipka et al., 2012) for compressed Mixed Linear Modelling (cMLM; Zhang et al. (2010)). Manhattan Plots of SNP associations were filtered to include only those with saf > 0.01 (256,397 SNPs), with light red and grey markers representing hemi-SNPs for which a genomic position could not be assigned, with red and

black markers representing simple and hemi-SNPs assigned with confidence to a genomic position (Havlickova et al., 2018). Both the 5% false discovery rate (FDR; Benjamini & Hochberg 1995) and Bonferroni (Dunn, 1961) corrections were used to set significance thresholds at $p < 0.05$.

Using a subsection of the 383 diversity panel, it was possible to perform SNP predictions to assess how allelic variation affects observed phenotypic traits (i.e. B concentration). For this, a '*test panel'* of 109 accessions was utilised (which has had no AT analysis performed) for testing the predictive nature of the most highly associated SNP markers chosen from the 274 '*training panel*' (for which AT had been performed). The RIPR panel was split roughly so that two thirds of all accessions were present in the 274 panel and one third in the 109 panel, with a comprehensive representation of all crop types found within each of the two panels. The most highly associated SNP markers within an association peak were ran through the custom R script (SNPfinderV2), which assigned the particular allele present in each accession in both the 274 and 109 panels. Next the allelic effect for each SNP was compared to the trait data. For this, the average of the trait data was calculated for each of the alleles observed in the 274 panel. This average was then multiplied by ± half of the allelic affect (i.e. the decreasing allele = -0.5* allelic affect; the increasing allele = +0.5* allelic affect; if three alleles were present then generally the allele with the median trait value would be assigned as zero, with the remaining two alleles assigned as previously stated). Data from the 274 panel was then standardised according to the 109 panel via subtracting the mean and dividing by the SD. Subsequently, data was rescaled to the 274 panel by multiplying by the SD and adding the mean (of the 274 panel). The predictive nature of the markers was then assessed by relating the standardised and rescaled data to the trait data; the correlation coefficients (R-value) and two-sided t-test (p-value), performed by the data analysis software GENSTAT 19th edition (VSN International, Hemel Hempstead, UK), were used to ascertain whether the markers were significantly predictive of observed variation in B phenotype.

2.2.2 GEM association analysis

GEM association analyses were carried out using previously described methods (Harper et al., 2012) with modifications. Custom R scripts were used to perform fixed-effect linear modelling, using the PKISO-derived Q-matrix and RPKM values as explanatory variables, and trait data as the response variable. Genomic controls (Devlin and Roeder, 1999) were applied to correct for spurious associations where the genomic inflation factor observed was > 1. For each regression, R^2 regression coefficients, constants and significance values were calculated (Havlickova et al., 2018). Manhattan Plots were generated, as described in section 2.2.1, with significant expression plotted for 53,889 CDS models.

The predictive ability of the most highly associated markers was assessed through splitting the 383 diversity panel split into the 'training' and 'test' panel (with these being the same panels as described above). From the GEM association analyses, both the gradient and intercept of the regression were outputted, which were then standardised and rescaled to the 109 panel. For this, the RPKM values of accessions in the 109 panel were multiplied by the intercept (from the 274 regression analysis) and the gradient added. The predictive nature of each marker was then assessed by again calculating the correlation coefficient and performing a two-sided t-test in GENSTAT, using the new 'predictive' RPKM and the original trait data.

2.2.3 Candidate gene identification and selection

The AT pipeline, summarised in Figure 2.1, generates graphical Manhattan Plots depicting the association between either SNP or GEM markers and the phenotypic trait of interest. Both SNP and GEM markers are aligned on the x-axis based on the pseudomolecule CDS gene model order, with the y-axis showing the significance of the association (as –log10P). In order to assess statistical stringency, two type one error tests were included: 1) the 5% false discovery rate showing the point at which 5% of markers are expected to be false positives (Benjamini & Hochberg 1995) and 2) the Bonferroni corrected threshold (Dunn, 1961). However, in some instances, neither the FDR nor Bonferroni significance thresholds were reached, an example of which is shown in Figure 2.2

(specifically, this shows seed B AT). In such cases, visually-determinable association peaks showing a clear clustering of markers, were analysed for potential genes related to the trait of interest (e.g. labelled 1 in Figure 2.2). Additionally, single standalone GEMs above the background noise of the markers were also analysed (e.g. labelled 2 in Figure 2.2). Significance thresholds were not lowered in instances where the Bonferroni or FDR were not reached. This was decided in order to maintain consistency, both within this research (i.e. analysis of Manhattan plots between seed and leaf), and between additional AT studies conducted using the RIPR diversity panel (Alcock et al., 2017, 2018; Havlickova et al., 2018). For example, AT analyses for both P and K accumulation by Alcock et al., (2018) similarly returned Manhattan Plots lacking both the FDR/Bonferroni corrected significant thresholds and were successfully analysed for markers related to the trait of interest.

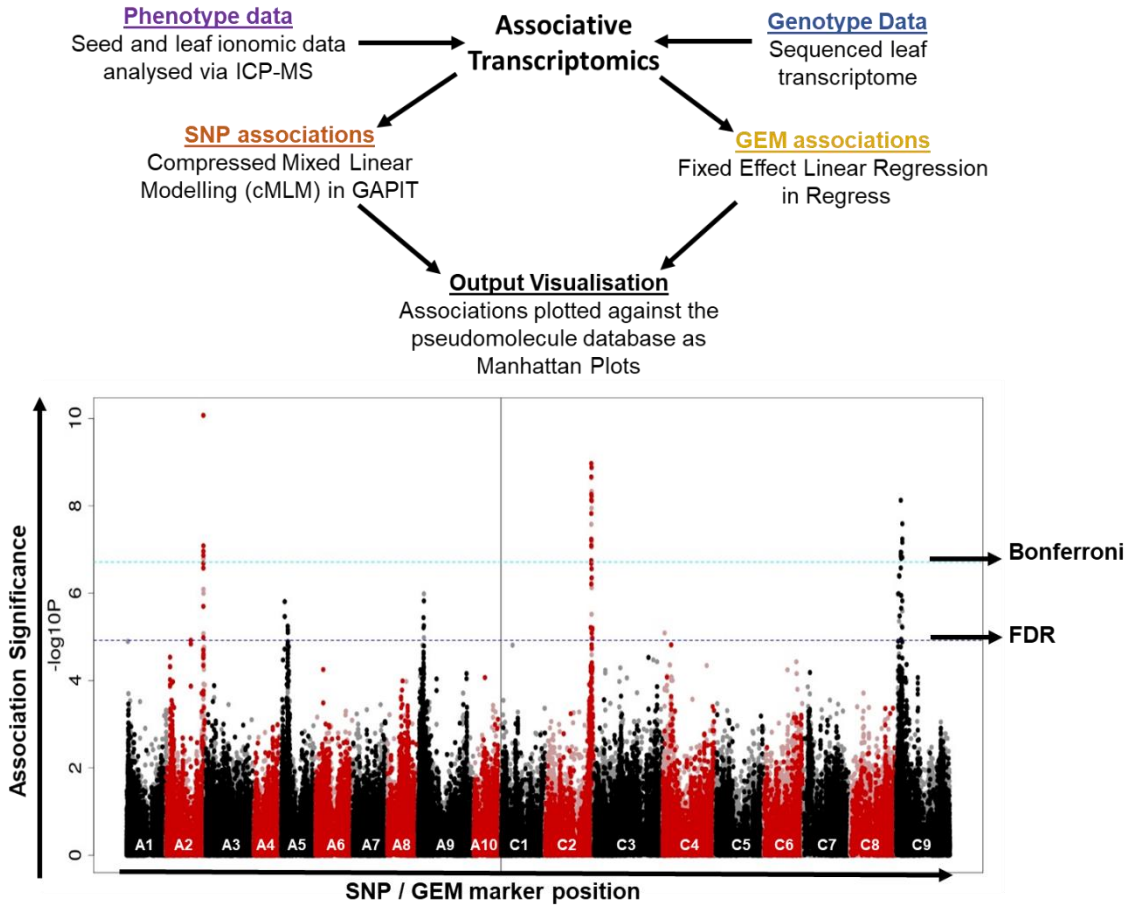


Figure 2.1 Associative Transcriptomic Pipeline

AT analysis allows for the identification of genetic loci controlling a trait of interest using phenotypic and genotypic information from the large RIPR diversity panel. Methods are outlined in Harper et al., (2012), with modifications allowing for the increased size of the RIPR diversity panel (Havlickova et al., 2018). SNP associations were carried out in R, utilising the package GAPIT (Genomic Association and Prediction Integrated Tool; Lipka et al., (2012)), which performs compressed mixed linear modelling (Zhang et al., 2010), with the PSIKO derived Q-matrix accounting for population structure (Popescu et al., 2014) and the auto-generation of the K-matrix. GEM associations were conducted in R using the in-house script Regress to perform fixed effect linear modelling, with the PKISO derived Q-matrix and RPKM values as explanatory variables, and trait data as the response variable. Manhattan Plots were generated, depicting the –log10P values of the SNP or GEM associations on the y-axis and the pseudomolecule position on the x-axis. The trait presented here is a SNP Manhattan Plot depicting the proportion of seed B out of the total seed nutrient profile from the 383 panel. The shift between black and red depicts a change in chromosome number (with the chromosome number provided; A1-A10 and C1-9), with the solid vertical line showing the split between the A and C genomes. For SNP associations, the light red and grey markers represent hemi-SNPs for which a specific genome cannot be assigned (marker may be present in either the A or C genome). Red and black markers represent simple and hemi-SNPs assigned with confidence to a genomic position. The grey and blue horizontal lines mark the 5% false discovery rate (Benjamini & Hochberg., 1995) and 0.05 Bonferroni significance (Dunn., 1961) used to set the significance thresholds to p <0.05.

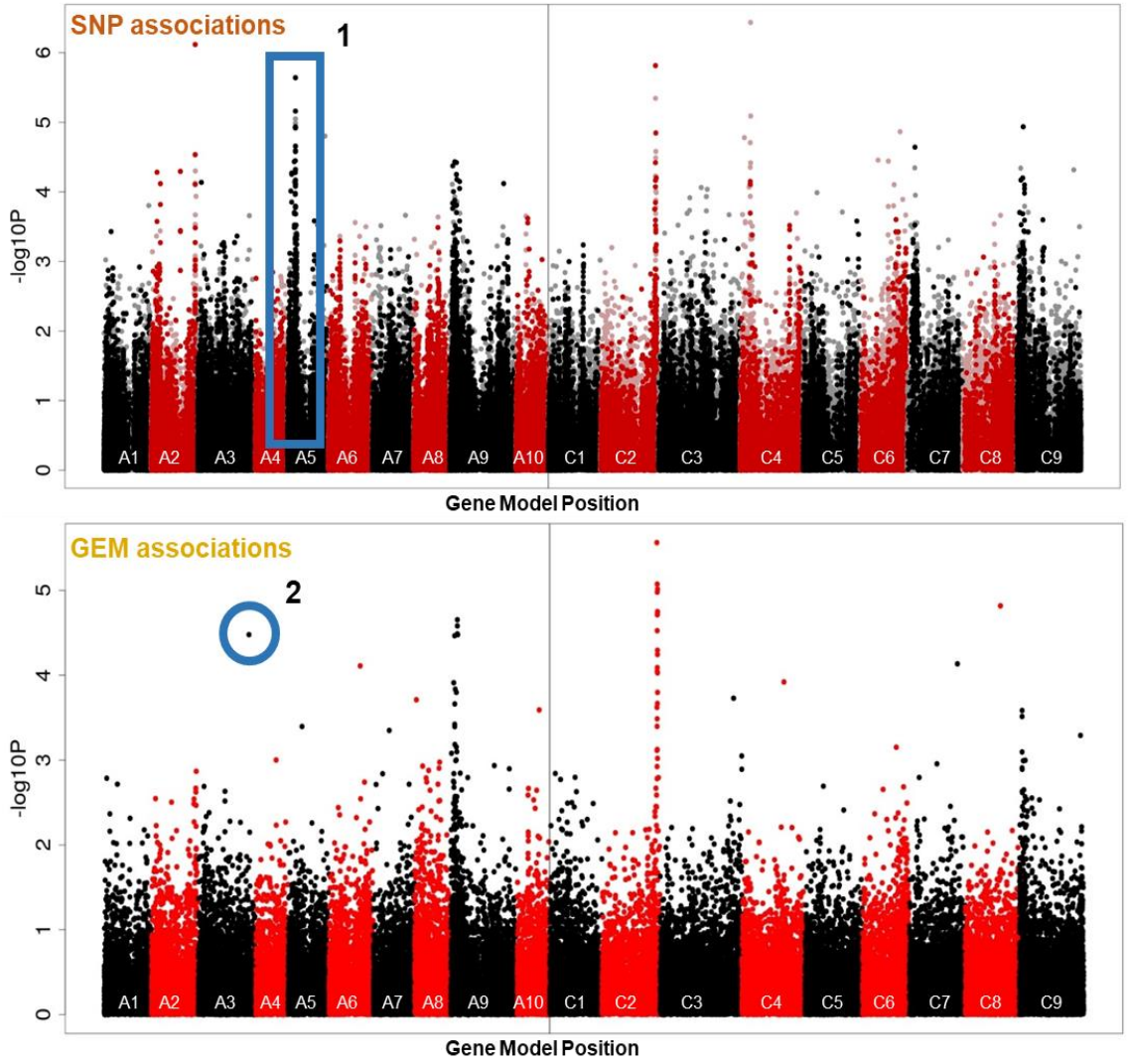


Figure 2.2 Analysis of Manhattan Plots lacking FDR or Bonferroni corrected significance thresholds

Manhattan Plots can either depict single nucleotide polymorphisms (SNPs; top) or transcript abundance (GEMs; bottom) in association with trait variation. The trait presented in this example is seed B concentrations in the 383 panel. Both SNP and GEM markers are aligned on the x-axis based on the pseudomolecule CDS gene model order, with the significance of the association (as –log10P) on the y-axis. In some instances neither the FDR nor Bonferroni corrected significance thresholds were reached (shown here as lacking both grey and blue dotted horizontal lines respectively). In such cases 1) clearly defined association peaks with a clear clustering of markers were analysed for genes related to the trait of interest and 2) single standalone GEMs above the background of the markers were also analysed.

Within SNP analyses the most highly associated SNP markers (i.e. the top markers within an association peak) were tested for trait predictability (see Sections 2.2.1 and 2.2.2). However, in association mapping studies it is unlikely the most highly SNP marker will be the underlying causative gene of interest. Therefore it is essential to look at genes within linkage disequilibrium (LD; Ecke et al., 2010) of the most highly associated markers within all defined assocaition peaks. Based on previously estimated LD decay in *B. napus*, a large number of CDS models were within LD of highly associated SNP markers (> 100 CDS gene models) (Alcock et al., 2017). An approximate region of ± 1.50 megabases (Mb) around significant SNPs has been reported as a threshold for identifying candidate genes within LD in *B. napus* (Wang et al., 2016), however this threshold was limited to ± 1.00 Mb within this research. This would typically equate to an approximate region of ± 150 CDS gene models around the most highly associated SNP markers, which were searched for candidate genes related to the trait of interest. The ordered pan-transcriptome based on the *B. napus* A and C CDS gene models (He et al., 2015) aided this identification of putative candidate genes causative of the observed associations (e.g. previously linked to B accumulation or could affect the ionomic profile of the tissue under observation). The *A. thaliana* orthologues of the *B. napus* markers were subsequently viewed on the *Arabidopsis* Information Resource (TAIR; Lamesch et al., (2012)). Furthermore, the PiiMS database (ionomicshub.org/home/PiiMS; Baxter et al. (2007)) was used to identify T-DNA lines in *A. thaliana* which had been previously characterised via ICP-MS. This was a useful tool when assessing whether a particular candidate had any documented ionomic disruption and consequently would not warrant further characterisation within the scope of this research.

In instances where no *A. thaliana* orthologue of the *B. napus* marker was present within the pan-transcriptome, the CDS of the relevant marker was compared with all other markers in the pseudomolecule database (using an in-house script). The unannotated marker was also compared across different species using NCBI BLAST (Basic Local Alignment Search Tool; Altschup et al., 1990). This technique was particularly relevant when analysing unannotated GEMs. Given that GEMs show the correlation between gene expression and the trait of interest, novel transporters may be elucidated using this technique.

Candidate genes were not selected for characterisation in *A. thaliana* if they already had a well-documented role in the control of B, or if ionomic profiles were available on PiiMS. Certain candidates which could be putatively linked to B (e.g. through cell wall organisation, control of root traits, water transpiration) were taken forward for further study, although it is acknowledged few candidates identified had previous research which specifically links them to the control of B concentrations. Certain candidates were identified in multiple association peaks or in both SNP and GEM analyses, and therefore were higher confidence candidates to characterise further. Candidates tested in *A. thaliana* have been summarised in Table 2.1, with greater detail provided in Table 3.6 and Table 4.5. Furthermore, additional candidate genes identified from AT analysis, whilst not characterised in *A. thaliana*, which may be controlling B concentrations were identified. A summary of these candidates is provided in Table 3.7 and Table 4.6.

The use of AT for complex trait analysis has been delineated in Havlickova et al., (2018), from which further details on methods are outlined.

2.3 Analysis of *A. thaliana* candidate genes

Candidate genes identified from AT analyses, which could be potentially affecting leaf or seed B concentrations, were tested using T-DNA insertion lines in their respective *A. thaliana* orthologue. From this, any ionomic disruption within the T-DNA line in comparison to the wild type control validated the candidate gene.

2.3.1 Plant material utilised

A list of all T-DNA insertion lines available for each candidate gene was provided on TAIR (Lamesch et al., 2012), with the position of the insertion viewed on the 'SeqViewer Nucleotide' function. SALK (Alonso et al., 2003) or SAIL (Sessions et al., 2002) T-DNA lines were chosen with a preference for exon T-DNA lines as opposed to intron (minimising the chance of the T-DNA being spliced out) or promoter insertions. Seed for the chosen T-DNA insertion lines was obtained from the Nottingham *Arabidopsis* Stock Centre (NASC; Nottingham, UK), with confirmed homozygous lines for an insertion (as opposed to a segregating line)

obtained where possible. A list of T-DNA lines is provided in Table 2.1. The Colombia-0 ecotype (Col-0) was used as a wild type control.

2.3.2 Growth of A. thaliana plant material

Once obtained from NASC, seed was imbibed at 5°C for 48 hours to break dormancy and promote uniform germination. Seeds were sown into soil (John Innes F2) within P24 trays (with each individual cell within the P24 being 5 cm^3) and watered daily. Mutant lines and Col-0 controls were grown in separate trays and arranged at the same level (i.e. on the same shelf) in a growth room using LED lighting (100 µM m^3) set to sixteen hours of light and eight hours dark. On average, temperatures within the growth room ranged from 20°C to 23°C. After 10 ten days growth, well-established seedlings were randomly thinned to one plant per pot of a P24 tray. gDNA was sampled from each plant (Section 2.3.3), with all lines subsequently genotyped via PCR (Section 2.3.5). Aracons and Aratubes (Arasystems, BETATECH BVBA, Ghent, Belgium) were placed on the plants to improve seed harvesting. Upon full silique development, watering was stopped and plants were allowed to dry completely before seed was collected into paper bags from which the seeds were threshed.

2.3.3 Material sampling for genotyping and gDNA extraction

Plants were sampled for gDNA extraction following ~ 3 week's growth. One leaf was excised from each plant with scissors and placed into a pre-labelled 2 mL Eppendorf tube (scissors were sterilised between each sample with 100% Ethanol; EtOH)**.** Material was placed on ice whilst sampling and stored at -20°C prior to gDNA extraction.

gDNA was extracted using a modified CTAB (Cetyl trimethylammonium bromide) extraction method. 500 µL 2x CTAB buffer [2% CTAB, 1.4 M NaCl, 100 mM Tris-Cl, 20 mM Na-EDTA] was added to frozen leaf samples, with samples then homogenised using a tissue lyser and metal beads. Samples were then incubated at 65°C for one hour. Next, 300 µL of chloroform: isoamyl alcohol (IAA; 24:1 ratio) was added, then vortexed thoroughly. Samples were centrifuged for five minutes at 14,000 g with approximately 400 µL from the upper aqueous layer removed

into a new Eppendorf. 1 mL of EtOH NaAc (ethanol sodium acetate) (960 µL EtOH + 40 µL 3M NaAc) was added, samples were inverted and then left for 30 minutes at room temperature, or alternatively overnight at -20°C. Samples were centrifuged for five minutes at 14,000 g with all supernatant removed. Pellet was rinsed with 500 µL 70% EtOH and centrifuged again. All supernatant was finally removed and pellet air died. gDNA was re-suspended in 100 µL dH_20 and stored at -20°C.

2.3.4 Primer design

Genomic primers for mutant screening were designed using the SALK T-DNA primer design tool (signal.salk.edu/tdnaprimers). 5' (left primer; LP) and 3' (right primer; RP) gene specific primers for each mutant line, along with SALK LBb1.3 and SAIL LB1 left border primer (LbP), were then obtained from Integrated DNA Technologies (IDT). Right border primers were not used for genotyping as homozygous mutants were obtained for all candidate genes in at least one T-DNA insertion line using the LbP. See Table 2.1 for primer sequences for each T-DNA mutant line.

Table 2.1 Summary of A. thaliana T-DNA insertion lines and genotyping primer sequences

The candidate locus is provided, as is the name given throughout this research and the how each gene is linked to the B phenotype. The T-DNA lines utilised, respective NASC ID code and insertion site for each T-DNA is stated. The 5' and 3' primers used for T-DNA genotyping are also provided. Information on the Columbia-0 control and left border primers has also been summarised

Locus	Given Name	Description/ putative link to trait	T-DNA	NASC ID	Insertion site	5' genomic primer	3' genomic primer
AT5G62000.3	ARF2	Controlling auxin development and various root traits	SALK_035537	N535537	Exon	TGTAAAGGTTC ACAAGCAGGG	CCAAAACCAAA AGTCAGCAAG
AT2G35610.1	XEG113_a	Cell wall organisation; root hair development	SALK_007511C	N685144	Exon	CAGAAAGCCAG TGAAATGGAC	TTGACCAGAAC CTACCAAACG
	XEG113_b		SALK_066991C	N662800	Exon	AATCTTTCTTCT CGCTCCTGC	ACAATGCAGGA GGTTTCATTG
AT2G36390.1	BE3	Amylopectin biosynthesis	SALK_030954C	N670497	Promoter	TACCAGTTGTA GCCGTTGTCC	TAAATTCGTCAT TGCCGAAAC
AT5G44480.1	DUR_a	Cell wall organisation; lateral root development	SALK_143736C	N658447	Exon	TAGCCACAAAA CCAATCCATC	AAAGCTCAGCC TCGAAAAGTC
	DUR_b		SALK_000741C	N683078	Intron	GCAGAATTTGG TAGCTTTGACC	AGCAAAATGCA TTACAGCGTC
AT1G16880.1	ACR11	Activator of GS2	SALK_206346C	N695904	Intron	CGTATCTTTTGC TGGAATTCG	TCTGTACGTGA CCGAGGAAAC
AT1G17890.1	GER2	Cell wall organisation	SALK_020145C	N653073	Exon	CCCAGTAAATT CCCAAACACC	AATGATAATTTC CACCCGGAG
AT3G28715.1	ATP_a	Transmembrane transport	SALK_093787	N593787	Exon	TTCCAGGAGGA AAACAAGGAC	GGTACGTAGAG TTGTACTGTCC G
	ATP_b		SALK_078674C	N657969	Exon	AAGAACTTAGG TTGCCACCAAG	TTGGTGATTTG GGAGAGTTTG

Locus	Given Name	Description/ putative link to trait	T-DNA	NASC ID	Insertion site	5' genomic primer	3' genomic primer
AT2G39510.1	UMAMIT14	Seed development	SALK_037123C	N685665	Exon	CCAAGGCAAAG GAATTAAAGG	GTTTCCATCAT CATCACGACC
AT2G37170.1	PIP2;2	Water transport; upregulated under B deficiency	SAIL_169_A03	N871747	Coding Region	CAACCATAAGC CTACCAAAAGG	TTATAGATTAC GGCAGCTCCG
AT2G34390.1	NIP2;1_a	Proposed facilitator of B transport (not confirmed *in planta*)	SAIL_439_A02	N820193	Exon	ACCTGACCACA TTTCGCATAC	GAAGACGCTTT GATGTTGAGC
	NIP2;1_b		SALK_023890	N523890	Coding Region	CCCAATGATTA GCCCCTCTAG	CGACCTGTGTC TTTTGCTCTC
AT1G01490.2	HMT_a	Heavy metal transport; metal ion transport	SALK_035720C	N674436	Exon	TTTTCTATCCGG AAGAGGAGC	CAAAATTCGTT GATACTCCCG
	HMT_b		SALK_060068C	N675415	Promoter	AACGGTATTTCA GCCACACAC	CAGCATGTTTT TATTCGGCTG
AT5G58350.1	WNK4	Phosphorylation; both SNP/GEM candidate	SALK_201692C	N689300	Coding Region	GCCCATAGCAT CATTGGTATG	GACCCTGAATG GGAAGAACTC
AT3G43960.1	CPS	Root hair development	SALK_152721.4	N652721	Promoter	ACCAATCTCAA CCACAAGCAC	CCCACTTTACC AATCCATGTG
	Colombia-0			N6000	Wild Type Control		
			LB1.3		SALK Left Border Primer		ATTTTGCCGAT TTCGGAAC
			LB1		SAIL Left Border Primer		GCCTTTTCAGA AATGGATAAAT AGCCTTGCTTC C

2.3.5 Polymerase chain Reaction (PCR)

Genotyping was carried out using a double PCR approach. LP and RP were used to amplify wild type product for the gene of interest (i.e. lacking T-DNA insertion). LbP1.3/LB1 and RP were used to check for presence of T-DNA insertion. Through a combination of the two PCR reactions it was possible to detect WT (product only in LP and RP reaction), heterozygous (product in each reaction) and homozygous (product only in LbP1.3/LB1 and RP reaction). Alternatively, a PCR reaction could utilise the three primers which would show all three genotyping combinations in one gel. See Figure 2.3 for protocol visualisation. Col-0 was used as a control for each reaction.

DNA was amplified following a standard 20 µL reaction containing 10 µL 2xMasterMix (Thermoscientific 2x MasterMix; Fisher Scientific UK Ltd), 8 µL dH_20, 0.5 µL LP primer and 0.5 µL RP/LbP primer. 1 µL of gDNA was added to each PCR reaction. Reactions were cycled using a Touchdown PCR programme to promote annealing specificity [94˚C for 5 minutes (94°C for 30 seconds, 60°C for 30 seconds {decrease by 1˚C every cycle}, 72°C for 1 minute) x 10 cycles, (94°C for 30 seconds, 55°C for 30 seconds, 72°C for 1 minute) x 29 cycles, 72°C for 7 minutes, 7°C hold]. After amplification, PCR products were electrophoresed on a 1% agarose gel and visualised under ultraviolet (UV) light.

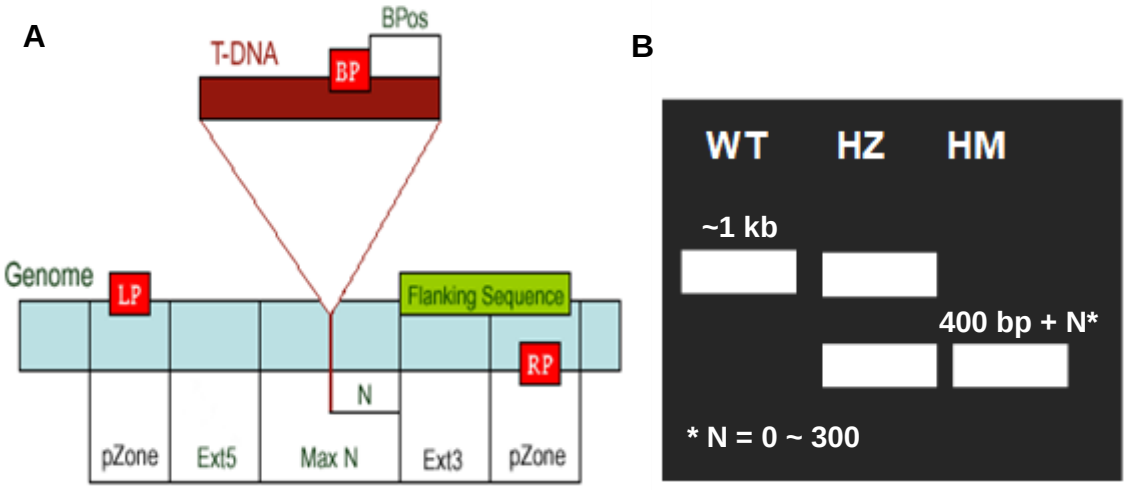


Figure 2.3 Protocol for SALK and SAIL T-DNA primer design

a. *The protocol for SALK/SAIL T-DNA primer design. The right primer (RP) zone is found on the 3' end of the insertion (flanking sequence) and the left primer (LP) zone is found on the 5' end. N; difference of the actual insertion site and the flanking sequence position usually ~0 - 300 bp. MaxN; maximum difference of the actual insertion site and the sequence, default 300 bp. pZone; regions used to pick up primers, default 100 bp. Ext5, Ext3; regions between the MaxN to pZone, reserved not for picking up primers. LP, RP; left and right genomic primers. BP; T-DNA border primer. LB - the left T-DNA border primer. BPos; the distance from BP to the insertion site. (Image adapted from signal.salk.edu/tdnaprimer)* ***b.*** *For genotyping, left border primers LbP1.3/LB1 and RP PCR combination will identify homozygous (HM) individuals, whereas LP and RP will identify wild type (WT). When a band is show in both PCR reactions, or where two bands are observed from a PCR reaction containing all three genotyping primers, there is a heterozygous (HZ) mutation. WT product would normally be in the region of ~1 kb, whereas HM product would be in the region of ~ 400-700 bp.*

2.3.6 A. thaliana ionome analysis

Plants for ionome analysis were grown under conditions outlined in Section 2.3.2. In brief, plants were sown into soil (John Innes F2) and grown under long day conditions in a growth room which utilised LED lighting (100 µM m^3). All candidates analysed via ICP-MS were once again genotyped to ensure homozygosity for the T-DNA insert, following methods outlined in Section 2.3.5. Genotyping PCRs for homozygous T-DNA insertion confirmation for lines

analysed via ICP-MS can be found in Appendix 3.1 and Appendix 4.1 (leaf and seed ICP-MS candidates respectively).

2.3.6.1 Leaf ionome candidates

Plants grown for leaf ionome analysis were sampled after approximately four weeks growth, (when rosette leaves were fully established, but prior to stem bolting). Plants were removed from soil and roots excised, with material thoroughly washed using Milli-Q H_2O to remove soil residue. Material was then oven-dried at 90°C for 48 hours. Samples were stored under air-tight conditions with silica gel included to restrict moisture damage prior to further analysis. A minimum of six individual plants for each mutant line and Col-0 control (the control line was kept constant across analyses, i.e. Col-0 material from original N6000 NASC stock) were then analysed by ICP-MS. For this, 100 mg of homogenised leaf material (all rosette leaf material sampled was homogenised) was weighed from each plant for each line. If 100 mg of material was not available, then the maximal quantity of material available was used. HNO_3 digestions were subsequently carried out as described in section 2.4. Candidates were analysed together (i.e. mutant and Col-0 control) in the same ICP-MS run. All leaf candidates were analysed across three ICP-MS runs at the University of Nottingham.

2.3.6.2 Seed ionome candidates

Plants grown for seed ionome analysis utilised Aracons and Aratubes (Arasystems, BETATECH BVBA, Ghent, Belgium) to prevent cross-pollination and improve seed harvesting. Aracons were placed on each plant whilst at rosette stage, with the Aratube placed upon stem bolting. Upon full silique development, watering was stopped and plants were allowed to dry completely before seed was collected into paper bags from which the seeds were threshed. As it was not possible to obtain 100 mg of seed per individual plant, seeds were pooled from ~15 plants according to line (T-DNA or Col-0). A minimum of six digestions containing 100 mg of pooled seed was then digested for each T-DNA and respective Col-0 control, according to methods described in section 2.4. All seed

candidates were analysed together across two ICP-MS runs at the University of Nottingham.

2.3.7 Statistical analyses

Leaf and seed material from *A. thaliana* mutants was characterised via ICP-MS (see Section 2.4), with data processed following methods outlined in Section 2.1.3. A minimum of six biological replicates for both T-DNA and mutant lines were analysed; leaf elemental concentrations were analysed from individual plants whereas seed was pooled according to line. Significant differences in elemental concentrations was then assessed using a two-sided t-test. Pearson correlation coefficients between B and additional elemental concentrations were also calculated. Differences were considered significant at $p < 0.05$. All statistical analysis was carried out using GENSTAT 19th edition (VSN International Ltd, Hemel Hempstead, United Kingdom).

2.4 Nitric acid digestion and Inductively Coupled Plasma- Mass Spectrometry analysis of *A. thaliana* and swede plant material

100 mg of seed or homogenised dried plant tissue was placed in a 15 mL centrifuge tube (polypropylene; Greiner Bio-One) and left overnight to pre-digest in 4 mL Trace Analysis Grade 70% nitric acid (HNO_3; Fisher Scientific). In some instances 100 mg material was not available (due to poor plant growth, for example), in such cases the maximum quantity of material obtained was digested. Two operational blanks (HNO_3 only) and one CRM (*B. oleracea* Wepal IPE 132, LGC Standards Limited; Table 2.2) were also included in each digestion run. Samples were digested in a dry heat-block at 100°C for 150 minutes. Once completely cooled, sample digest was poured into a new 15 mL centrifuge tube containing 11 µL Milli-Q H_2O. Samples were filtered using 11 µm 70 mm filter paper (Munktell) and were then suitable for long-term storage at room temperature.

Prior to ICP-MS analysis, samples were diluted 1:10 in Milli-Q H2O (1 mL sample: 9 mL Milli-Q H2O) into ICP-MS tubes. Samples were subsequently analysed by ICP-MS (Thermo Fisher Scientific iCAPQ) at the University of Nottingham, with data then processed as described in section 2.1.3 according to previously outlined methods (Thomas et al., 2016; Alcock et al., 2017).

Percentage recovery for B within CRM samples for all ICP-MS runs from which *A. thaliana* and swede material was analysed was > 93%, with a minimum percentage recovery for analysis of ≥ 85%.

Table 2.2 Consensus values for Brassica oleracea IPE 132

Certified reference material IPE 132 (LGC Standards Limited) was used as an external standard for HNO_3 digestions and subsequent ICP-MS analysis. Consensuses values were calculated from a minimum of 16 data-points, with a coefficient of variation (CV %) < 25%. Boron, for example, shows a consensus value of 30.9 mg/kg from 174 data points (N). (MAD = median of absolute deviation). Adapted from Certificate of Analysis: WEPAL IPE 132 Brassica oleracea

Method: Inorganic Chemical Composition								
Element	**Unit**	**Mean**	**Std.Dev.**	**CV %**	**N**	**Median**	**MAD**	**95 % confidence limits**
As	µg/kg	289	70.0	24.2	30	288	49.0	263 - 315
B	mg/kg	30.9	2.63	8.5	174	31.0	1.80	30.52 - 31.31
Ba	mg/kg	3.95	0.397	10.0	23	3.85	0.280	3.78 - 4.12
Ca	g/kg	25.4	1.95	7.7	266	25.4	1.32	25.18 - 25.65
Cd	µg/kg	118	17.6	15.0	77	120	12.0	114 - 122
Cl (as Cl)	g/kg	6.51	0.478	7.3	65	6.52	0.319	6.39 - 6.63
Co	µg/kg	125	28.5	22.7	42	124	19.3	116 - 134
Cu	mg/kg	4.49	0.788	17.6	252	4.52	0.535	4.39 - 4.58
Fe	mg/kg	256	31.1	12.2	249	254	21.0	251.9 - 259.6
Hg	µg/kg	12.5	2.55	20.4	41	12.7	1.72	11.7 - 13.3
K	g/kg	38.3	2.76	7.2	265	38.4	1.91	37.99 - 38.66
Mg	g/kg	1.49	0.117	7.9	266	1.50	0.081	1.472 - 1.500
Mn	mg/kg	25.7	2.62	10.2	270	25.9	1.80	25.44 - 26.06
Mo	µg/kg	2130	208	9.8	53	2100	143	2070 - 2190
N - Kjeldahl (as N)	g/kg	35.2	1.95	5.5	181	35.1	1.33	34.93 - 35.50
N - NO3 (as N)	mg/kg	4650	587	12.6	34	4610	423	4440 - 4850
Na	mg/kg	1530	155	10.1	181	1540	109	1512 - 1558
Ni	µg/kg	808	146.5	18.1	74	803	100.0	774 - 842
P (as P)	g/kg	6.31	0.431	6.8	260	6.30	0.295	6.25 - 6.36
S (as S)	g/kg	9.85	0.870	8.8	137	9.88	0.609	9.71 - 10.00
Se	µg/kg	156	26.3	16.8	20	160	18.5	144 - 168
Sr	mg/kg	87.5	5.47	6.3	24	87.2	3.55	85.1 - 89.8
Zn	mg/kg	27.6	2.68	9.7	269	27.9	1.85	27.32 - 27.97
Method: Real totals								
Element	**Unit**	**Mean**	**Std.Dev.**	**CV %**	**N**	**Median**	**MAD**	**95 % confidence limits**
C - elementary	g/kg	401	10.6	2.6	35	402	7.0	397.8 - 405.0
N - elementary	g/kg	37.9	2.21	5.8	81	37.6	1.47	37.4 - 38.4
Method: Acid extractable (So-called totals)								
Element	**Unit**	**Mean**	**Std.Dev.**	**CV %**	**N**	**Median**	**MAD**	**95 % confidence limits**
Al	mg/kg	228	55.4	24.3	39	226	39.0	210 - 246

Chapter 3: Identifying genetic loci controlling leaf boron concentrations in *B. napus*

3.1 Introduction

A plants requirement for B within its leaves can vary widely. Variation occurs both within and between species, with various factors requiring consideration; including accession, environmental conditions or tissue under investigation. Oertli, (1993), for example, reported a 100-fold difference in tomato (*Lycopersicon esculentum)* leaf B concentrations between the base and tips of older leaves. As an extension to this, Huang, Ye and Bell, (1996) found that leaf B critical concentrations in *B. napus* varied dependent upon both the tissue sampled and its age (where critical concentration was based upon the Smith, (1986) definition as *the minimum concentration of the nutrient in a specified plant part required for producing 90% of maximum biomass or plant yield*). Young leaves exhibited much lower critical concentrations in this study (relative to successive older leaves), with an approximate concentration of 10 - 14 mg/kg (dry weight). In another study, Dell and Huang, (1997) found mean leaf B concentrations in *B. napus* accessions grown under sufficiency as 17 mg/kg (and 10 mg/kg when grown under deficiency). More recently, a comparison of B concentrations between two *B. napus* accessions; one determined B-efficient and one B-inefficient, found concentrations in leaves of 25-day old OSR plants of ~ 60 mg/kg when grown under B sufficiency. Conversely under B deficiency, mean B concentrations was found to be < 10 mg/kg. Whilst no significant difference in B concentrations was observed between accessions across the two treatments, the efficient accession did exhibit greater shoot weight when grown under low B (as compared to the inefficient accession) (Pommerrenig et al., 2018).

Levels of salinity can affect B accumulation; with increasingly saline conditions decreasing leaf B concentrations. Conversely increased transpiration (when salinity is lower) leads to leaf B accumulation (Ben-Gal and Shani, 2002). Additionally, the specific sugar alcohol produced by the plant can affect B transport and ultimate concentrations within the leaves. B-polyol complexes allow for ready B-translocation (Shelp et al., 1995), with species that translocate

sorbitol as their primary assimilate exhibiting increased B-mobility (Brown and Hu, 1996). Brown and Shelp, (1997) found wide variation between leaf B concentrations in different developmental stages (expanding, mature and old leaves), dependent on whether sorbitol or mannitol was the primary sugar within the plant.

Hu, Brown and Labavitch, (1996) assessed leaf B concentrations in fourteen species, including wheat, barley and *B. oleracea* (broccoli and cauliflower) and *B. rapa* (turnip) crop types. Notably they found that leaf B concentration was positively correlated with cell wall pectin concentration. Given that the key role of B *in planta* is the cross-linkage of RG-II pectic polysaccharide molecules (Kobayashi, Matoh and Azuma, 1996), it is unsurprising that plants which have higher pectin concentrations subsequently have higher B requirements. Generally, dicots (which includes *B. napus*) have higher percentages of pectin relative to monocots and therefore exhibit higher B requirements. The requirements of these fourteen species, as stated by Hu, Brown and Labavitch, (1996), is summarised in Table 3.1. Asparagus and onion, which are both monocots, actually had the greatest B requirements relative to the other species studied. However, both are particularly pectin-rich species (relative to both monocots and dicots) and therefore exhibit higher requirements. The authors do however state than a slight over-estimation in their leaf B requirements may have been reported due to the variation in sampling methods for these two species specifically (due to smaller sample-sizes whole shoots of asparagus and whole tops of onion were used in the analysis, rather than just leaf blade material). Furthermore, only one accession was used within the study, which may reduce the replicability of these results, given that a wide variation in B-concentrations between different accessions may be exhibited. Assessing B requirements from multiple accessions within each species would have increased the validity of these results. For example, Baxter et al., (2012) reported leaf B concentrations across 96 *A. thaliana* accessions sampled at 5-weeks old. Hydroponically grown plants had B concentrations ranging from 41.75 mg/kg - 75.50 mg/kg, whilst soil grown plants ranged from 60.81 mg/kg - 338.38 mg/kg. Furthermore, within this research, leaf B concentrations varied 5.40-fold across the 383 accessions under investigation (Figure 3.1), with a heritability (effect of genotypic variation on phenotypic variation within a population) estimation of 0.22 (Thomas et al., 2016).

Table 3.1 Leaf B (mg/kg) requirements in fourteen plant species

The species (and cultivar if provided) and common name of fourteen plant species is given, along with their classification as either a monocot or dicot. The requirement of B (mg/kg dry weight) as reported by Hu, Brown and Labavitch, (1996) is also provided.

Species	Common Name	Classification	Leaf B Requirement (mg/kg dry weight)
Hordeum vulgare	Barley	Monocot	4 - 10
Triticum aestivum	Wheat	Monocot	4 - 10
Zea mays var. sacchara	Sweetcorn	Monocot	4 - 10
B. oleracea var. italic	Broccoli	Dicot	18 - 25
Cucumis sativa	Cucumber	Dicot	22 - 26
Lycopersicon esculentum	Tomato	Dicot	22 - 27
B. oleracea var. botrytis	Cauliflower	Dicot	25 - 35
B. oleracea var. acephala	Collard	Dicot	25 - 35
Raphanus sativa	Radish	Dicot	25 - 35
Daucus carota var. sativus	Carrot	Dicot	25 - 30
Pisum sativum	Pea	Dicot	25 - 35
B. rapa var. rapifer	Turnip	Dicot	25 - 35
Allium cepa	Onion	Monocot	30 - 45
Asparagus officinalis	Asparagus	Monocot	43 - 55

3.2 Results

3.2.1 Leaf B concentration variation and elemental correlations across the RIPR diversity panel

Variation in leaf B across the RIPR diversity panel varied 5.40-fold, from 6.70 to 36.21 mg/kg (dry weight), with the means for each crop type as follows; winter OSR 12.47, semiwinter OSR 12.78, swede 13.15, winter fodder 13.43, spring OSR 14.97. Spring OSR accessions had higher mean leaf B concentrations than winter OSR ($p < 0.001$), however no significant difference was observed for mean leaf B concentrations for other crop types (Figure 3.1).

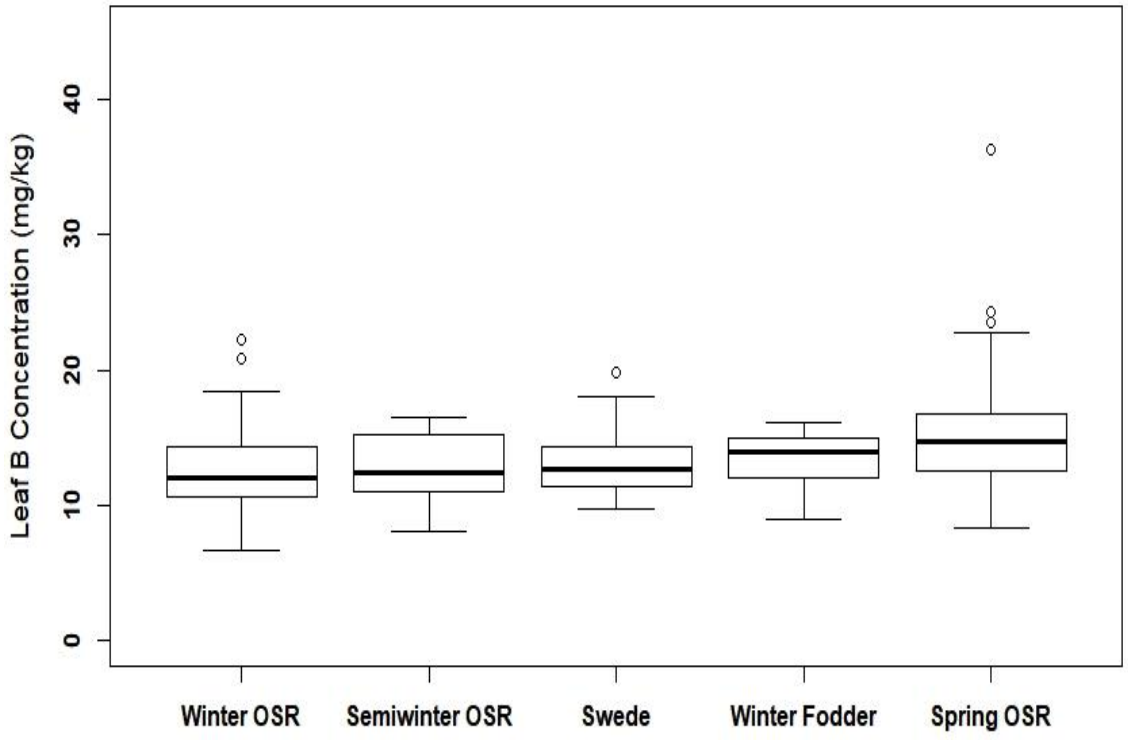


Figure 3.1 Leaf B (mg/kg) concentration across B. napus crop types

Data presented shows the means of leaf B concentrations of 343 B. napus accessions within the RIPR diversity panel whose crop type had been assigned, including; 160 winter OSR, 7 semiwinter OSR, 35 swede, 15 winter fodder and 128 spring OSR. Boxes represent the first and third quartiles with the median drawn. The whiskers represent the lowest and highest data points < 1.5 times the upper and lower quartiles. Circles represent outliers > 1.5 times the upper and lower quartiles. Boxes are organised by mean leaf B from lowest (left) to highest (right).

B concentrations in leaf were significantly positively correlated with nine of the eleven essential elements tested (all p <0.001) (chloride and nitrate data not available) (Table 3.2). B is most highly correlated with Mn, with a strong relationship between leaf Mn/B also identified within another *B. napus* diversity panel (Bus et al., 2014). To briefly explore this further, AT analyses was also conducted for leaf Mn concentrations, with markers within three SNP peaks (A7, A10, C9) found to co-localise to the same regions seen for leaf B (see Appendix I for leaf Mn Manhattan plots). Further discussion of potential candidates within these specific regions is provided in Section 3.3.2.4.

Table 3.2 Correlation between leaf B and other essential elements

Correlation coefficient for leaf B with eleven essential nutrients (n = 383). Significant values are in bold and underlined.

Element Vs B	Correlation Coefficient (r)	Significance (p)
Ca	0.333	**<0.001**
Cu	0.297	**<0.001**
Fe	0.230	**<0.001**
K	0.281	**<0.001**
Mg	0.310	**<0.001**
Mn	0.578	**<0.001**
Mo	0.426	**<0.001**
Ni	0.029	0.574
P	0.246	**<0.001**
S	0.383	**<0.001**
Zn	0.086	0.092

3.2.2 Leaf B associative transcriptomics: outputs and predictions

Association analyses for leaf B concentrations across the 383 diversity panel is shown in Figure 3.2. Only two SNPs, found in association peaks on A7, with a saf > 0.01, passed the Bonferroni (p = 0.05) corrections (Figure 3.2a). The light-coloured SNP markers observed on chromosome C6 represent hemi-SNPs for which the genome of the polymorphism cannot be assigned (marker position may be on A7 or C6). Additional association peaks whose markers passed the 5% FDR correction were seen in chromosomes A3, A6, A9, A10, C2, C4, C8. Furthermore, minor association peaks were observed in chromosomes A2 and C9. No GEMs passed the Bonferroni or FDR (Figure 3.2b).

The predictive capabilities of SNP markers were assessed by splitting the 383 diversity panel in two, forming a '*training panel*' on which AT was performed, and a '*test panel*' in which no AT was performed (see Section 2.2.1 for implementation and analysis of SNP predictions). The correlation coefficient and significance of the SNP predictions performed for leaf B are provided in Table 3.3.

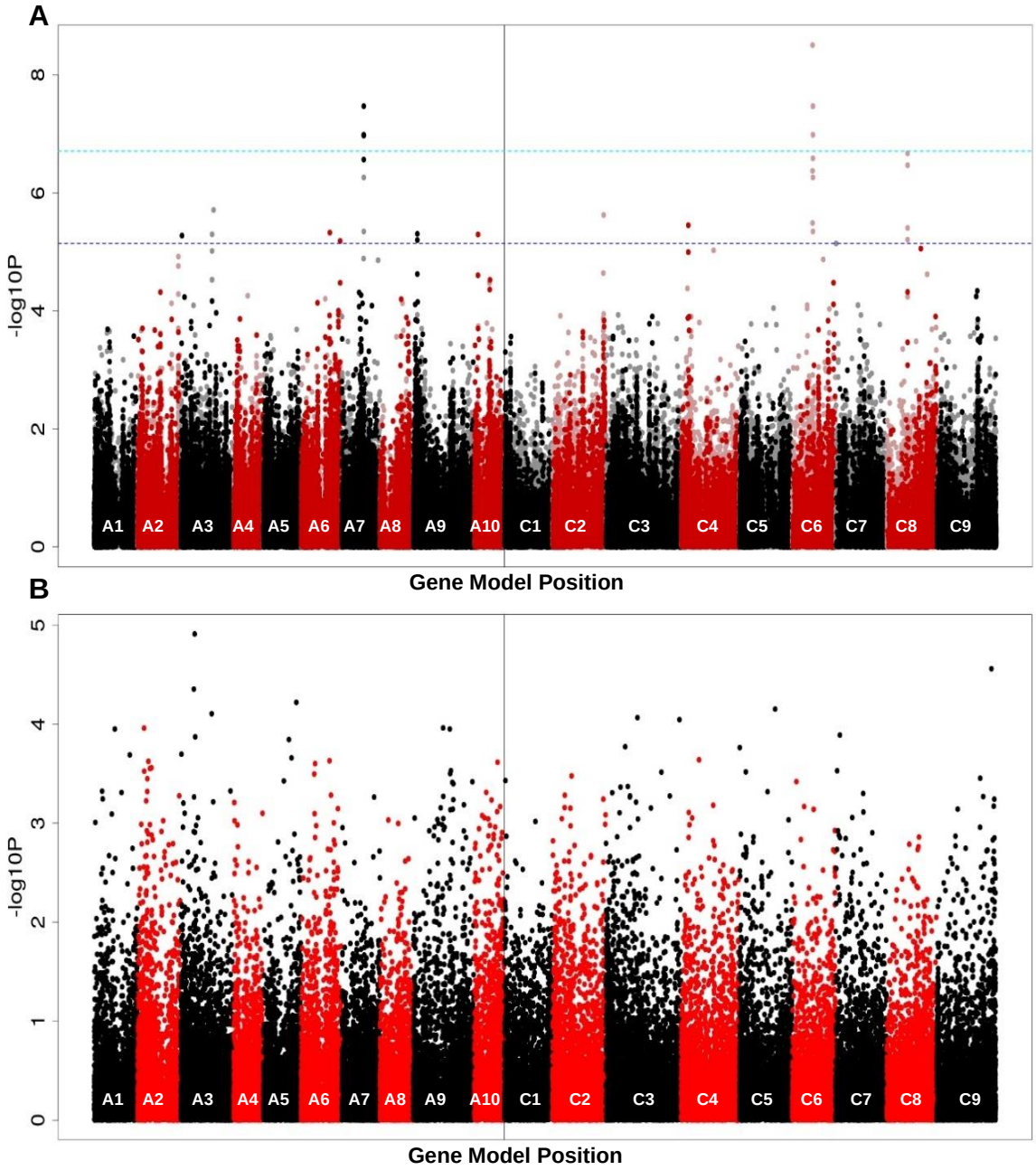


Figure 3.2 Genome wide distribution of (a) SNP and (b) GEM markers associating with B concentrations (mg/kg) in leaves from 383 accessions

SNP associations were calculated with the R script GAPIT (Lipka et al., 2012) using compressed mixed linear modelling, with inference of population structure by Q-matrix produced using PSIKO (Popescu et al., 2014). GEM associations were performed using Regress, performing fixed-effect linear modelling, using PKISO-derived Q-matrix and RPKM values as explanatory variables, and leaf B as the response variable. The –log10P value of SNP and GEM associations were plotted against the pseudomolecule database (representative of the nineteen B. napus chromosomes; labelled A1-A10 and C1-C9) based of the CDS gene model order. In the SNP analyses, light red and grey points represent hemi-SNP markers that have not been linkage mapped to a genome, whereas black and red markers represent markers assigned with confidence to a genomic position. The grey and blue lines on the SNP Manhattan Plot show the 5% FDR (Benjamini & Hochberg 1995) and the Bonferroni corrected significance threshold of 0.05 (Dunn, 1961) respectively.

Table 3.3 Predictive capability of single nucleotide polymorphism markers from leaf B 274 AT analysis

The most highly associated SNP markers within association peaks were tested for their predictive capability of leaf B concentration. The SNP marker and the –log10P from the 274 AT analysis are given. Next the mean leaf B concentration (mg/kg) for accessions in the panel with the reference allele and the most frequent allele variant are shown, with the significance (p) in the difference of these concentrations. The blue and pink colouration shows relative low and high B concentrations respectively. Finally, the correlation coefficient (r) and significance (p) of the predictions using the 109 panel are shown. Significant values are in bold and underlined.

Association Peak	SNP	AT 274 -log10P	B (mg/kg): Reference Allele	B (mg/kg): Allelic Variant	B (mg/kg): Allelic Variation Significance (p)	109 Prediction: Correlation Coefficient (r)	109 Prediction: Significance (p)
A3	Cab004562.1:255:T	4.200	13.332	18.805	**<0.001**	-0.041	0.669
A6.1	Cab006587.2:564:G	3.180	12.714	14.750	**<0.001**	**0.386**	**<0.001**
A6.2	Cab042239.1:1176:G	5.260	13.285	19.737	**<0.001**	-0.001	0.993
A7	Cab012686.1:1527:A	6.640	13.312	24.032	**<0.001**	-0.035	0.722
A9	Cab046445.1:2445:T	5.470	13.239	18.006	**<0.001**	**0.242**	**0.011**
A10	Cab029600.1:363:T	5.580	13.319	14.544	0.391	0.145	0.134
C2	Bo2g163990.1:1902:T	5.220	12.653	14.189	**<0.001**	**0.281**	**0.003**
C4	Bo4g026630.1:1541:T	2.958	12.709	14.640	**<0.001**	**0.203**	**0.034**
C6	Bo6rg076770.1:831:T	6.640	13.312	24.032	**<0.001**	0.091	0.345
C8	Bo8g066850.1:237:G	4.100	13.356	20.288	**<0.001**	0.015	0.881

Only four SNP markers (those on association peaks A6.1, A9, C2, and C4) of the ten lead markers analysed were found to be predictive of leaf B phenotype. Cab006587.2:564:G, the orthologue of which in *A. thaliana* is DOWNY MILDEW RESISTANT 1, a homoserine kinase (*DMR1*; AT2G17265.1), producing O-phospho-L-homoserine (HserP) used for methionine and cystathionine γ-synthase for threonine biosynthesis (Lee et al., 2005). Cab046445.1:2445:T, the orthologue of which in *A. thaliana* is one of the subunits of the vacuolar H^+-ATPase (*ATP*; AT3G28715.1), is involved in the acidification of intracellular compartments. Such proton pumps have been linked to elemental transport in addition to other ions and metabolites (Sze et al., 2002). Bo2g163990.1:1902:T, the orthologue of which in *A. thaliana* is a DEAD-box RNA helicase (*RH7*; AT5G62190.1) is involved in RNA metabolism (Huang et al., 2016). Finally, Bo4g026630.1:1541:T the orthologue of which in *A. thaliana* is AT6G25070.1. This gene is described as 'neurofilament light protein', however additional functional information is unavailable. Of these four SNP markers shown as predictive for leaf B concentration, the marker Cab046445.1 was taken for further study using *A. thaliana* T-DNA insertion lines.

Although no GEMs passed the Bonferroni or FDR, predictions using the most highly associated markers were similarly performed by splitting the 383 diversity panel into the training and test panels. For this, the most significantly associated GEMs on each chromosome were chosen and have been summarised in Table 3.4. Where GEMs had notably lower –log10P values than other markers (e.g. chromosome C8), no markers were tested for trait predictability. Again, Manhattan Plots for AT analyses on the 274 panel used for trait predictions can be found in Appendix I.

Table 3.4 Predictive capability of gene expression markers from leaf B 274 AT analysis

The most highly associated GEMs were tested for their predictive capability of leaf B concentrations. The GEM and the –log10P from the 274 AT analysis are given, with the correlation coefficient (R) and significance (p) of the predictions using the 109 panel shown. Significant values are in bold and underlined. Finally, the A. thaliana orthologue and TAIR (Lamesch et al., 2012) description is provided.

Chromosome	GEM	AT 274 -log10P	Prediction: Correlation Coefficient (r)	Prediction: Significance (p)	*A. thaliana* Orthologue	*A. thaliana* Annotation
A1	Cab020459.1	3.21	**0.211**	**0.026**	AT4G37610.1	BTB AND TAZ DOMAIN PROTEIN 5
A2	Cab032697.1	2.71	**0.253**	**0.007**	AT5G58350.1	WITH NO LYSINE (K) KINASE 4
A3	Cab002565.1	4.10	**0.310**	**<0.001**	AT1G05870.4	PROTEIN OF UNKNOWN FUNCTION
A3	Cab004528.1	3.61	**0.346**	**<0.001**	*No orthologue*	--
A3	Cab002472.4	3.45	**0.380**	**<0.001**	AT5G10140.4	FLOWERING LOCUS C
A3	Cab003267.1	3.03	**0.197**	**0.038**	AT2G45660.1	SUPPRESSOR OF OVEREXPRESSION OF CO 1
A3	Cab015407.1	3.02	**0.340**	**<0.001**	AT5G03040.3	IQ-DOMAIN 2
A4	Cab019228.1	3.50	0.039	0.681	AT3G55240.1	PLANT PROTEIN 1589 OF UNKNOWN FUNCTION
A5	Cab031463.1	3.23	**0.344**	**<0.001**	AT3G43960.1	CYSTEINE PROTEINASES SUPERFAMILY PROTEIN

Chromosome	GEM	AT 274 -log10P	Prediction: Correlation Coefficient (r)	Prediction: Significance (p)	*A. thaliana* Orthologue	*A. thaliana* Annotation
A8	Cab036025.1	3.02	0.110	0.244	AT4G34030.1	3-METHYLCROTONYL-COA CARBOXYLASE
A8	Cab040722.1	2.76	**0.218**	**0.026**	AT4G30935.1	WRKY DNA-BINDING PROTEIN 32
A9	BnaA09g49870D	4.73	0.065	0.501	AT1G06570.1	PHYTOENE DESATURATION 1
A9	Cab021504.1	2.78	**0.270**	**0.004**	AT3G50530.2	CDPK-RELATED KINASE
C2	Bo2g028880.1	2.76	**0.264**	**0.005**	AT5G58050.1	SHV3-LIKE 4
C3	Bo3g032510.1	3.61	**0.232**	**0.015**	AT2G38480.1	CASP-LIKE PROTEIN 4B1
C3	Bo3g036530.1	3.31	0.179	0.056	AT2G43290.1	MULTICOPY SUPPRESSORS OF SNF4 DEFICIENCY IN YEAST 3
C4	Bo4g046140.1	3.20	0.083	0.386	AT2G32150.1	HALOACID DEHALOGENASE-LIKE HYDROLASE SUPERFAMILY PROTEIN
C4	Bo4g046140.1	3.20	**0.268**	**0.004**	AT1G69420.1	DHHC-TYPE ZINC FINGER FAMILY PROTEIN
C5	Bo5g002450.1	2.82	**0.435**	**<0.001**	AT1G01490.2	HEAVY METAL TRANSPORT/DETOXIFICATION SUPERFAMILY PROTEIN
C5	Bo5g010440.1	2.92	**0.261**	**0.006**	AT1G09970.1	LEUCINE-RICH RECEPTOR-LIKE PROTEIN KINASE FAMILY PROTEIN
C7	Bo7g014510.1	2.85	**0.434**	**<0.001**	AT2G17410.1	AT RICH INTERACTION DOMAIN 2
C9	Bo9g174920.1	4.20	**0.259**	**0.006**	AT5G07960.1	ASTERIX-LIKE PROTEIN

GEMs found to be predictive of leaf B concentration were also correlated with RPKM values from the 383 panel (n = 381) in order to specifically assess the effect of gene expression on leaf B concentrations (i.e. does increased gene expression lead to increased leaf B concentrations, or vice versa). All GEMs tested were significantly correlated with leaf B concentrations; increased expression for twelve GEMs (Cab020459.1, Cab032697.1, Cab002565.1, Cab004528.1, Cab003267.1, Cab015407.1, Cab040722.1, Bo3g032510.1, Bo4g046140.1, Bo5g002450.1, Bo5g010440.1, Bo9g174920.1) was associated with an increase in leaf B concentration, whereas increased expression for the remaining five GEMs (Cab002472.4, Cab031463.1, Cab021504.1, Bo2g028880.1, Bo7g014510.1) was associated with a decrease in leaf B concentration (all $p < 0.001$). GEM expression profiles were also compared between certain crop types, specifically winter (n = 160) and spring (n = 127) OSR, to assess whether these crop types exhibited any variation in expression patterns. The mean gene expression for all GEMs tested varied significantly between crop types (all $p < 0.001$ apart from Bo4g046140.1, where $p < 0.01$) (Table 3.5).

Table 3.5 Effect of transcript abundance on leaf B concentrations between B. napus crop types

GEMs tested as predictive for leaf B concentration and their A. thaliana orthologue information obtained from TAIR (Lamesch et al., 2012) are given. The correlation coefficient (r) and significance (p) of the correlation between transcript abundance (as reads per kb per million aligned reads; RPKM) and leaf B concentration is then given (n = 381). Next the mean gene expression of markers in winter (n = 160) and spring (n = 127) OSR accessions is shown, with the significance (p) of this variation also stated. The blue and pink colouration shows relative low and high expression respectively. Any significant values are in bold and underlined.

Chromosome	GEM	*A.* thaliana Annotation	383: Correlation Coefficient (r)	383: Significance (p)	Winter OSR: Mean Gene Expression	Spring OSR: Mean Gene Expression	Expression variation significance (p)
A1	Cab020459.1	BTB AND TAZ DOMAIN PROTEIN 5	0.269	**<0.001**	1.623	4.15	**<0.001**
A2	Cab032697.1	WITH NO LYSINE KINASE 4	0.297	**<0.001**	18.14	9.33	**<0.001**
A3	Cab002565.1	PROTEIN OF UNKNOWN FUNCTION	0.357	**<0.001**	3.882	17.122	**<0.001**
A3	Cab004528.1	No annotation	0.325	**<0.001**	0.339	2.528	**<0.001**
A3	Cab002472.4	FLOWERING LOCUS C	-0.337	**<0.001**	13.415	3.368	**<0.001**
A3	Cab003267.1	SOC1	0.262	**<0.001**	4.993	32.346	**<0.001**
A3	Cab015407.1	IQ-DOMAIN 2	0.306	**<0.001**	4.911	8.306	**<0.001**
A5	Cab031463.1	CYSTEINE PROTEINASES SUPERFAMILY PROTEIN	-0.315	**<0.001**	12.020	3.227	**<0.001**
A8	Cab040722.1	WRKY DNA-BINDING PROTEIN 32	0.271	**<0.001**	2.657	3.835	**<0.001**

Chromosome	GEM	*A.* thaliana Annotation	383: Correlation Coefficient (r)	383: Significance (p)	Winter OSR: Mean Gene Expression	Spring OSR: Mean Gene Expression	Expression variation significance (p)
A9	Cab021504.1	CDPK-RELATED KINASE	-0.276	**<0.001**	10.539	7.743	**<0.001**
C2	Bo2g028880.1	SHV3-LIKE 4	-0.281	**<0.001**	1.293	0.388	**<0.001**
C3	Bo3g032510.1	CASP-LIKE PROTEIN 4B1	0.306	**<0.001**	2.288	4.165	**<0.001**
C4	Bo4g046140.1	DHHC-TYPE ZINC FINGER FAMILY PROTEIN	0.237	**<0.001**	10.330	17.202	**<0.01**
C5	Bo5g002450.1	HEAVY METAL TRANSPORT/DETOXIFICATION SUPERFAMILY PROTEIN	0.297	**<0.001**	32.236	52.293	**<0.001**
C5	Bo5g010440.1	LEUCINE-RICH RECEPTOR-LIKE PROTEIN KINASE FAMILY PROTEIN	0.310	**<0.001**	1.722	2.793	**<0.001**
C7	Bo7g014510.1	AT RICH INTERACTION DOMAIN 2	-0.317	**<0.001**	5.583	1.779	**<0.001**
C9	Bo9g174920.1	ASTERIX-LIKE PROTEIN	0.340	**<0.001**	1.115	4.669	**<0.001**

Many of the GEMs were not well characterised. Cab002565.1, for example, the orthologue of which in *A. thaliana* is AT1G05870.4, is the most highly associated GEM from AT analyses and correlates significantly with leaf B concentration. However, no functional information or related publications are available on TAIR. The second most highly associated GEM is Bo9g174920.1, the orthologue of which in *A. thaliana* is described as an ASTERIX-LIKE PROTEIN (AT5G07960.1), similarly has no functional annotation. Bo9g174920.1 was also the 5th most highly associated marker within leaf Mn AT analyses. Cab004528.1 was the 7th most highly associated GEM with leaf B AT and was significantly predictive for leaf B concentrations. However, no *A. thaliana* annotation was present in the pan-transcriptome. The CDS for this gene model was compared against all others in the pan-transcriptome (see Section 2.2.3), however no gene models with *A. thaliana* orthologue annotations were found. Furthermore, comparing the CDS to other plant species by BLAST (Altschup et al., 1990) similarly did not yield any orthologous genes. Further characterisation of such unannotated genes (or genes lacking little functional information) would be of interest to assess their role in the control of leaf nutrient concentrations.

The third most highly associated GEM (out of a total of 53,889 CDS models for which significant expression was detected) from the 383 AT analyses was the CDS model Cab002472.4, the orthologue of which in *A. thaliana* is FLOWERING LOCUS C (*FLC*; AT5G10140.4). *AtFLC* is a regulator of flowering time, responsible for repressing floral induction prior to vernalisation (Sheldon et al., 2000). The expression of Cab002472.4 was significantly higher ($p < 0.001$) in winter than spring OSR (Table 3.5). This is unsurprising given the higher vernalisation requirements in winter OSR, with such expression profiles in concordance with an earlier study looking at *BnaFLC* (Tadege et al., 2001). Another homoeologue of *FLC* (CDS model Bo3g024250.1) was also found to be highly associated with leaf B concentration, being the 45th most associated GEM. *AtFLC* has previously been characterised on the PiiMS database; a significant decrease in Li, B, As, Se and Mo, and a significant increase in Mg, Ca, Fe and Ni concentrations was reported in the mutant line as compared to WT. A further flowering time regulator identified from GEM analyses was the CDS model Cab003267.1, the orthologue of which in *A. thaliana* is SUPPRESSOR OF OVEREXPRESSION OF CO 1 (*SOC1*; AT2G45660.1). *AtFLC* is a repressor of

AtSOC1 and therefore expression is repressed in plants exhibiting high FLC expression (Lee et al., 2000). Again, this follows observed expression patterns, with higher Cab003267.1 expression in spring OSR (where FLC expression was low) ($p < 0.001$). Both Cab007661.1 and Bo3g038880.1 were also two of the highly associated markers within leaf Mn analyses.

Of the remaining GEMs, certain candidates were of interest. One example is the CDS model is Cab020459.1, the orthologue of which in *A. thaliana* is described as BTB AND TAZ DOMAIN PROTEIN 5 (*BT5*; AT4G37610.1), a gene involved in plant development and is induced by auxin (Robert et al., 2009). A homoeologue of Cab020459.1 was also one of the most highly associated markers on chromosome C1 (CDS model Bo1g003510.1). Another is the CDS model Bo3g032510.1, the orthologue of which in *A. thaliana* is CASP-LIKE PROTEIN 4B1 (*CASPL4B1*; AT2G38480.1). CASP proteins are involved in the formation of the Casparian strip, with CASP-like proteins thought to be involved in membrane formation or cell wall modifications (Roppolo et al., 2014). This is an interesting candidate that could potentially perturb multiple elements, given that the Casparian strip is essential for multiple physiological processes, including the control of nutrient transport (Geldner, 2013). This GEM was also identified in AT analyses for essential elements by another member of the Bancroft research group and was therefore tested in *A. thaliana* using T-DNA insertion lines. However, no significant difference for leaf B concentrations was observed (Sweeney, 2018; PhD thesis).

The ionome of *A. thaliana* orthologues for five GEMs has previously been characterised on PiiMS. One of these was *AtFLC*, which as previously mentioned, had significantly decreased leaf B concentrations in the mutant line in comparison to WT. The remaining four candidates, namely; WRKY DNA-BINDING PROTEIN 32 (Cab040722.1/AT4G30935.1), CDPK-RELATEDKINASE (Cab021504.1/ AT3G50530.2), SHV3-LIKE 4 (Bo2g028880.1/AT5G58050.1) and LEUCINE-RICH RECEPTOR-LIKE PROTEIN KINASE (Bo5g010440.1/AT1G09970.1), were not found to be significantly different in the mutant lines on PiiMS. SHV3-LIKE 4 was also one of the most highly associated GEMs from leaf Mn AT, however again, no significant variation in leaf Mn concentrations was documented on PiiMS.

Two GEMs were taken forward for further study; Cab031463.1, the orthologue of which in *A. thaliana* is a putative cysteine proteinase (*CPS*; AT3G43960.1) and Bo5g002450.1, the orthologue of which in *A. thaliana* is a HEAVY METAL TRANSPORT/DETOXIFICATION SUPERFAMILY PROTEIN (*HMT*; AT1G01490.2). *AtCPS* has been linked to the control of root hair elongation, with mutants exhibiting shorter root hairs under P deficiency (Lin et al., 2011). B deficiency has also been shown to affect root hair formation and elongation (Martín-Rejano et al., 2011) and therefore this gene was an interesting candidate for further study, given that limited characterisation has previously been conducted. The second candidate is similarly not well characterised, with no publications specifically focusing on *AtHMT*. Whilst not the most highly associated marker, this GEM was the most significantly predictive candidate for leaf B concentration tested (Table 3.4). Furthermore, element transport has been linked to similar metal binding domain proteins (De Abreu-Neto et al., 2013). Resultantly, it is an interesting candidate for further characterisation.

3.2.3 Leaf B associative transcriptomics: candidates within linkage disequilibrium of highly associated SNP markers chosen for further characterisation in A. thaliana

Within association mapping studies, the most highly associated marker is not necessarily the causative locus due to non-random allele association (linkage disequilibrium, LD). It is feasible for any genes within LD of significantly associated SNPs to be causative of the observed associations. For information regarding analysis of SNP association peaks for candidates within LD, see Section 2.2.3.

Of the genes previously identified as controlling B concentrations in either *B. napus* or *A. thaliana* (including *NIP* or *BOR* orthologues), only one copy of *BnaNIP6;1* (CDS model Cab014832.1) was identified in leaf B AT analyses. This was within a minor peak observed on chromosome A2, with a lead marker (most highly associated SNP marker) of BnaA02g19760D:925:T (Figure 3.3). Due to the breadth of previous characterisation *NIP6;1* (*c.f.* Tanaka et al., 2008; Hua et al., 2017), no further research was warranted within this thesis' scope.

Through analysing additional association peaks overserved within leaf B AT, additional putative candidates were identified. Twelve candidate genes, orthologous to eight genes in *A. thaliana,* were chosen for further characterisation using T-DNA insertion lines. Whilst it is acknowledged that no previous work conducted on these genes specifically links them to a role in controlling leaf B concentrations, putative roles in root system architecture (RSA), cell wall organisation and the control of water transpiration make them interesting candidates for further characterisation. Information regarding these genes is provided hereafter and summarised in Table 3.6 (GEM candidates discussed previously are also included within this table). Additional candidates within LD of highly associated SNP markers, whilst not characterised using T-DNA insertion lines, that would be potential targets for future study, are discussed in Section 3.3.2 and summarised in Table 3.7.

Table 3.6 Summary of candidates from leaf AT analyses tested using A. thaliana T-DNA insertion lines

Markers identified from either SNP or GEM AT are first given, along with the A. thaliana orthologue and functional annotation taken from TAIR (Lamesch et al., 2012). Finally, the unique code for each SALK or SAIL T-DNA insertion line is stated.

SNP/GEM	Marker	*A. thaliana* Orthologue	Given Name	Annotated Function	SALK/SAIL Code
SNP	Cab021707.1 Cab006100.2 Bo2g161810.1	AT5G62000.3	*ARF2*	AUXIN RESPONSE FACTOR 2	SALK_035537
SNP	Cab042227.2	AT5G44480.1	*DUR*	DEFECTIVE UGE IN ROOT	SALK_143736C SALK_000741C
SNP GEM	Cab017365.1 Bo9g135000.1 Cab032697.1	AT5G58350.1	*WNK4*	WITH NO LYSINE (K) KINASE 4	SALK_201692C
SNP	Bo4g030820.1	AT2G37170.1	*PIP2;2*	PLASMA MEMBRANE INTRINSIC PROTEIN 2	SAIL_169_A03
SNP	Bo4g034640.1	AT2G36390.1	*BE3*	STARCH BRANCHING ENZYME 2.1	SALK_030954C
SNP	Bo8g066850.1	AT1G16880.1	*ACR11*	ACT DOMAIN REPEATS 11	SALK_206346C
SNP	Bo8g067520.1	AT1G17890.1	*GER2*	NAD(P)-BINDING ROSSMANN-FOLD SUPERFAMILY PROTEIN	SALK_020145C
SNP	Cab046445.1	AT3G28715.1	*ATP*	ATPASE, V0/A0 COMPLEX, SUBUNIT C/D	SALK_093787 SALK_078674C

SNP/GEM	Marker	*A. thaliana* Orthologue	Given Name	Annotated Function	SALK/SAIL Code
GEM	Cab031463.1	AT3G43960.1	*CPS*	CYSTEINE PROTEINASES SUPERFAMILY PROTEIN	SALK_152721
GEM	Bo5g002450.1	AT1G01490.2	*HMT*	HEAVY METAL TRANSPORT/DETOXIFICATION SUPERFAMILY	SALK_035720C SALK_060068C

Although not passing the Bonferonni or FDR corrected significance thresholds, discernible association peaks were seen on chromosome A2 (Figure 3.3) (for analysis of SNP association peaks with markers not passing significance thresholds see Section 2.2.3). The peak designated A2.1 contained the orthologue of *AtNIP6;1*, as discussed above. Within LD of Cab021724.1:1068:C in peak A2.2 is the CDS model Cab021707.1 the orthologue of which in *A. thaliana* is AUXIN RESPONSE FACTOR 2 (*ARF2*; AT5G62000.3). Orthologues of *AtARF2* were also found within two other association peaks; on chromosome C2 (CDS model Bo2g161810.1) and a minor peak on chromosome A6, designated as A6.3 (CDS model Cab006100.2) (Figure 3.4). Interestingly *AtARF2* is involved in the regulation of both lateral root (Marin et al., 2010) and root hair (Choi, Seo and Cho, 2018) formation. Some of the earliest symptoms association with B deficiencies are related to the plants' root system architecture (RSA) and the (Martín-Rejano et al., 2011), whilst root hair mutants have been shown to accumulate less B (Tanaka et al., 2014).

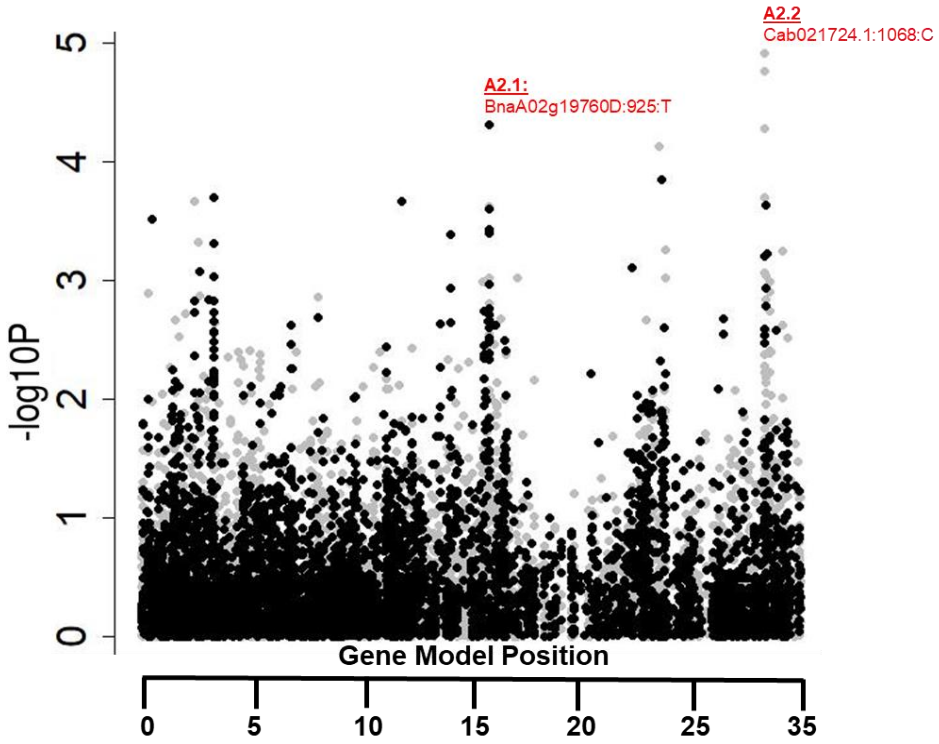


Figure 3.3 A2 leaf B SNP association peak

The –log10P value of leaf B SNP associations for chromosome A2 were plotted against the pseudomolecule database based on the CDS gene model order. These peaks have been designated as A2.1 and A2.2, with lead markers of BnaA02g19760D:925:T and Cab021724.1:1068:C respectively. The scale represents the approximate size of the chromosome and position of the marker in megabases.

Two association peaks on chromosome A6 contained SNPs passing the FDR, with lead markers of Cab006587.2:564:G and Cab042239.1:1176:G, designated as A6.1 and A6.2 respectively. Whilst the lead marker on A6.1 was significantly predictive for leaf B concentration (Table 3.3) no candidate genes were identified within this region. Within LD of the lead marker on peak A6.2 was the CDS model Cab042227.2, the orthologue of which in *A. thaliana* is DEFECTIVE UGE IN ROOT (*DUR*; AT5G44480.1). A paralogue of *AtDUR* is *AtMUR4* (At1G30620.2); an enzyme involved in the catalysis of UDP-D-Xylose to UDP-L-Arabinose. Previous characterisation of *Atmur4* has demonstrated that mutants exhibit a dwarfed phenotype, thought to be due to defected RG-II production (Burget et al., 2003). Furthermore, recent evidence has highlighted the importance of *AtDUR* in lateral root elongation (Zhao et al., 2019). This link to RG-II is of interest,

particularly given limited characterisation of this gene has previously been conducted.

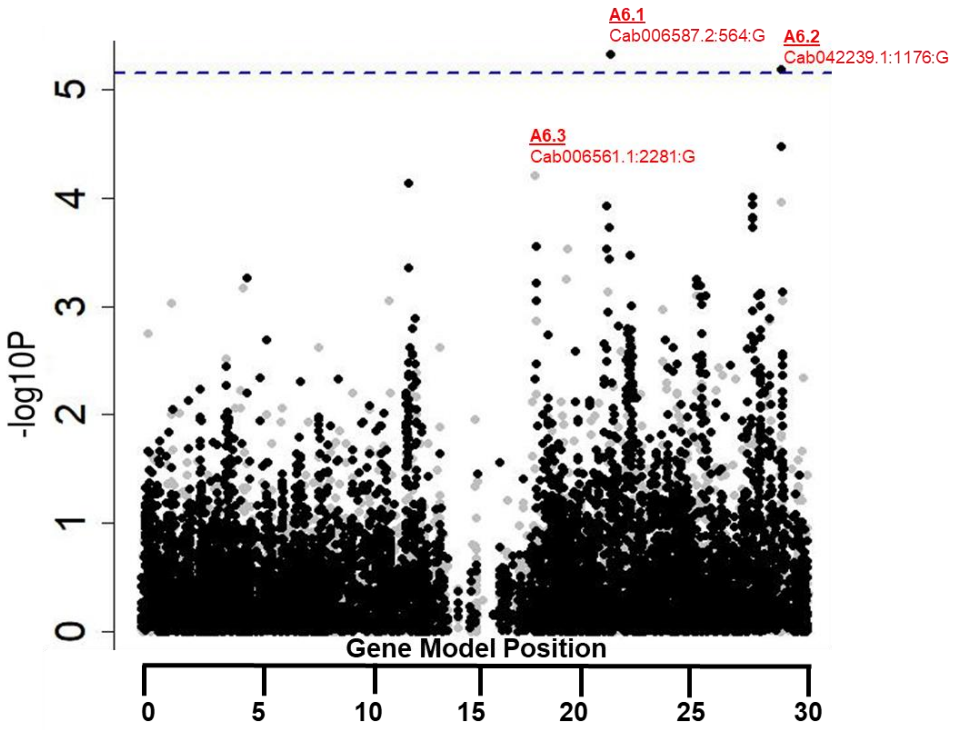


Figure 3.4 A6 leaf B SNP association peak

The –log10P value of leaf B SNP associations for chromosome A6 were plotted against the pseudomolecule database based on the CDS gene model order. These peaks have been designated as A6.1 and A6.2, with lead markers of Cab006587.2:564:G and Cab042239.1:1176:G respectively. An additional minor peak of interest has been characterised as A6.3, with a lead marker of Cab006077.1:299. The scale represents the approximate size of the chromosome and position of the marker in megabases.

The association peak designated as A10.2 co-localised with an association peak observed for leaf Mn concentrations, with lead markers of Cab017304.1:432:T and Cab017378.1:1044:T (B and Mn respectively). One candidate within this region is the CDS model Cab017365.1, the orthologue of which in *A. thaliana* is WITH NO LYSINE KINASE 4 (*WNK4*; AT5G58350.1). A homoeologue of Cab017365.1 was also found within the minor association peak found on chromosome C9 (CDS model Bo9g135000.1). Furthermore, an additional copy of *WNK4* was also found within the GEMs, as one of the more highly associated markers observed on chromosome A2. The transcript abundance for this GEM

(CDS model Cab032697.1) correlated significantly with leaf B concentration ($p < 0.001$). Gene expression also varied significantly ($p < 0.001$) between spring and winter OSR (Table 3.5). Although limited functional information is available for this candidates, given markers were observed in both SNP (in multiple association peaks) and GEM AT analyses it was an interesting candidate to test further.

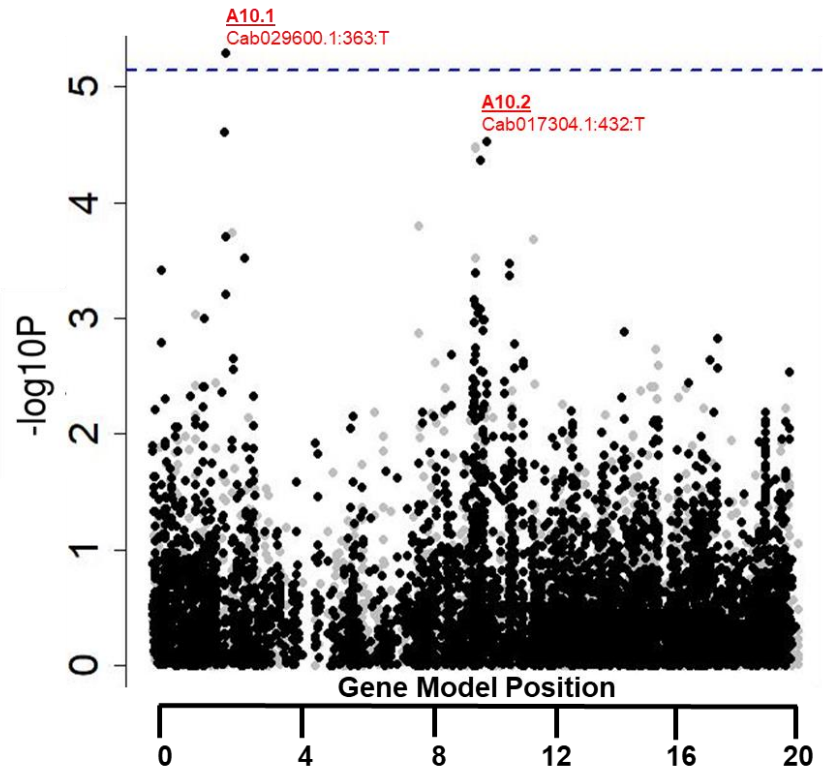


Figure 3.5 A10 leaf B SNP association peak

The –log10P value of leaf B SNP associations for chromosome A10 were plotted against the pseudomolecule database based on the CDS gene model order. These peaks have been designated as A10.1 and A10.2, with lead markers of Cab029600.1:363:T and Cab017304.1:432:T respectively. The scale represents the approximate size of the chromosome and position of the marker in megabases.

The lead SNP marker on chromosome C4, Bo4g026630.1:1541:T, was found to be significantly predictive of leaf B concentrations (Table 3.3). Two candidate genes within this region were chosen for further characterisation, both of which were also identified with seed B AT analyses. Firstly, the CDS model Bo4g030810.1, the orthologue of which in *A. thaliana* is PLASMA MEMBRANE

INTRINSIC PROTEIN 2 (*PIP2;2*; AT2G37170.1). *AtPIP2;2* is a member of the AQP family of proteins (which includes *AtNIP5;1*) and is essential for whole-plant water transport (Postaire et al., 2010; Da Ines et al., 2010). The upregulation of *BnaC04.PIP2;2* under B-deficiency has been reported (Hua et al., 2017; Table 1.1). A second candidate within this region is the CDS model Bo4g034640.1, the orthologue of which in *A. thaliana* is STARCH BRANCHING ENZYME 2.1 (*BE3*; AT2G36390). A SNP within the C4 association peak was also identified in AT outputs (Bo4g034640.1:2223:C) and subsequent predictive testing found this SNP to be negatively correlated with leaf B concentration ($R = -0.235$, $p = 0.014$).

The lead marker chromosome C8 was Bo8g066850.1:237:G and two candidates identified within this region were taken forward for further characterisation. Firstly, the CDS model Bo8g066850.1, the orthologue of which in *A. thaliana* is ACT DOMAIN REPEATS 11 (*ACR11*; AT1G16880.1). *AtACR11* is an activator of plastid glutamine synthetase (*GS2*), which is essential for N assimilation (Osanai et al., 2017). Whilst *ACR11* has not previously been linked to a control of B concentrations, *GS2* was proposed as a candidate gene in QTL analyses for shoot growth under low B (Shi et al., 2013; Table 1.3). A second candidate within this region is the CDS model Bo8g067520.1, the orthologue of which in *A. thaliana* is NAD(P)-BINDING ROSSMANN-FOLD SUPERFAMILY PROTEIN (*GER2*; AT1G17890.1), an enzyme putatively involved in GDP-fucose formation (Rhomberg et al., 2006). *AtMUR1* is also involved in this process (see Section 1.3.1.4), mutants of which have been characterised as fucose deficient with the phenotype restored with boric acid supplementation (O'Neill et al., 2001). Although *GER2* is not well characterised (with only one associated publication on TAIR), the link to fucose formation is of interest and therefore this candidate was also analysed further.

Whilst it is possible that these candidate genes are within the vicinity of an association peak by chance, all markers were within LD of highly associated SNP markers (following guidelines outlined in Section 2.2.3).

3.2.4 Leaf B candidate gene analysis: five candidate genes exhibit altered rosette leaf B concentrations

Of the *A. thaliana* T-DNA lines ordered for candidate analysis (Table 3.6), two lines failed to germinate (*DURa*; SALK_143736C & *CPS*; SALK_152721). Furthermore, upon genotyping two lines failed to yield any homozygous individuals (*ATPa*; SALK_093787 & *HMTa*; SALK_035720C). The remaining ten T-DNA lines were either segregating or confirmed homozygous. Unfortunately, due to time constraints, additional T-DNA lines were not obtained for *CPS* and therefore this candidate was not characterised further. Lines analysed via ICP-MS were compared individually with their respective Col-0 controls.

As the whole elemental profile was available via ICP-MS analyses, all essential elements whose percentage recovery was ≥ 85% were analysed. The following elements were included; B, Mg, P, S, K, Ca, Mn, Zn and Mo. All lines saw significant differences in elemental concentrations for at least one element as compared to WT. For ease of comparisons between multiple elements, the subsequent figures show the percentage difference from Col-0 (as outlined in Baxter, (2010)) for all elements studied. Specific statistical analyses (two-sided t-test with T-DNA line element vs WT control element) for all leaf ionome candidates is provided throughout Appendix III.

A significant variation in B concentrations was observed in five lines; a significant decrease in *arf2* (Figure 3.6), *wnk4* (Figure 3.7), *pip2;2* (Figure 3.8) and *acr11* (Figure 3.9), and a significant increase in *be3* (Figure 3.10). Whilst no significant variation in leaf B concentration was observed for the remaining four candidates (*dur*, *ger2*, *atp* & *hmt*), the elemental profiles are provided in Figure 3.11 - Figure 3.14 respectively.

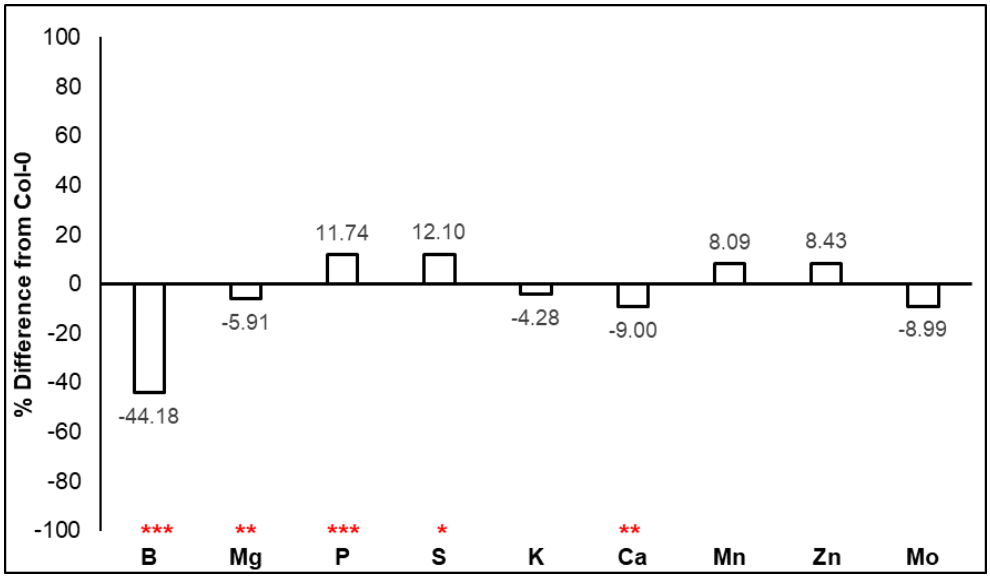


Figure 3.6 Percentage difference of elemental concentrations in leaf of arf2 compared to Col-0

Percentage difference of mean elemental concentrations (mg/kg) between control (Col-0) and arf2 T-DNA (SALK_035537) A. thaliana rosette leaves characterised via ICP-MS. Only essential elements whose percentage recovery was ≥ 85% were tested. Single, double and triple stars represent significance at p <0.05, p <0.01 and p <0.001 respectively, as determined via a two-sided t-test for each element individually. A summary of statistical data is provided in Appendix 3.1.

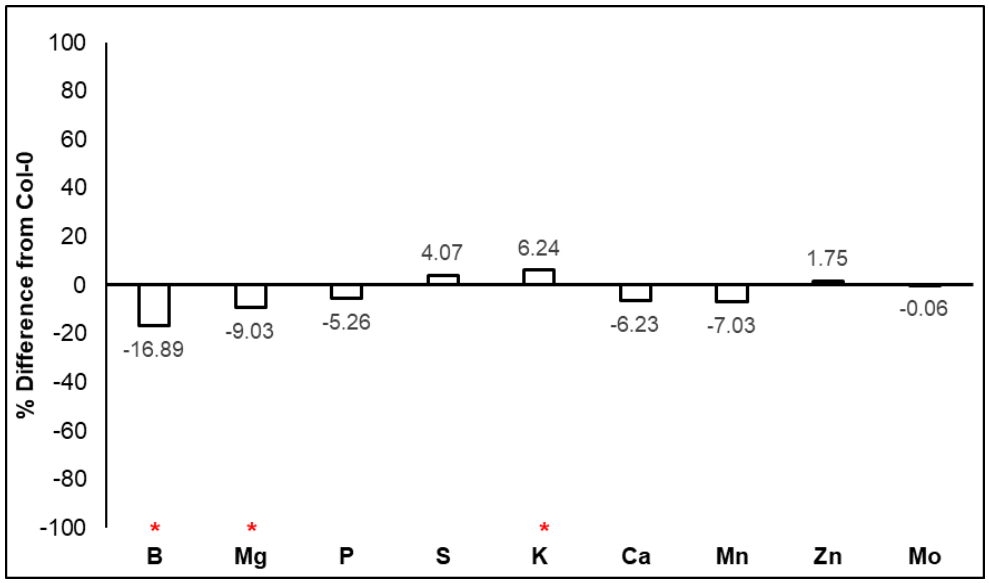


Figure 3.7 Percentage difference of elemental concentrations in leaf of wnk compared to Col-0

Percentage difference of mean elemental concentrations (mg/kg) between control (Col-0) and wnk T-DNA (SALK_201692C) A. thaliana rosette leaves characterised via ICP-MS. Only essential elements whose percentage recovery was ≥ 85% were tested. Single stars represent significance at p <0.05, as determined via a two-sided t-test for each element individually. A summary of statistical data is provided in Appendix 3.3.

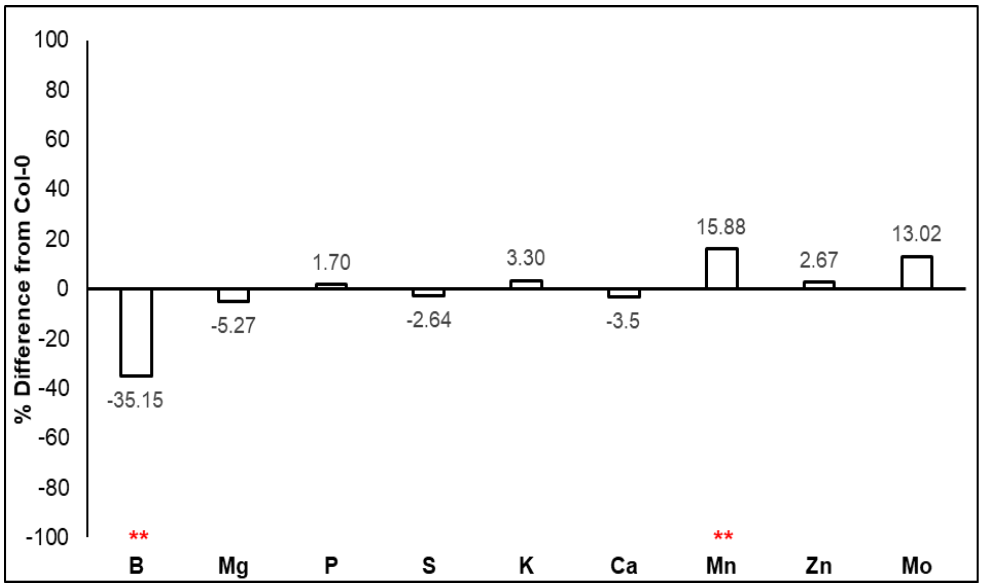


Figure 3.8 Percentage difference of elemental concentrations in leaf of pip2;2 compared to Col-0

Percentage difference of mean elemental concentrations (mg/kg) between control (Col-0) and pip2;2 T-DNA (SAIL_169_A03) A. thaliana rosette leaves characterised via ICP-MS. Only essential elements whose percentage recovery was ≥ 85% were tested. Double stars represent significance at p <0.05, as determined via a two-sided t-test for each element individually. A summary of statistical data is provided in Appendix 3.4.

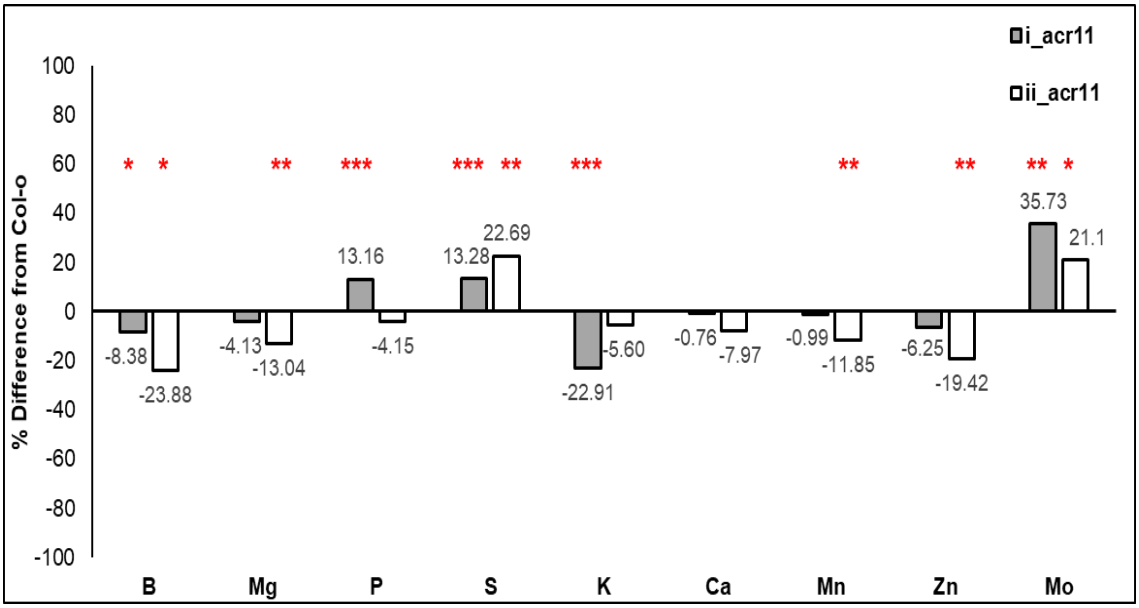


Figure 3.9 Percentage difference of elemental concentrations in leaf i_arc11 and II_acr11 compared to Col-0

Percentage difference of mean elemental concentrations (mg/kg) between control (Col-0) and acr11 T-DNA (SALK_296346C) A. thaliana rosette leaves characterised via ICP-MS. Data presented is from two separate experiments (8 biological replicates of A. thaliana rosette leaf material grown in two separate instances). Only essential elements whose percentage recovery was ≥ 85% were tested. Single, double and triple stars represent significance at p <0.05, p <0.01 and p <0.001 respectively, as determined via a two-sided t-test for each element individually. A summary of statistical data is provided in Appendix 3.6.

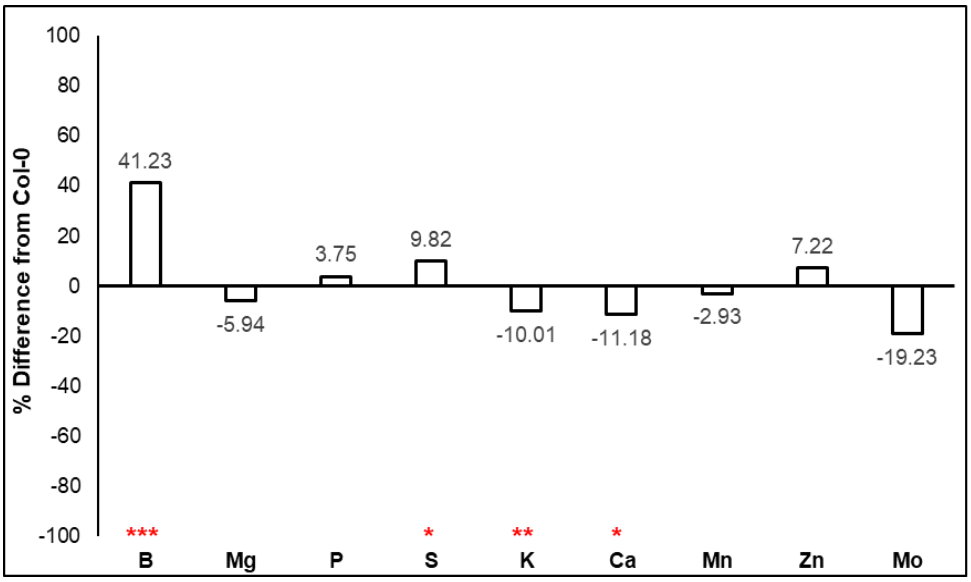


Figure 3.10 Percentage difference of elemental concentrations in leaf of be3 compared to Col-0

Percentage difference of mean elemental concentrations (mg/kg) between control (Col-0) and be3 T-DNA (SALK_030954C) A. thaliana rosette leaves characterised via ICP-MS. Only essential elements whose percentage recovery was ≥ 85% were tested. Single, double and triple stars represent significance at p <0.05, p <0.01 and p <0.001 respectively, as determined via a two-sided t-test for each element individually. A summary of statistical data is provided in Appendix 3.5.

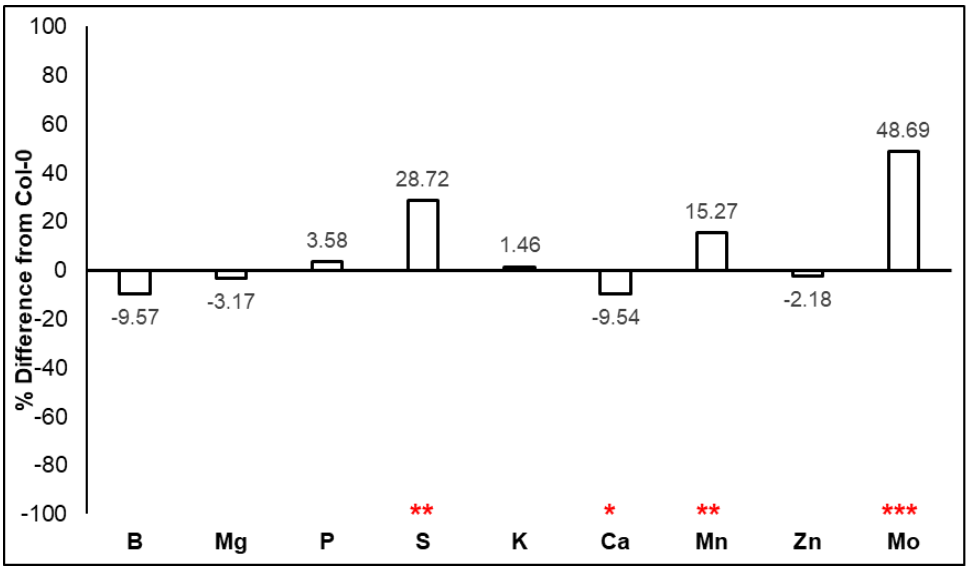


Figure 3.11 Percentage difference of elemental concentrations in leaf of dur compared to Col-0

Percentage difference of mean elemental concentrations (mg/kg) between control (Col-0) and dur T-DNA (SALK_000741C) A. thaliana rosette leaves characterised via ICP-MS. Only essential elements whose percentage recovery was ≥ 85% were tested. Single, double and triple stars represent significance at p <0.05, p <0.01 and p <0.001 respectively, as determined via a two-sided t-test for each element individually. A summary of statistical data is provided in Appendix 3.2.

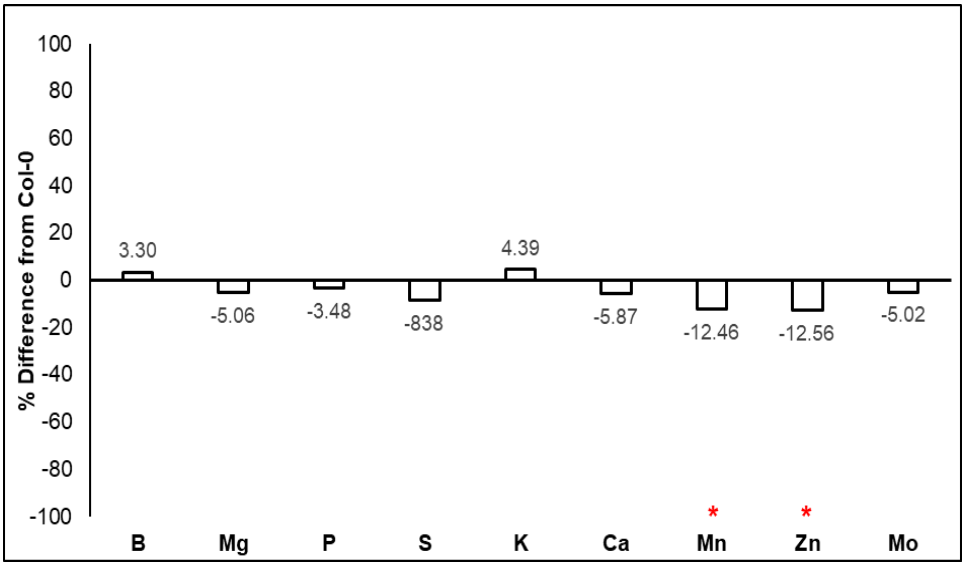


Figure 3.12 Percentage difference of elemental concentrations in leaf of ger2 compared to Col-0

Percentage difference of mean elemental concentrations (mg/kg) between control (Col-0) and ger2 T-DNA (SALK_020145C) A. thaliana rosette leaves characterised via ICP-MS. Only essential elements whose percentage recovery was ≥ 85% were tested. Single stars represent significance at p <0.05, as determined via a two-sided t-test for each element individually. A summary of statistical data is provided in Appendix 3.8.

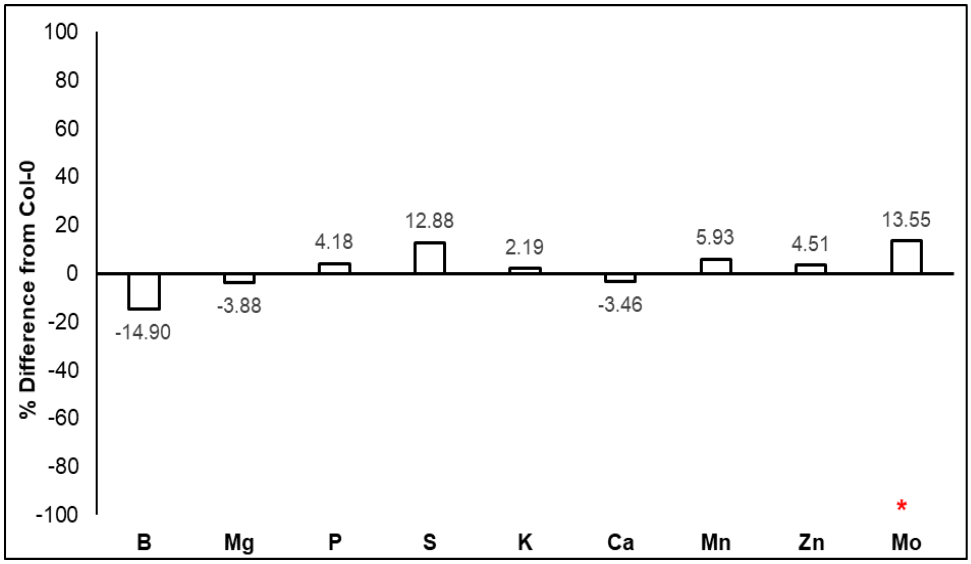


Figure 3.13 Percentage difference of elemental concentrations in leaf of atp compared to Col-0

Percentage difference of mean elemental concentrations (mg/kg) between control (Col-0) and atp T-DNA (SALK_078674C) A. thaliana rosette leaves characterised via ICP-MS. Only essential elements whose percentage recovery was ≥ 85% were tested. Single star represent significance at p <0.05, as determined via a two-sided t-test for each element individually. A summary of statistical data is provided in Appendix 3.7

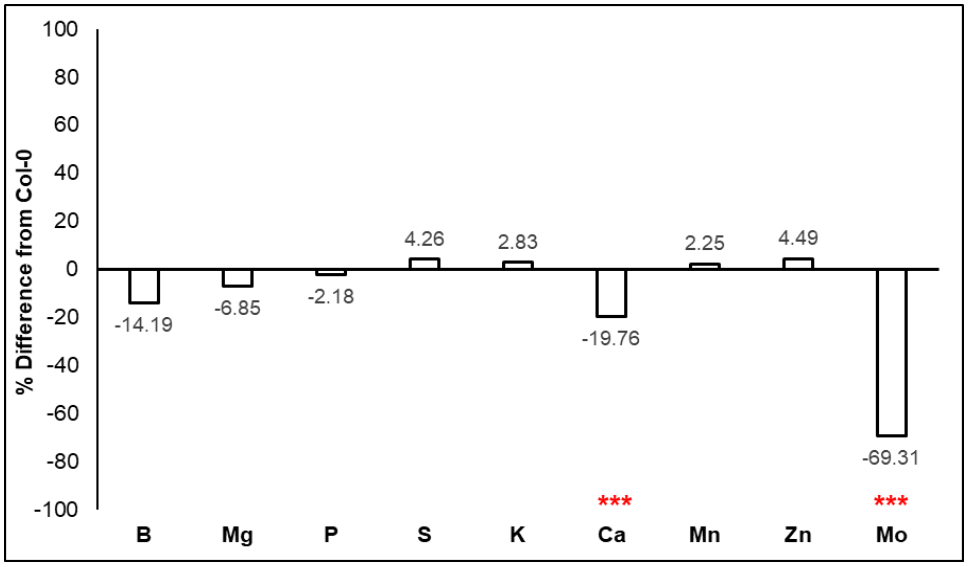


Figure 3.14 Percentage difference of elemental concentrations in leaf of hmt compared to Col-0

Percentage difference of mean elemental concentrations (mg/kg) between control (Col-0) and hmt T-DNA (SALK_060068C) A. thaliana rosette leaves characterised via ICP-MS. Only essential elements whose percentage recovery was ≥ 85% were tested. Triple stars represent significance at <0.001, as determined via a two-sided t-test for each element individually. A summary of statistical data is provided in Appendix 3.9.

3.3 Discussion

3.3.1 Analysis of leaf B concentrations in A. thaliana T-DNA insertion lines

3.3.1.1 Candidate gene analysis reveals A. thaliana mutants with altered rosette leaf B concentrations

In total 14 candidate genes, orthologous to 10 genes in *A. thaliana,* were identified from SNP and GEM AT analyses and further characterised using T-DNA insertion lines. Any perturbation within the mutant line as compared to WT validated the candidate within the scope of this thesis' research. Five T-DNA lines were found to vary significantly in leaf B concentrations. More information on candidates found to have significantly different leaf B concentrations as compared to WT has been provided hereafter. Furthermore, previously documented *A. thaliana* gene expression profiles for certain candidates have been summarised, as identified from Genevestigator (Hruz et al., 2008).

Leaf B concentration was significantly reduced in *arf2* (AT5G62000.3) ($p < 0.001$). This line exhibited the greatest percentage difference in B concentrations between mutant and control, with a mean decrease of 44.18% (Figure 3.6). Three orthologues of *AtARF2* were identified within AT analyses (within association peaks on chromosomes A2, A6 and C2). *ARF2* is an auxin response factor; transcription factors involved in the regulation of auxin responsive genes, with *AtARF2* acting as a transcriptional repressor (Guilfoyle and Hagen, 2007). Expression profiles obtained from Genevestigator reveal *AtARF2* is highly expressed across plant tissues and different developmental time-points (Figure 3.15).

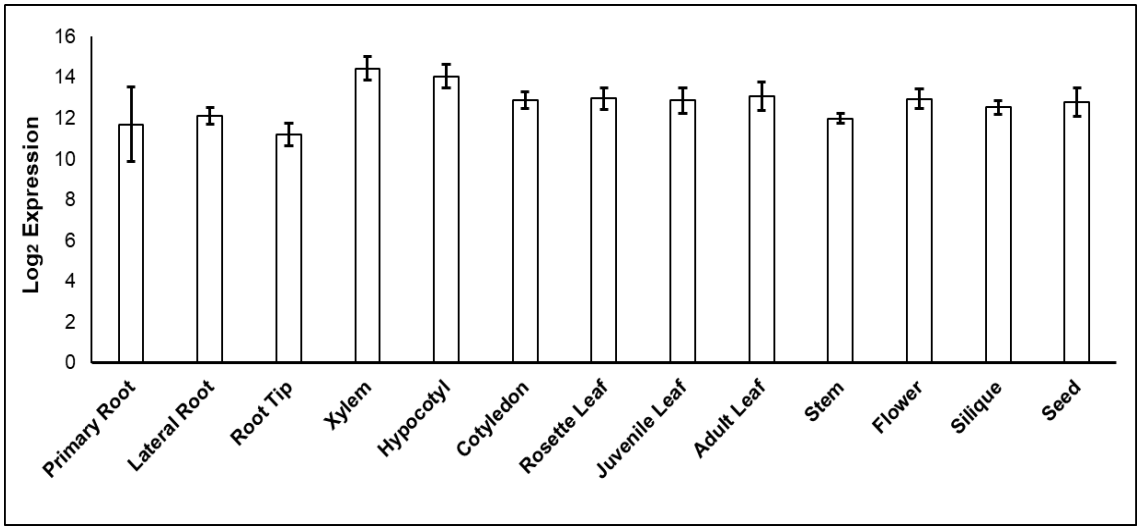


Figure 3.15 Expression of AtARF2 in different plant tissues

AtARF2 is expressed in various root, leaf and inflorescence tissues. Data was obtained from Genevestigator (Hruz et al., 2008), with data presented being the mean expression ± SD.

Previous characterisation of *Atarf2* mutants have shown multiple developmental phenotypes, including darkened rosette leaves, delayed flowering and abnormal floral morphology. These pleiotropic effects have not been directly attributed to auxin signalling, but rather through the regulation of downstream pathways involved in cell growth (Okushima et al., 2005). Interestingly, similar auxin-related developmental defects have been characterised for *Atrh7* mutants (Liu, Tabata and Imai, 2016) (DEAD-box RNA helicase predictive of leaf B concentrations; Table 3.3). Aside from the regulation of auxin-mediated development, there has also been a role identified within the control of K concentrations, whereby *AtARF2* acts as a transcriptional repressor of HIGH AFFINITY K^+ TRANSPORTER 5

(*HAK5*). Under conditions of K-sufficiency, *AtARF2* is bound to the *AtHAK5* promoter, with phosphorylation of *AtARF2* under deficient conditions relieving transcription inhibition. *Atarf2* mutants are more tolerant to low-K conditions and exhibit longer primary root length as compared to WT. Conversely, *ARF2*-overepressing lines are sensitive to low-K stress (Zhao et al., 2016). K concentrations were not perturbed within this study, which may be a result of the growing conditions (i.e. under nutrient sufficiency). Further characterisation has shown *AtARF2* is an inhibitor of root hair growth. For example, Choi, Seo and Cho, (2018) demonstrated *AtARF2* over-expressing lines exhibit shorter root hairs, whilst loss-of-function mutants exhibit longer root hairs. However it should be noted, Zhao et al., (2016) only observed differences in root hair morphology of *Atarf2* mutants under low-K conditions, and not under nutrient-sufficiency. Root hairs have previously been shown to affect B concentrations, for example root hair-less mutants (such as the NR23 mutant) accumulate less shoot B (Tanaka et al., 2014). Furthermore, White et al., (2013) proposes that increased number and longer root hairs can improve B acquisition. Therefore, this link to root hair growth within this mutant may be pertinent. However, if indeed *Atarf2* mutants do exhibit increased root hair length, then reasoning behind a decrease in B concentration as opposed to an increase, remains unclear. However, as root phenotyping was not conducted as part of this research, further characterisation of RSA (particularly assessing whether root hairs were perturbed in the *Atarf2* mutant) is required. As such, growing the T-DNA mutants using tissue culture methods (i.e. on agar plates) would be the next step for this research. Initially, nutrient sufficient conditions should be maintained to keep methods constant with how candidates were identified (i.e. diversity panel was grown under nutrient sufficiency and AT conducted to identify loci controlling B uptake in conditions of sufficiency). However, different nutritional stresses could be subsequently studied.

Multiple homoeologues of the kinase *WNK4* (*A. thaliana* orthologue AT5G58350.1) were identified in AT analyses. Two homoelogous copies were observed in SNP association peaks on chromosome A10 and C9. Furthermore a GEM (Cab032697.1) on chromosome A2 was positively correlated with B concentrations, with increased expression associated with increased leaf B (Table 3.5). As such, it was expected that a knock-out of this gene would lead to

decreased B concentrations in the mutant line. This was subsequently observed, with leaf B concentrations significantly reduced in the mutant as compared to WT ($p < 0.05$), with a mean reduction of 16.88% (Figure 3.7). The function of *AtWNK4* has not previously been reported, although it may be under the control of the circadian clock (Akamichi et al., 2002). Additional *WNK* family genes are involved in flowering time regulation via indirect regulation of *FLC* (Cao-Pham et al., 2018) and the phosphorylation of subunit C in the vacuolar H+-ATPase, which potentially mediates ion transport (Hong-Hermesdorf et al., 2006). However, specific phosphorylation targets of *AtWNK4* are unknown. As such, further characterisation is required to elucidate such targets, which may subsequently elucidate reasons why B concentrations were perturbed.

The aquaporin *Atpip2;2* (AT2G37170.1) exhibited significant differences in leaf B concentrations ($p < 0.01$), with a mean decrease of 35.14% in the mutant line (Figure 3.8). *PIP2;2* was arguably the most interesting candidate gene identified within leaf AT analyses that was subsequently tested using *A. thaliana* T-DNA insertion lines. This is primarily due to the fact that PIPs are AQPs, which as discussed previously, are a family of channel proteins which includes known B transporters (reviewed in Section 1.3.1.6). Furthermore, although PIPs have not been shown to be specific B transporters, yeast strains expressing two rice PIPs, *OsPIP1;3* and *OsPIP2;6*, exhibited functional influx and efflux B transport (Mosa et al., 2016). Similar characterisation has been conducted using the barley PIPs *HvPIP1;3* and *HvPIP1;4* (Fitzpatrick and Reid, 2009). Only Mn also differed significantly in the *pip2;2* mutant, with a significant increase in Mn concentrations ($p < 0.01$). Interestingly, Yao et al., (2014) found that AQPs, including PIPs, were downregulated in conditions of Mn toxicity. However no subsequent work has followed on from this.

BnaC04.PIP2;2 is upregulated under conditions of both B deficiency and toxicity (Hua et al., 2017) (Table 1.1), with similar expression profiles reported by Yuan et al., (2017). In *A. thaliana* the expression of *AtPIP2;2* is high across plant tissues and different developmental time-points (Figure 3.16).

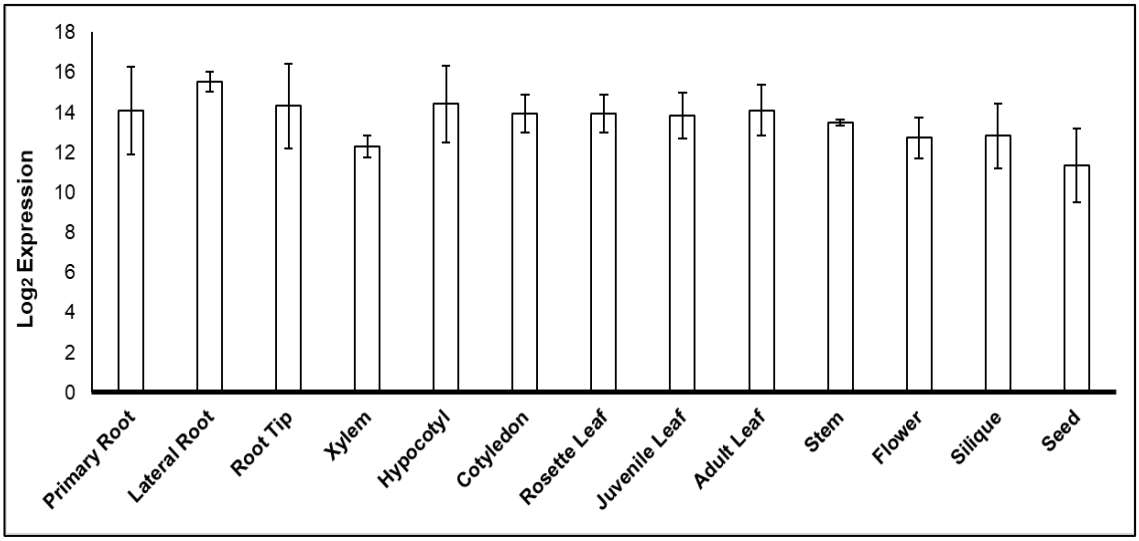


Figure 3.16 Expression of AtPIP2;2 in different plant tissues

AtPIP2;2 is expressed in various root, leaf and inflorescence tissues. Data was obtained from Genevestigator (Hruz et al., 2008), with data presented being the mean expression ± SD.

Previous research has primarily attributed the role of PIPs as facilitators of water transport; with *Atpip2;2* exhibiting significantly reduced relative water flux into rosette leaves (Da Ines et al., 2010; Postaire et al., 2010). This control of water transportation has also been linked to B accumulation, with repression of *AtPIP2;2* in both roots and shoots under B-toxicity proposed as a mechanism to control B concentrations (Macho-Rivero et al., 2018). However, irrespective that expression profiles have been studied for this gene, relative concentrations of B in different tissues of *Atpip2;2* mutants have, to the authors knowledge, not previously been characterised. Although there is no current evidence to support that *AtPIP2;2* is directly involved in B transport, at the very least a control of water balance via AQPs may be a mechanism to regulate the control of B accumulation. Clarification of whether *AtPIP*s and *BnaPIP*s are able to facilitate B transport would therefore be the next step within the research, which could initially be conducted using complemented yeast strains, similar to the aforementioned studies. Assessing the transport capacity of each of the seven *BnaPIP2;2* homoeologues may also be of importance, in case of function-specificity or sub-functionalisation of the different gene copies. If no transport capacity for *BnaPIP2;2*s were identified, the correlation between water flux and B, both within different tissues and under different nutritional statuses, may further clarify the

relationship between *BnaPIP*s and B concentrations (for example is greater B accumulation observed with increased water transport).

AtBE3 (AT2G36390.1) corresponded to a SNP within the C4 association peak (Bo4g034640.1:2223:C) with trait predictions finding this SNP to be negatively correlated with B concentrations. A significantly increased leaf B concentration was observed in the mutant line as compared to WT ($p < 0.001$), with a mean difference of 41.23% (Figure 3.10). *AtBE3* is a starch branching enzyme, with mutants exhibiting atypical starch phenotypes. Specifically, changes to amylopectin chain length distributions are observed (Dumez et al., 2006). B availability does affect starch accumulation, with increased starch content under B deficiency (*c.f.* Han et al., 2007; Bogiani, Amaro and Rosolem, 2013). It is postulated that this is owing to a reduced carbon demand due to growth inhibition under deficiency. However, as *AtBE3* does not affect overall starch content, rather amylopectin organisation, how this gene is specifically linked to B concentration is as yet unclear.

The final mutant found to have a significantly altered B phenotype was *AtACR11* (AT1G16880.1), with a significant decrease in leaf B concentrations ($p < 0.05$). This candidate was analysed twice (2 x 8 biological replicates), with the mutants found to have a mean decrease in leaf B concentrations of 8.37% and 23.87% as compared with WT (Figure 3.9). *AtACR11* is an activator of plastid glutamine synthetase (*GS2*), which catalyses the formation of glutamine from glutamate and ammonium, essential for N assimilation. Expression levels of *AtACR11* and *AtGS2* are highly correlated (Osanai et al., 2017). Various research has linked B availability to N metabolism, whilst increased expression of *GS2* has been identified as a response to short-term B deficiency in *N. tabacum* roots (Kobayashi, Mutoh and Matoh, 2004; Beato et al., 2014). Furthermore, *GS2* was a candidate loci underlying a QTL for *B. napus* shoot growth under low B (Shi et al., 2012) (Table 1.3), however the specific reasoning behind designating this as a candidate gene was unclear. *AtACR11* is also postulated to modulate reactive oxygen species (ROS) production, with *Atacr11* mutants accumulating ROS (Singh et al., 2018). ROS production is also an important mechanism in the control of RSA under B deficiency (Camacho-Cristobal et al., 2015) and may be involved in signalling nutrient deprivation (Schachtman and Shin, 2007). Further

characterisation is required to fully understand the role *ACR11* plays within the control of B concentrations; whether this is through the regulation of GS, ROS production, or some other unknown phenomenon.

3.3.1.2 Critical review of leaf mutant characterisation

This research presents a method for validating candidate genes putatively involved in leaf B accumulation, as identified through AT and subsequently tested in *A. thaliana* using T-DNA insertions lines. Characterisation of candidate genes identified through AT for leaf Ca and Mg accumulation was similarly conducted using T-DNA insertion lines by Alcock et al., (2017). Whilst this method was quick (rosette leaf harvesting was ~ 4 weeks post sowing); inexpensive (~ £6 per T-DNA line) and high-throughput (multiple T-DNA lines could be analysed on every ICP-MS run), certain factors require consideration.

Where available, confirmed homozygous T-DNA lines were ordered for elemental characterisation, however in some instances homozygous individuals were not identified through genotyping. Additional lines were segregating or all WT. Resultantly, only one T-DNA insertion line was used for ionomic analysis for each of the candidate genes tested in leaves. Whilst acceptable for initial screening of candidate function, additional T-DNA insertion lines should be analysed for further validation. Furthermore, transcript quantification of T-DNA lines, which was not analysed as part of this research, should be assessed. This would ensure gene knock-out within the mutant line (as opposed to knock-down). Additionally, if feasible, the use of segregating WT, rather than Col-0, should be used as a control for future research. This would allow for a more direct comparison between mutant and WT (for example in case the T-DNA line originated from varying Columbia ecotypes, or if there had been any contamination in the WT seed stock). However, as the majority of lines ordered for this research were not segregating, this would not have been possible for all T-DNA lines. Additionally, the lack of additional phenotyping of *A. thaliana* candidates (for example analysing RSA, rosette leaf morphology, days to flowering etc) is a more general criticism of this research. As such, further phenotyping of candidates, rather than just focusing on the ionomic profile, would be pertinent. For example, as previously discussed, analysing the RSA of *Atarf2* would enable further

clarification as to whether the root hairs were altered in this mutant, which could explain why B concentrations were perturbed.

All of the T-DNA lines tested showed ionomic variation in at least one element. Multi-element phenotypes are common in ionomic studies, for example Lahner et al., (2003) reported that only 11% of the mutants in their study exhibited variation in a single element. Within this research *Atacr11* exhibited the most perturbed ionome (Figure 3.9), with significant decreases in B, Mg, K, Mn, Zn, and significant increases in P, S and Mo concentrations. This candidate was tested twice with eight biological replicates in each instance. However, it should be noted that whilst the general trend within the elemental profile was consistent (i.e. elements were either increased or decreased), some elements did not differ significantly in both replications, highlighting the importance of repeating ionomic experiments. However this itself poses difficulties as the ionome may be impacted by multiple factors, such as growth stage, media, or environment (Huang and Salt, 2016). For example, hypothetically speaking, sampling plants for ionomic analysis on day 24 in experiment-X and experiment-Y would not necessarily result in the plant's exhibiting developmental homogeneity. Resultantly, this could impact on the ionomic profile. Therefore minimising experimental variation is essential. Growing plants in growth rooms (as opposed to glasshouses), as was conducted for this research, was decided upon in order to minimise environmental variation (e.g. light intensity and temperatures should resultantly be less variable). However, undoubtedly, growing conditions would have been inconsistent when growing plants for analysis in different batches. Using agar plates, as opposed to soil, would have accounted for differences in soil nutrient availability (and also have allowed root phenotyping and growing mutants under B deficiency and toxicity). However, this was decided against to maintain conditions in which plants were initially grown in the AT panel (i.e. in soil under nutrient sufficiency). Furthermore, this would not have been a viable option for analysis of seed AT candidates. Alternatively, hydroponics could have feasibly been utilised, although again this would have resulted in different growth methods than those used for the diversity panel. For example Liu et al., (2017) reported the difficulties of phenotyping in one growth system to subsequently predict those same phenotypes in alternative systems.

To conclude, whilst numerous criticisms of the experimental method can be put forward, utilising *A. thaliana* T-DNA insertion lines was a beneficial technique for an initial screen of leaf AT candidates. Five T-DNA insertion lines exhibited disruption in B concentrations relative to WT controls. The most interesting candidate which warrants further validation is undoubtedly *AtPIP2;2*. Further characterisation of this gene would be pertinent, particularly in clarifying whether it exhibits functional B transport capacities. Subsequent validation using knock-out mutants in *B. napus*, for example using TILLING lines (see Section 3.3.3), would also be of interest.

3.3.2 Candidates within LD of leaf B association peaks may be causative of variation in leaf B concentrations

As discussed, the remaining four AT candidates tested using *A. thaliana* T-DNA insertion lines (*dur*, *ger2*, *atp* & *hmt*) did not show a significant variation in leaf B concentrations (Figure 3.11- Figure 3.14). However, there were numerous additional candidates identified from AT which were not analysed in *A. thaliana*. Primarily these candidates were identified within LD of highly associated SNPs. Some of these candidate genes have previously been characterised, either as a known B transporter or on the PiiMS database, and therefore did not warrant further validation within the scope of this thesis. Others could be linked to B, or have the potential to affect the wider ionome. These candidates have been grouped and discussed hereafter and summarised in Table 3.7.

Table 3.7 Candidate genes within linkage disequilibrium of highly associated SNP markers from leaf B AT analyses

Candidates within LD of the most highly associated SNP makers within an association peak are summarised. The lead marker within the association peak and the candidate locus are given, as is the A. thaliana orthologue and annotated function on TAIR (Lamesch et al., 2012). If significnant variaition in leaf B concentrations between mutant and WT A. thaliana lines has been previoulsy reported on PiiMS (Baxter et al., 2007), this is also stated.

Association Peak	Lead SNP Marker	Candidate Locus	*A. thaliana* orthologue	Given Name	Annotated Function	PiiMS Leaf B variation
A2.1	BnaA02g19760D:925:T	Cab014832.1	AT1G80760.1	*NIP6;1*	NOD26-LIKE INTRINSIC PROTEIN 6;1	Yes
A2.2	Cab021724.1:1068:C	Cab021707.1	AT5G62000.3	*ARF2*	AUXIN RESPONSE FACTOR 2	-
A2.2	Cab021724.1:1068:C	Cab021721.1	AT5G61850.1	*LFY*	LEAFY	-
A6.2	Cab042239.1:1176:G	Cab042227.2	AT5G44480.1	*DUR*	DEFECTIVE UGE IN ROOT	-
A6.3	Cab006077.1:299	Cab006084.1	AT5G61850.1	*LFY*	LEAFY	
A6.3	Cab006077.1:299	Cab006100.2	AT5G62000.3	*ARF2*	AUXIN RESPONSE FACTOR 2	
A7	Cab012686.1:1527:A	Cab012603.1	AT3G58060.1	*MTP8*	MTP8	-
A7	Cab012686.1:1527:A	Cab012735.1	AT3G61250.1	*LMI2*	LATE MERISTEM IDENTITY 2	-
A9	Cab046445.1:2445:T	Cab046455.1	AT3G28860.1	*ABCB19*	ATP BINDING CASSETTE SUBFAMILY B19	Yes
A9	Cab046445.1:2445:T	Cab038423.1	AT5G25890.1	*IAA28*	INDOLE-3-ACETIC ACID INDUCIBLE 28	-
A10.1	Cab029600.1:363:T	Cab029648.2	AT1G05180.1	*AXR1*	AUXIN RESISTANT 1	Yes
A10.1	Cab029600.1:363:T	Cab036837.1	AT1G06180.1	*MYB13*	MYB DOMAIN PROTEIN 13	-
A10.2	Cab017304.1:432:T	Cab017371.3	AT5G58270.1	*ABCB25*	ABC TRANSPORTER OF THE MITOCHONDRION 3	Yes

Association Peak	Lead SNP Marker	Candidate Locus	*A. thaliana* orthologue	Given Name	Annotated Function	PiiMS Leaf B variation
A10.2	Cab017304.1:432:T	Cab017365.1	AT5G58350.1	*WNK4*	WITH NO LYSINE (K) KINASE 4	-
A10.2	Cab017304.1:432:T	Cab017166.1	AT5G60660.1	*PIP2;4*	PLASMA MEMBRANE INTRINSIC PROTEIN 2;4	-
C2	Bo2g163990.1:1902:T	Bo2g161810.1	AT5G62000.3	*ARF2*	AUXIN RESPONSE FACTOR 2	-
C2	Bo2g163990.1:1902:T	Bo2g161690.1	AT5G61850.1	*LFY*	LEAFY	-
C4	Bo4g026630.1:1541:T	Bo4g030810.1	AT2G37170.1	*PIP2;2*	PLASMA MEMBRANE INTRINSIC PROTEIN 2	-
C4	Bo4g026630.1:1541:T	Bo4g034640	AT2G36390	*BE3*	STARCH BRANCHING ENZYME 2.1	-
C6	Bo6rg076410.1:483:A	Bo6rg074430.1	AT3G58060.1	*MTP8*	MTP8	-
C6	Bo6rg076410.1:483:A	Bo6rg077600.1	AT3G61250.1	*LMI2*	LATE MERISTEM IDENTITY 2	-
C8	Bo8g066790.1:1116:T	Bo8g066850.1	AT1G16880.1	*ACR11*	ACT DOMAIN REPEATS 11	-
C8	Bo8g066790.1:1116:T	Bo8g067520.1	AT1G17890.1	*GER2*	NAD(P)-BINDING ROSSMANN-FOLD SUPERFAMILY PROTEIN	-
C8	Bo8g066790.1:1116:T	Bo8g065110.1	AT1G12950.1	*RHS2*	ROOT HAIR SPECIFIC 2	No
C8	Bo8g066790.1:1116:T	Bo8g064880.1	AT1G12560.1	*EXP7*	EXPANSIN A7	-
C9	Bo9g134980.1:180:C	Bo9g133340.1	AT5G57620.1	*MYB36*	MYB DOMAIN PROTEIN 36	Yes
C9	Bo9g134980.1:180:C	Bo9g135000.1	AT5G58350.1	*WNK4*	WITH NO LYSINE (K) KINASE 4	-
C9	Bo9g134980.1:180:C	Bo9g134880.1	AT5G58270.1	*ABCB25*	ABC TRANSPORTER OF THE MITOCHONDRION 3	Yes
C9	Bo9g134980.1:180:C	Bo9g134740.1	AT5G58050.1	*SVL4*	SHV3-like 4	No

3.3.2.1 NIP and PIP AQPs were found within LD of multiple association peaks

As reviewed in Section 1.3.1.6, aquaporins (AQPs) are a large protein family involved in water transport, gas-transfer, solute distribution and nutrient uptake (Tyerman, Niemietz and Bramley, 2002). Some AQP candidates were identified within AT analysed, as discussed hereafter. However, it should be noted, that this is a large protein family and there is the potential that an AQP candidate could be within the vicinity of an association peak by chance. In *B. napus* 121 full length AQPs have been characterised. This includes 43 PIPs, 35 TIPs (tonoplast intrinsic proteins), 32 NIPs and 11 SIPs (small basic intrinsic proteins) (Yuan et al., 2017).

An orthologue of one copy of NOD26-LIKE INTRINSIC PROTEIN 6;1 (*NIP6;1*; AT1G80760.1; CDS model Cab014832.1) was identified within LD of the lead marker in the minor association peak designated as A2.1. *AtNIP6;1* is a plasma membrane boric acid channel involved in xylem - phloem transfer and preferential distribution of B from roots to shoots, with transcript accumulation greatest in both young rosette leaves and shoot apices. Furthermore, although upregulated under B-deficiency, *AtNIP6;1* is also expressed under sufficient B conditions, particularly in stem tissue (Tanaka et al., 2008). This is perhaps unsurprising given the need for B transport to growing tissues, where B supply is essential (Dell and Huang, 1997). No other NIP, previously shown to be involved in B transport, including orthologues of *AtNIP5*;1 and *AtNIP7*;1 (Takano et al., 2006; Li et al., 2011), was observed in leaf B association analyses. This is in contrast to multiple previous QTL studies in *B. napus* which identified various orthologues of *AtNIP5;1,* as reviewed in Section 1.3.1.9. However, these studies were primarily focusing on B-efficiency traits (such as seed yield in plants grown in low B conditions), rather than the control of B concentrations in specific tissues. In *A. thaliana* the expression of *AtNIP5;1* and *AtNIP7;1* are primarily thought to be root (Takano et al., 2006) and anther (Li et al., 2011) specific respectively. Whilst significant expression of *AtNIP7;1* in inflorescence, shoot and root has been identified (Lindsay and Maathuis, 2016), such expression profiles have not been reported in *B. napus* (Hua et al., 2017). Furthermore, there was no significant expression (no RPKM > 0.4) for either copy of *BnaNIP7;1* within the leaf transcriptome data used within this research (transcriptome from the 2nd true leaf

stage, see Section 2.1.2). Conversely, low but significant expression (~ ≤ 1.50 RPKM) for three out of the total of six *BnaNIP5;1s* (*BnaA03.NIP5;1, BnaC02.NIP5;1* and *BnaC03.NIP5;1*) was detected. The limited significant expression of *BnaNIP5;1*s and *BnaNIP7;1*s may be a result of both the tissue sampled (young leaf only rather than root or inflorescence) and the methods in which the plants in the diversity panel were grown (i.e. nutrient sufficient conditions). In turn this could explain why these loci were not identified in AT analyses. This is a general criticism of the experimental methods used to detect loci controlling nutrient concentrations within this research. For further information regarding this, see Section 3.3.3.

Whilst no other NIPs were identified within AT analyses, several PIPs were found within LD of multiple association peaks. Although PIPs have not been shown to be specific B transporters, yeast strains expressing rice and barley PIPs exhibit functional influx and efflux B transport capacity (Fitzpatrick and Reid, 2009; Mosa et al., 2016). Interestingly, the expression of multiple *BnaPIPs* is induced under both B deficiency and toxicity (Table 1.1). Within the C4 association peak the CDS model Bo4g030810.1 was identified, the orthologue of which in *A. thaliana* is PLASMA MEMBRANE INTRINSIC PROTEIN 2 (*PIP2;2*; AT2G37170.1). This candidate was tested using *A. thaliana* T-DNA insertion lines; with the mutant exhibiting significantly decreased leaf B concentrations as compared to WT (Figure 3.8). Furthermore, within the association peak designated A10.2 was the CDS model Cab017166.1, the orthologue of which in *A. thaliana* is PLASMA MEMBRANE INTRINSIC PROTEIN 2;4 (*PIP2;4;* AT5G60660.1). In *B. napus, BnaPIP2;4*s are down-regulated under B deficiency (Yuan et al., 2017). Similar expression profiles have been reported in *A. thaliana* (Alexandersson et al., 2010) but interestingly not in *B. oleracea* (Diehn et al., 2015). *AtPIP2;4* also affects P availability, with mutants exhibiting longer root hairs under deficiency. Interestingly, control plants also exhibited longer root hairs, indicating a role within root hair development (Lin et al., 2011), however no subsequent work has followed on from this.

3.3.2.2 The regulation of flowering time may be an important indicator of leaf B concentrations

Several flowering time regulators were identified from AT analyses. These include *FLC* (CDS models Cab002472.4 and Bo3g024250.1) and *SOC1* (CDS model Cab003267.1) as three of the GEMs highly associated with leaf B concentrations (Table 3.4). *AtFLC* represses floral induction prior to vernalisation (Sheldon et al., 2000) and is also a repressor of *AtSOC1* (Lee et al., 2000), with both essential members of floral regulatory network. It is unlikely that either of these genes are specifically controlling B uptake. Given that B concentrations were significantly lower in winter than spring OSR ($p < 0.001$; Figure 3.1), the observed associations may be a consequence of underlying population structure within the diversity panel resulting in spurious associations between leaf nutrient concentrations and flowering time difference between crop types. However, the implementation of control methods within AT, specially the Q and K matrices (for both SNP and GEM analyses; see Section 1.3.4.2) should account for this. Conversely the different vernalisation requirements of crop types within the diversity panel may lead to expression patterns that may (directly or indirectly) be controlling leaf B concentrations. Additional flowering time regulators were identified within LD on multiple SNP association peaks. The CDS models Cab021721.1, Cab006084 and Bo2g161690, the orthologue of which in *A. thaliana* is LEAFY (*LFY*; AT5G61850.1) were identified within the association peaks A2.2 A6.3 and C2. *AtLFY* is a well characterised flowering time regulator involved in the early stages of flower initiation (Blázquez et al., 1997). Characterisation has similarly been performed in *B. juncea* (Dhakate et al., 2017). Furthermore, *AtLFY* is positively regulated by *AtSOC1* (Lee et al., 2008). A direct interactor of *AtLFY* is *AtLMI2* (Pastore et al., 2011), for which two homoelogous copies were found within LD of association peaks on chromosomes A7 and C6 (CDS models Cab012735.1 and Bo6rg077600.1). *At*LMI2 encodes a myb domain protein specifically involved in meristem transition (Pastore et al., 2011). Like *AtFLC* and *AtSOC1*, it is unlikely that these candidates are directly affecting leaf B concentrations. However, the control of vegetative growth and subsequent floral initiation by these regulatory genes may affect leaf B concentrations and would therefore be suitable markers for further characterisation. Additionally floral regulators, including orthologues of *AtFLC* and *AtSOC1*, have previously been

identified in AT analyses for leaf macronutrient (P, NO_3^-, Ca and Mg) concentrations (Alcock et al., 2017, 2018), indicating the importance of flowering time within the control of leaf nutrient accumulation.

3.3.2.3 Auxin-related and root hair development genes may affect leaf B acquisition

The relationship between auxin and B has primarily been discussed in relation to phytohormonal alterations under B deficiency. Primary root elongation, for example, is inhibited under low B supply, regulated in part through auxin signalling via *AtAUX1* and *AtPIN2* (Martín-Rejano et al., 2011; Camacho-Cristobal et al., 2015). Various genes related to auxin signalling were identified in AT analyses. Within the GEMs, the CDS models Cab020459.1 and Bo1g003510.1, the orthologue of which in *A. thaliana* is BTB AND TAZ DOMAIN PROTEIN 5 (*BT5*; AT4G37610.1), were associated markers on chromosomes A1 and C1. *AtBT5* is an auxin-inducible gene thought to be involved within plant development (Robert et al., 2009) although specific functions are unclear.

Several auxin-related genes within LD of multiple association peaks were also identified. Within LD of the lead marker on association peak A9 is the CDS model Cab038423.1, the orthologue of which in *A. thaliana* is INDOLE-3-ACETIC ACID INDUCIBLE 28 (*IAA28*; AT5G25890.1). *AtIAA28* is a regulator of lateral root formation in response to auxin signalling (Notaguchi, Wolf and Lucas, 20012). Also within this region was the CDS model Cab046455.1, the orthologue of which in *A. thaliana* is ATP-BINDING CASSETTE B19 (*ABCB19*; AT3G28860.1) a gene involved in polar auxin transport (Noh et al., 2001). Significant perturbations (both increased and decreased) in B concentrations in different *Atabcb19* T-DNA mutant lines was observed on PiiMS. Within LD of the lead marker on association peak A10.1 is the CDS model Cab029648.2, the orthologue of which in *A. thaliana* is AUXIN RESISTANT 1 (*AXR1*; AT1G05180). *AtAXR1* is auxin sensitive (Leyser et al., 1993), with root hairs formation affected within mutant lines (Cernac et al., 1997). A significant decrease in leaf B concentration of *Ataxr1* was observed on PiiMS. Several root specific candidates were also identified, particularly within LD of the lead marker of the C8 association peak. Root hair formation is induced under B deficiency (González-Fontes et al., 2016) with root

hair defective mutants in *A. thaliana* accumulating less shoot B (Tanaka et al., 2014). The CDS models Bo8g064880.1 and Bo8g065110.1, the orthologues of which in *A. thaliana* are EXPANSIN A7 (EXP7; AT1G12560.1) and ROOT HAIR SPECIFIC 2 (RHS2; AT1G12950.1) were identified within LD of the C8 association peak. Both *AtEXP7* and *AtRHS2* are involved in root hair elongation (Won et al., 2009; Lin, Choi and Cho, 2011). Additionally, the GEM Cab031463.1, the orthologue of which in *A. thaliana* is the putative cysteine proteinase (*CPS*; AT3G43960.1) has also been linked to the control of root hair elongation (Lin et al., 2011). Whilst further characterisation of these markers is required, the potential for root hair traits to modulate B acquisition would be an interesting area for further research.

3.3.2.4 Association peaks for leaf B and Mn concentrations co-localise on chromosomes A7, A10 and C9

Correlations between elements in plants has been discussed (Baxter, 2015), with various studies identifying multiple significant correlations between pairs or groups of elements (*c.f.* Broadley et al., 2008; Bus et al., 2014; Pinson et al., 2015). However, these elemental correlations can vary, and are often dependent upon numerous factors; including tissues, environments and populations utilised. QTLs for different elements have also been found to co-localise to the same genomic regions within these studies. Within the RIPR diversity panel, leaf B was most highly correlated with leaf Mn (Table 3.2). This in concordance with another study in *B. napus* (Bus et al., 2014), which also found a strong positive correlation between B/Mn concentrations in OSR seedlings. Interestingly, several association peaks co-localised to the same genomic regions for both leaf B and Mn concentrations (for leaf Mn Manhattan Plot see Appendix 1.3). The association peak on chromosome A7 was notably defined for both elements, with markers passing the Bonferroni corrected significance thresholds. The lead marker within both B and Mn SNP analysis was the same CDS model (Cab012686.1), although the specific polymorphisms did differ (B: Cab012686.1:527:A, Mn: Cab012686.1:1425:T). The CDS model Cab012603.1, the orthologue of which in *A. thaliana* is *MTP8* (AT3G58060.1), a cation efflux protein and a well characterised Mn transporter (Eroglu et al., 2016), was

identified within LD of the lead markers on chromosome A7. It is unclear if, or how, *MTP8* may affect B concentrations and unfortunately no ionomic information for *AtMTP8* is available on PiiMS. There is the possibility that the co-localisations of loci is occurring by chance, however this seems unlikely given the strong positive correlation between B and Mn and that SNP markers within the association peaks passed the Bonferroni corrected significance thresholds (hence should not be false positives). Further characterisation of *MTP8* would be pertinent. Initial screening using *A. thaliana* T-DNA insertion lines, as conducted within this research, would be a simple test to assess how B concentrations are affected.

The association peak for leaf B, designated as A10.2 (Figure 3.5), co-localised with an association peak observed for leaf Mn concentrations, with lead markers of Cab017304.1:432:T and Cab017378.1:1044:T (B and Mn respectively). Within this region was the CDS model Cab017365.1, the orthologue of which in *A. thaliana* is WITH NO LYSINE KINASE 4 (*WNK4*; AT5G58350.1). A homoeologue of Cab017365.1 was observed in the C9 association peak and additionally within GEM analyses on chromosome A2. *WNK4* was tested using *A. thaliana* T-DNA insertion lines. B concentrations were significantly reduced in the mutant line, however no significant perturbation in Mn concentrations was observed (Figure 3.7). Additional candidates within this region, which may be affecting both leaf B and Mn concentrations, were also identified. Within LD of both lead markers is the CDS model Cab017371.3, the orthologue of which in *A. thaliana* is ABC TRANSPORTER OF THE MITOCHONDRION 3 (*ABCB25*; AT5G58270.1). A significant decrease in both B and Mn concentrations was observed on PiiMS for *Atabcb25* mutants as compared to WT. Whilst no known role for ABC transporters within the control of B transport has been elucidated, one ABC transporter (*BnaC02.ABCG21*) was a candidate gene for the QTL underlining ‘B-efficiency coefficient’ (Hua et al., 2016b), as reviewed in Section 1.3.1.9.

A minor association peak on chromosome C9 was observed for both leaf B and Mn concentrations, with lead markers of Bo9g134980.1:180:C and Bo9g135080.1:1365:A respectively. This association peak contained the CDS models Bo9g133340.1 and Bo9g134880.1, the orthologues of which in *A. thaliana* are *WNK4* (AT5G58350.1) and *ABCB25* (AT5G58270.1), homoeologues

to the candidates discussed above. Another CDS model within this region is Bo9g133340.1, the orthologue of which in *A. thaliana* is MYB DOMAIN PROTEIN 36 (*MYB36*; AT5G57620.1). *AtMYB36* positively regulates the expression of various genes involved in Casparian strip formation, including CASPARIAN STRIP MEMBRANE DOMAIN PROTEIN 1 (*CASP1*; AT2G36100.1) and ENHANCED SUBERIN 1 (*ESB1*; AT2G28670.1). An orthologue of *AtCASP1* was identified within seed B SNP analyses (see seed chapter). *Atmyb36* mutants exhibit multi-element phenotypes, with a significantly altered leaf ionome (Kamiya et al., 2015). Characterisation on PiiMS shows a significant decrease in both B and Mn concentrations within mutant lines as compared to WT. Finally, within this region is the CDS model Bo9g134740.1, the orthologue of which in *A. thaliana* is SHV3-like 4 (*SVL4*; AT5G58050.1). A paralogue of Bo9g134740.1 was also found within the GEM analyses for both B and Mn concentrations (CDS model Bo2g028880.1). This GEM was found to be predictive and negatively correlated with leaf B concentration (Table 3.4). However, no significant ionomic disruption was characterised on PiiMS for either B or Mn concentrations. *AtSVL4* is a Glycerophosphodiester phosphodiesterase (GDPD) and whilst its role is not well understood, other family members are thought to be involved in the control of phosphate concentrations (Cheng et al., 2011). Additional shared GEMs were identified for B and Mn concentrations, such as the floral regulators *FLC* and *SOC1*.

3.3.3 Utilising AT and A. thaliana T-DNA lines to understand the genetic control of leaf B concentrations

Given that multiple QTL studies in *B. napus* identified orthologues of *AtNIP5;1,* as controlling phenotypic differences in B utilisation between accessions (Section 1.3.1.9 and Table 1.3), it was perhaps surprising that no markers were identified within AT analyses. Furthermore, whilst numerous association peaks were observed within AT analyses for leaf B concentrations, none of the other major loci previously linked to the control of B concentrations were identified (aside from one copy of *NIP6;1* in a minor association peak). This may be a result of various aspects relating to the initial experimental design. Transcriptome sequencing was only conducted on leaf material sampled at the 2nd true leaf stage and resultantly

may have not been representative of the leaves when they were sampled for ionomic analysis (at the 6th - 8th true leaf stage). Furthermore, genes which are only expressed in certain developmental stages (e.g. reproductive growth), tissues (e.g. roots) or environmental conditions (e.g. B deficiency) may not have been discovered. However, the successful implementation of AT for seed (Havlickova et al., 2018) and stem (Wood et al., 2017) traits have been conducted using this same data set (i.e. RNA sampled at 2nd true leaf). This is due to power of GWAS-based studies in identifying allelic variation in genes within LD of the causative gene, which results in the observed associated regions on Manhattan Plots (Havlickova et al., 2018). Nevertheless, it could be argued that having a more representative RNAseq dataset, from different developmental time points, would have been beneficial. Furthermore, subjecting the plants to nutrient deficiency and toxicity, rather than just sufficiency, may have yielded more candidates. It would also have been interesting to quantify B concentrations specifically within the cell wall, where the majority of cellular B is located (~ 95% under sufficiency; Blevins and Lukaszewski, (1998)). This may have given a greater indication of the scope of B variation/utilisation-efficiency across the panel. However, these benefits of having such data have to be offset by the greatly increased timescales that would have resulted in sampling material on multiple occasions (2160 plants were in the full panel, from which 387 accessions were genotyped/phenotyped in five replicates). Furthermore, as the RIPR collaboration assessed a wide range of phenotypic traits (and not just the ionome), this would have been a highly impractical task.

Regardless of *in silico* analysis of candidate genes identified from AT analyses, it is essential for validation to be conducted *in planta*. Whilst T-DNA insertion lines in *A. thaliana* were utilised within this research, various other techniques for validation could feasibly have been employed. Mutagenesis methods, such as TILLING (targeting induced local lesions in genomes), were initially developed as an alternative to insertional mutagenesis (Mccallum et al., 2000). Typically chemical mutagens (such as ethyl methanesulphonate; EMS) are used, which result in a high number of point mutations distributed randomly throughout the genome. TILLING technologies also produces a wider spectrum of mutations (such as missense and truncation) as opposed to insertional mutagenesis (which generally only results in gene knockout) (Kurowska and Daszkowska-golec,

2011). However recent advances in targeted genome editing make such techniques advantageous over traditional mutagenesis approaches. Older methodologies such as zinc finger nucleases (ZFNs; Kim, Cha and Chandrasegaran, (1996)) and transcription activator-like effector nucleases (TALENs; Christian et al., (2010)) have been increasingly replaced by RNA-guided nuclease systems (Bortesi and Fischer, 2015), with CRISPR-Cas9 (Jinek et al., 2012) arguably the most well-known. CRISPR-Cas9 has been used in *B. napus*, for example Braatz et al., (2017) was able to target two homoeologues genes involved in the pod shatter response. However, current European legislation designates that CRISPR-Cas9 edited organisms are genetically modified (GM) (Callaway, 2016), which therefore limits the commercial applications of material derived through this technique. RNA interference (RNAi; reviewed in (Wilson and Doudna, 2013)) can also be utilised for gene mutation, although this technique has variable success rates and is labour intensive (requiring vector construction and subsequent transformation) (Parry et al., 2009). Insertional mutation by endogenous transposons (*c.f.* Miura, Yonebayashi and Watanabe, 2001) can also be conducted, although this technique is not widely used and is not available for all plant species (e.g. is more suitable for maize due to the high transposon number) (Parry et al., 2009).

Whilst the majority of the aforementioned methods can be used for gene mutation in *B. napus*, the complexities of polyploidy mean that is it a technically difficult process. As previously discussed, up to six orthologues of each *A. thaliana* gene can be found in *B. napus* (Chalhoub et al., 2014), with homoeologues exhibiting minimal nucleotide divergence (Trick et al., 2009b). Furthermore, genetic redundancy within the *Brassica* genome (Lukens et al., 2004) makes the analysis of multiple gene copies intensive. As a result, insertional mutation (T-DNA) in *A. thaliana* (Krysan, Young and Sussman, 1999) was utilised within this research. However it should be noted that the successful implementation of T-DNAs for insertional mutagenesis is based on availability of efficient transformation techniques in the species of interest (Parry et al., 2009). In *A. thaliana* however, it is a widely established and utilised method.

The use of T-DNA lines to validate candidate genes in *A. thaliana*, as utilised within this research, was viable due to the extensive genome collinearity between

B. napus and *A. thaliana* (Cheung et al., 2009). However, given that the two species diverged approximately 20 mya (Koch, Haubold and Mitchell-Olds, 2001) gene function may not be conserved between species. Resultantly, candidates may have been incorrectly rejected (or indeed validated) using this technique. For a primary screen of gene function this was determined an acceptable compromise given the high number of candidates selected for characterisation and the much shorter time-frames for ionomic profiling (when using T-DNA in *A. thaliana*). However, further validation of gene function in *B. napus*, for example using TILLING, would be pertinent for all candidates exhibiting disruption in B concentrations identified from this thesis' research.

3.4 Summary and conclusions

The specific focus of the research presented in this chapter was assessing whether natural variation in leaf B concentrations, across the *B. napus* diversity panel, is a result of underlying genetic variation. SNP and GEM AT outputs were successfully analysed for candidate genes linked to the control of B concentrations in leaf.

Multiple SNP association peaks were observed within leaf B AT; the peak on chromosome A7 passed the Bonferroni corrected significance threshold, with additional peaks only passing the FRD on chromosomes A3, A6, A9, A10, C2, C4, C8. Furthermore, minor association peaks were observed in chromosomes A2 and C9 (Figure 3.2a). Although no GEMs passed the Bonferroni or FRD, some markers had notably higher –log10P values (Figure 3.2b), with expression for all tested GEMs correlating significantly with leaf B concentrations ($p < 0.001$) (Table 3.5). Furthermore, expression for all GEMs tested varied significantly between winter and spring OSR crop types (all $p < 0.001$ apart from Bo4g046140.1, where $p < 0.01$) (Table 3.5).

Lead SNP markers on peaks passing the FDR were tested for trait predictability, with four SNPs found to be predictive of leaf B concentration (Table 3.3). The effect of allelic variation was also tested, with B concentrations in accessions with the reference allele and the most frequent allelic variant compared. Up to a 1.80-fold change in mean B concentration was observed between the lead marker and

allelic variant, with a significant difference ($p < 0.001$) observed in all cases, except for the lead marker on chromosome A10 (in which significant B variation was not observed between alleles). In all cases, accessions with the allelic variant showed increased concentrations of leaf B (Table 3.3). This is indicative of the potential scope for improving leaf B accumulation, for example if accessions with these alleles were targeted in future breeding programmes. GEMs with notably higher –log10p values were also tested for trait predictability, with fifteen GEM markers predictive of leaf B concentrations (Table 3.4).

Five candidates identified in AT analyses and subsequently tested in *A. thaliana,* exhibited disruptions in leaf B concentrations relative to WT controls. *BnaPIP2;2* is a particularly interesting candidate due to the past research which links PIP AQPs to both B transport and a control of B accumulation via water flux (Macho-Rivero et al., 2018). Whilst further characterisation is required, this is a promising candidate which warrants additional research. Furthermore, the other four candidates which exhibited perturbed B concentrations within leaves would benefit from further investigation. Initially, characterising additional T-DNA insertion lines should be conducted before further phenotyping (e.g. growing mutants under differing B availabilities) and subsequent validation in *B. napus* (e.g. through TILLING). Additional candidates were also identified from GEM analyses and within LD of highly associated SNP markers. Some of these candidates were previously characterised on PiiMS and therefore were not tested further within the scope of this research. For example, an orthologue of MYB DOMAIN PROTEIN 36 was found within LD of the lead marker on chromosome C9, whilst an orthologue of NOD26-LIKE INTRINSIC PROTEIN 6;1 was found within LD of the lead marker on chromosome A2. Markers identified within leaf AT could ultimately be useful for marker-assisted selection (MAS) strategies to improve B-usage or accumulation in these tissues.

Chapter 4: Identifying genetic loci controlling seed boron concentrations in *B. napus*

4.1 Introduction

B availability during reproductive growth can significantly affect seed development, with B deficiency known to inhibit anther formation, pollen viability and the growth of the pollen tube (Shorrocks, 1997). For example, Marschner, (1995) reports that B availability during pollen tube growth is a major factor responsible for the high B demand during seed production. The inflorescence act as a B-sink, with Pommerrenig et al., (2018) demonstrating *B. napus* accessions deemed more B-efficient (estimated through various shoot and root growth parameters) were able to allocate B more effectively into the inflorescence than inefficient accessions.

Subsequent B concentrations within seeds are critical for seedling vigour and promoting root development during germination (Eggert and von Wiren, 2013). The release of B from the seed coat (where ~ 55% of B is found within the seed) during germination is reportedly low, and resultantly seed B nutrition is a major factor inhibiting seedling establishment (Eggert and von Wiren, 2013; Eggert and von Wirén, 2016). This is particularly relevant for growth in conditions of B-deficiency. A study in soybean, for example, looked at seedling establishment from seeds containing either 10 mg/kg or 20 mg/kg B. In B-sufficiency no significant differences in seedling growth was observed, however during deficiency, 80% of seedlings containing 10 mg/kg were abnormal, as compared to 20% from seeds containing 20 mg/kg (Rerkasem et al., 1997). Bellaloui, Smith and Mengistu, (2017) also identified strong positive correlations between seed coat B and subsequent successful germination.

The particular species under investigation can impact seed B concentrations, with various figures reported. Many studies have also focused on different environmental or developmental factors impacting upon concentrations. Mehmood et al., (2018), for example, looked at the effect of plant density and foliar B application on the resulting seed B concentration in cotton (*Gossypium hirsutum*). Higher plant densities (88,888 plants per hectare $(ha)^{-1}$, as compared

to 53,333 plants per ha^{-1}) significantly impacted seed B content, reducing average concentrations from 48 mg/kg to 38 mg/kg. Conversely, applying foliar B (at a concentration of 1200 mg/L) increased average seed B from 37 mg/kg to 48.3 mg/kg. Foliar B application was also found to significantly increase B concentrations in winter OSR from 8.6 mg/kg to 9.4 mg/kg (Jankowski et al., 2016). Whilst Eggert and von Wirén, (2016) found seed B concentrations in spring OSR to be 5 - 15 mg/kg. Harvest date can impact both total seed B and seed coat B, with delayed harvest decreasing average concentrations in soybean from ~ 44 mg/kg to ~ 29 mg/kg (Bellaloui, Smith and Mengistu, 2017). Undoubtedly however, plant accession can have a large effect on seed B, although few diversity panels (specifically *Brassica* species) focusing on seed B traits have been published. Baxter et al., (2012) reported seed B concentrations in 96 *A. thaliana* accessions, where seed B ranged from 8.98 - 40.59 mg/kg. Furthermore, heritability was estimated to be 0.39. Within the RIPR diversity panel seed B varied 3.38-fold (Figure 4.1), although estimated heritability was lower, at 0.13 (Thomas et al., 2016).

4.2 Results

4.2.1 Seed B concentration variation and elemental correlations across the RIPR diversity panel

Variation in seed B across the RIPR diversity panel varied 3.38-fold, from 26.15 to 88.46 mg/kg, with the means for each crop types as follows; winter fodder 49.25, semiwinter OSR 50.69, swede 52.20, winter OSR 53.16, spring OSR 56.56. Spring OSR had the highest mean seed B concentrations, with accessions exhibiting significantly higher mean seed B concentrations than winter fodder ($p < 0.001$). No significant difference was observed between other crop types. There was no significant correlation between seed B and leaf B concentrations ($p = 0.089$).

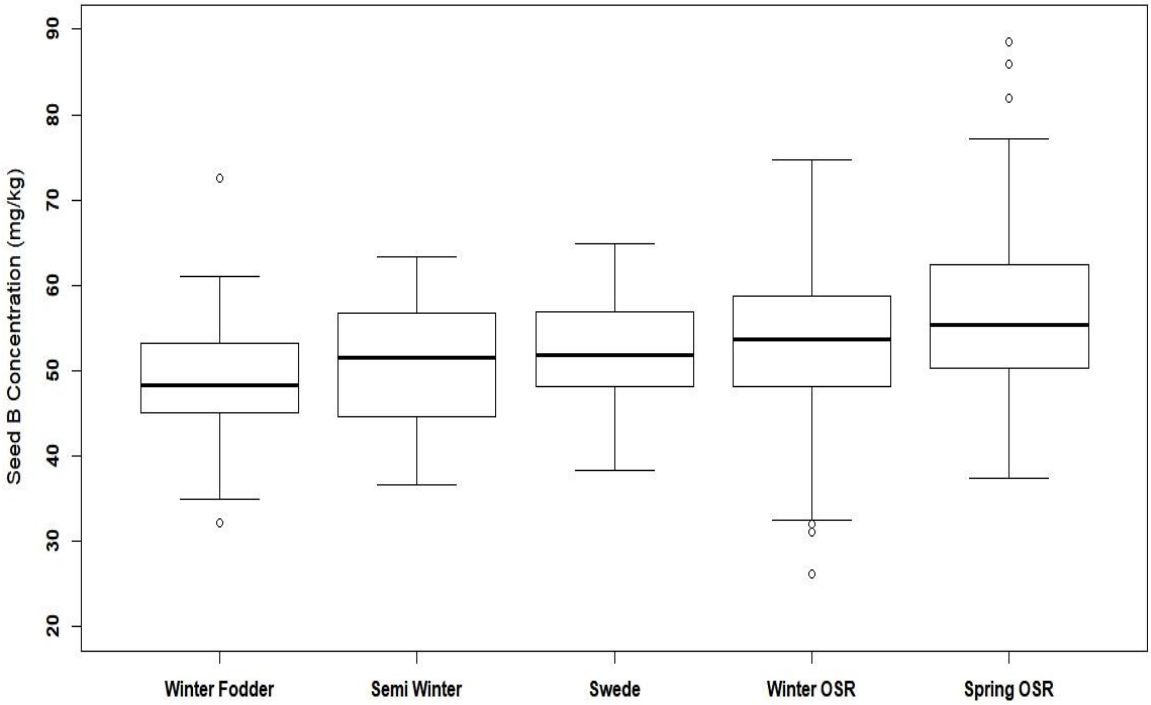


Figure 4.1 Seed B (mg/kg) concentration across crop types

Data presented shows the means of seed B concentrations of 338 B. napus accessions within the RIPR diversity panel whose crop type had been assigned, including; 15 winter fodder, 8 semiwinter OSR, 30 swede, 159 winter OSR, 126 spring OSR. The whiskers represent the lowest and highest data points > 1.5 times the upper and lower quartiles. Circles represent outliers < 1.5 times the upper and lower quartiles. Boxes are organised by mean seed B from lowest (left) to highest (right).

B concentrations in seed were significantly negatively correlated with eight of the eleven elements tested. Additionally, B was positively correlated with Mo (Table 4.1). Within this panel, B is most highly correlated with S, with this found to be the second most correlated pair of elements within the RIPR diversity panel, second only to S/Mo (Thomas et al., 2016). Association peaks on chromosomes A2, C2, A9 and C9 observed for seed S were found to co-localise to the same region as those seen for seed B (see Appendix II for seed S Manhattan plots). Given this association, a further detailed discussion of this can be seen within Section 4.3.2.2. Seed Mn, although not as highly correlated as in leaf (Table 3.2), was once again analysed, with Manhattan Plots for seed Mn compared with seed B. Again, SNP association peaks were found to co-localise, with a particularly defined peak on chromosome A5 observed (see Appendix II for seed Mn

Manhattan plots). Further discussion of potential candidates within this region are discussed throughout this chapter.

Table 4.1 Correlations between seed B and other essential elements

Correlation coefficient for seed B with eleven essential nutrients (n = 380). Significant values are in bold and underlined.

Element Vs B	Correlation Coefficient (r)	Significance (p)
Ca	-0.263	**<0.001**
Cu	-0.148	**0.004**
Fe	-0.118	**0.022**
K	-0.083	0.107
Mg	-0.095	0.064
Mn	-0.178	**<0.001**
Mo	0.241	**<0.001**
Ni	-0.152	**0.003**
P	-0.201	**<0.001**
S	-0.435	**<0.001**
Zn	-0.423	**<0.001**

4.2.2 Seed B associative transcriptomics: outputs and predictions

Association analyses for seed B concentrations across the 383 diversity panel is shown in Figure 4.2. No SNP or GEMs passed the Bonferroni or FDR corrections. However, discernible SNP association peaks were seen on chromosome A2, A5, A9, C2, C4 and C9. Association peaks within the GEMs were seen also on chromosome A9, C2 C9, with a minor discernible peak on A2. Additional GEMs with notably higher –log10P values were also observed. Again, predictions using the most highly associated SNP and GEM markers were carried out using split training and test panels, containing 274 and 109 accessions respectively. Trait

predictions with SNP and GEM markers can be seen in Table 4.2 and Table 4.3 respectively, with Manhattan Plots for AT analyses on the 274 panel found in Appendix II.

The association peaks seen within both SNP and GEM AT on chromosomes A2, A9, C2 and C9 correspond to regions of genome deletion containing orthologues of *AtHAG1* (At5g61420), a regulator of aliphatic glucosinolate (GSL) production (Hirai et al., 2007). The relationship between seed B and GSL was not investigated as part of this research, however further discussion of this is provided in Section 4.3.2.2. As such, this chapter will focus on additional candidates, in particular those identified within LD of association peaks observed on chromosomes A5 and C4.

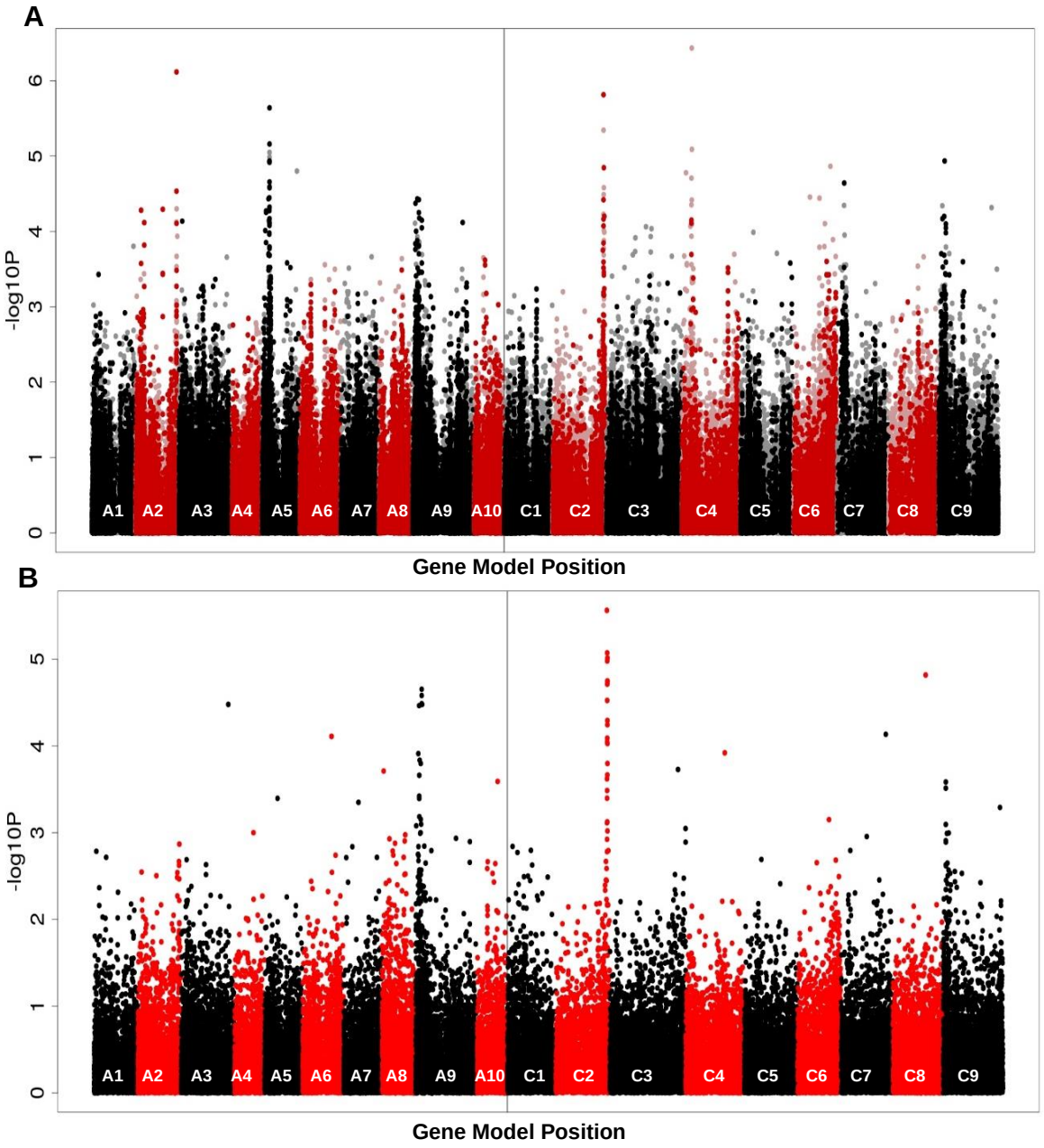


Figure 4.2 Genome wide distribution of (a) SNP and (b) GEM markers associating with B concentrations (mg/kg) in seeds from 383 accessions

SNP associations were calculated with the R script GAPIT (Lipka et al., 2012) using compressed mixed linear modelling with inference of population structure by Q-matrix produced using PSIKO (Popescu et al., 2014). GEM associations were performed using Regress, performing fixed-effect linear modelling, using PKISO-derived Q-matrix and RPKM values as explanatory variables, and seed B as the response variable. The –log10P value of SNP and GEM associations were plotted against the pseudomolecule database (representative of the nineteen B. napus chromosomes; labelled A1-A10 and C1-C9) based of the CDS gene model order. In the SNP analyses, light red and grey points represent hemi-SNP markers that have not been linkage mapped to a genome, whereas black and red markers represent markers assigned with confidence to a genomic position.

Table 4.2 Predictive capability of single nucleotide polymorphism markers from seed B 274 AT analysis

The most highly associated SNP markers within association peaks were tested for their predictive capability of seed B concentration. The SNP marker and the –log10P from the 274 AT analysis are given. Next the mean seed B concentration (mg/kg) for accessions in the panel with the reference allele and the most frequent allele variant are shown, with the significance (p) in the difference of these concentrations. The blue and pink colouration shows relative lower and high B concentration respectively. Finally, the correlation coefficient (r) and significance (p) of the predictions using the 109 panel are shown. Significant values are shown in bold and underlined.

Association Peak	SNP	AT 274 -log10P	B (mg/kg): Reference Allele	B (mg/kg): Allelic Variant	B (mg/kg): Allelic Variation Significance (p)	109 Prediction: Correlation Coefficient (r)	109 Prediction: Significance (p)
A2	Cab021724.1:682:G	6.44	55.95	48.21	**<0.001**	0.232	**0.016**
A5	Cab046472.1:561:A	5.15	52.51	55.00	0.065	0.106	0.276
A9	Cab044629.1:1830:G	4.36	56.40	47.42	**<0.001**	0.148	0.127
C2	Bo2g161640.1:589:C	5.54	55.99	50.30	**<0.001**	0.221	**0.022**
C4	Bo4g034470.1:987:C	5.23	53.71	53.04	0.625	0.095	0.330
C9	Bo9g017380.1:255:A	4.08	49.63	56.54	**<0.001**	0.162	0.094

Two SNP makers (those on association peaks A2 and C2) of the six lead markers analysed were found to be predictive of seed B phenotype (Table 4.2). The orthologue in *A. thaliana* for both Cab021724.1:682:G and Bo2g161640.1:589:C is CALNEXIN 1 (*CNX1*; AT5G61790.1). Both of these markers were closely located to *HAG1* homoeologues, being within five CDS models. *AtCNX1* is localised to the endoplasmic reticulum and is involved in protein folding (Liu et al., 2015), with pollen viability and tube growth particularly affected in *Atcnx1* (Vu et al., 2016). The ionome of leaf material in *Atcnx1* has previously been characterised on PiiMS, however there was no significant difference in B concentrations between mutant and WT lines.

Neither of the two most highly associated GEMs (that were not located with the A2/C2 and A9/C9 association peaks) (CDS models Bo8g091260.1 and Bo3g164310.1) had an annotated *A. thaliana* orthologue on the pseudomolecule database. The CDS of these gene models was compared against all others in the pan-transcriptome (see Section 2.2.3). No gene models with *A. thaliana* orthologue annotations was found for Bo8g091260.1. However, Bo3g164310.1 was most orthologous to AT4G12560.1 (CONSTITUTIVE EXPRESSER OF PR GENES 1; *CPR1*), a gene involved in the pathogen defence response (Gou et al., 2009). *Atcpr1* has previously been characterised on PiiMS, however no significant variation in leaf B concentrations was observed (however this would not necessarily mean that seed B concentrations would not be perturbed). However, neither of these two GEMs was predictive for seed B concentrations (Table 4.3) and therefore no further analysis was conducted.

Of the additional GEMs, five were found to be predictive of seed B concentrations, including Cab038298.3 (*AtHAG1*) as previously discussed. However, none of these genes had a well elucidated function. Cab026097.1, the orthologue of which in *A. thaliana* is PATHOGENESIS-RELATED THAUMATIN SUPERFAMILY PROTEIN (*PRT*; AT2G28790.1) is a cell wall protein identified from extracts in elongating cells (Bayer et al., 2006; Irshad et al., 2008). Cab008178.3, the orthologue of which in *A. thaliana* is LETHAL UNLESS CBC 7 B (*LUC7B*; AT5G17440.1). *AtLUCB7* is involved in mRNA splicing and thought to be important in the regulation of plant development (Francisco-Amorim et al., 2018), although limited characterisation has been carried out to date.

Bo7g110710.1, the orthologue of which in *A. thaliana* is RIBOSOMAL PROTEIN L18AE FAMILY (*L18ae*; AT4G26060.1), this candidate has not been characterised, with no publications specifically focused on *AtL18ae* on TAIR. Finally Bo2g161850.1, the orthologue of which in *A. thaliana* is SEUSS-LIKE 2 (*SLK2*; AT5G62090.2) and is involved in embryogenesis and post-embryonic development (Bao, Azhakanandam and Franks, 2010; Lee et al., 2014).

GEMs found to be predictive of seed B concentration were also correlated with RPKM values from the 383 panel (Table 4.4). All GEMs tested were negatively correlated with seed B concentrations, with increased expression associated with a decrease in seed B (all $p < 0.001$). Gene expression profiles were also compared between crop types, specifically winter ($n = 117$) and spring ($n = 91$) OSR, with mean expression varying significantly between Cab008178.3 ($p < 0.001$), Bo2g161850.1 and Bo7g110710.1 ($p < 0.05$).

Table 4.3 Predictive capability of gene expression markers from seed B 274 AT analysis

The most highly associated GEMs were tested for their predictive capability of seed B concentrations. The GEM and the –log10P from the 274 AT analysis are given, with the correlation coefficient (r) and significance (p) of the predictions using the 109 panel shown. Significant values are in bold and underlined. Finally, the A. thaliana orthologue and TAIR (Lamesch et al., 2012) description is provided.

Chromosome	GEM	AT 274 – log10P	Prediction: Correlation Coefficient (r)	Prediction: Significance (p)	*A. thaliana* orthologue	*A. thaliana* Annotation
A3	Cab016183.1	3.15	0.047	0.624	AT5G19430.1	RING/U-BOX SUPERFAMILY PROTEIN
A4	Cab026097.1	2.73	**0.192**	**0.045**	AT2G28790.1	PATHOGENESIS-RELATED THAUMATIN SUPERFAMILY PROTEIN
A5	BnaA05g13040D	3.00	0.144	0.139	AT3G29390.1	RS2-INTERACTING KH PROTEIN
A8	BnaA08g16910D	2.87	0.099	0.310	AT4G38790.1	ER LUMEN PROTEIN RETAINING RECEPTOR FAMILY PROTEIN
A9	Cab038298.3	3.62	**0.262**	**0.006**	AT5G61420.1	HIGH ALIPHATIC GLUCOSINOLATE 1
A10	Cab008178.3	2.79	**0.198**	**0.039**	AT5G17440.1	LETHAL UNLESS CBC 7 B
C2	Bo2g164020.1	4.78	0.154	0.109	AT5G62210.1	EMBRYO-SPECIFIC PROTEIN 3, (ATS3)
C2	Bo2g161850.1	4.30	**0.249**	**0.008**	AT5G62090.2	SEUSS-LIKE 2
C2	Bo2g161810.1	4.27	0.181	0.058	AT5G62000.3	AUXIN RESPONSE FACTOR 2
C3	Bo3g164310.1	3.64	0.132	0.147	*No orthologue*	--
C4	Bo4g142360.1	3.59	0.177	0.063	AT5G62130.2	PER1-LIKE FAMILY PROTEIN
C7	Bo7g110710.1	3.56	**0.214**	**0.025**	AT4G26060.1	RIBOSOMAL PROTEIN L18AE FAMILY

Chromosome	GEM	AT 274 – log10P	Prediction: Correlation Coefficient (r)	Prediction: Significance (p)	*A. thaliana* orthologue	*A. thaliana* Annotation
C8	Bo8g091260.1	4.37	0.173	0.071	*No orthologue*	--
C9	Bo9g176670.1	2.87	0.106	0.268	AT5G05970.2	TRANSDUCIN/WD40 REPEAT-LIKE SUPERFAMILY PROTEIN

Table 4.4 Effect of transcript abundance on seed B concentration

GEMs tested as predictive for seed B concentration and their A. thaliana orthologue information obtained from TAIR (Lamesch et al., 2012) are given. The correlation coefficient (r) and significance (p) of the correlation between transcript abundance (as reads per kb per million aligned reads; RPKM) and seed B concentration is then given (n = 381). Next the mean expression of winter (n=160) and spring (n=127) OSR is shown, with the significance (p) of this variation also stated. The blue and pink colouration shows relative low and high expression respectively. Any significant values are in bold and underlined.

Chromosomes	GEM	*A. thaliana* Annotation	383: Correlation Coefficient (r)	383: Significance (p)	Winter OSR: Mean Expression	Spring OSR: Mean Expression	Expression variation significance (p)
A4	Cab026097.1	PATHOGENESIS-RELATED THAUMATIN SUPERFAMILY PROTEIN	-0.253	**<0.001**	2.658	2.450	0.498
A9	Cab038298.3	HIGH ALIPHATIC GLUCOSINOLATE 1	-0.312	**<0.001**	12.933	14.687	0.450
A10	Cab008178.3	LETHAL UNLESS CBC 7 B	-0.282	**<0.001**	21.056	17.222	**<0.001**
C2	Bo2g161850.1	SEUSS-LIKE 2	-0.320	**<0.001**	11.065	14.033	**0.027**
C7	Bo7g110710.1	RIBOSOMAL PROTEIN L18AE FAMILY	-3.302	**<0.001**	2.410	1.950	**0.042**

4.2.3 Seed B associative transcriptomics: candidates within linkage disequilibrium of highly associated SNP markers chosen for further characterisation in A. thaliana

Analysis of seed B AT was primarily focused on the association peaks observed on chromosomes A5 and C4. These regions were homoelogous, with lead SNP marker of Cab046474.1:417:G and Bo4g034820.1:975:A respectively. As can be seen in the SNP AT Manhattan Plot (Figure 4.2a), the majority of markers within the association peak on C4 were light red points, representing hemi-SNP markers which have not been linkage mapped to a genome (marker can be in C4 or the homoelogous region on A5). Two candidates from this region on chromosome C4 were also identified within leaf AT and ultimately were found to vary significantly in leaf B concentrations. Firstly, Cab046461.3/ Bo4g034640.1 the orthologue for which in *A. thaliana* is STARCH BRANCHING ENZYME 2.1 (*BE3*; AT2G36390.1). A SNP (Bo4g034640.1:2223:C) within the leaf C4 association peak was found to be predictive of leaf B concentration. When tested, the mutant line had a significantly increased leaf B concentration ($p < 0.001$) (Figure 3.10). Secondly Cab025641.1/ Bo4g030810.1, the orthologue of which in *A. thaliana* is PLASMA MEMBRANE INTRINSIC PROTEIN 2 (*PIP2;2*; AT2G37170). A significant decrease in leaf B concentration was observed in the mutant line ($p <0.01$) (Figure 3.8). To assess whether any further ionome disruption could be characterised, both *BE3* and *PIP2;2* were also analysed for seed B concentrations. Additionally, the candidate identified within leaf AT, *ARF2*, was also observed within LD of the A2/C2 and A9/C9 association peaks within SNP AT. One homoeologue (CDS model Bo2g161810.1) was the fifth most highly associated GEM, although ultimately was not found to be predictive of seed B concentration (Table 4.3). Regardless, as this candidate had previously shown a significant decrease ($p <0.001$) in leaf B concentration within the mutant line (Figure 3.6), additional characterisation within the seed would assess whether any further ionomic disruption would be observed.

A further nine candidate genes, orthologues to three genes in *A. thaliana,* were chosen for further characterisation using T-DNA insertion lines. Information regarding these genes is provided hereafter and summarised in Table 4.5.

Table 4.5 Summary of candidates from seed AT analyses tested using A. thaliana T-DNA insertion lines

Markers identified from either SNP or GEM AT are first given, along with the A. thaliana orthologue and functional annotation taken from TAIR (Lamesch et al., 2012). Finally, the unique code for each SALK or SAIL T-DNA insertion line is stated.

SNP/GEM	Marker	*A. thaliana* Orthologue	Given Name	Annotated Function	SALK/SAIL Code
SNP	Cab046461.3 Bo4g034640.1	AT2G36390.1	*BE3*	STARCH BRANCHING ENZYME 2.1	SALK_030954C
SNP	Cab025641.1 Bo4g030810.1	AT2G37170.1	*PIP2;2*	PLASMA MEMBRANE INTRINSIC PROTEIN 2	SAIL_169_A03
SNP	Cab033998.1 Cab033997.1 Bo4g038360.1	XEG113	*XEG113*	XYLOGLUCANASE 113	SALK_007511C SALK_066991C
SNP	Cab025460.1 Bo4g026300.1	AT2G39510.1	*UMAMIT14*	NODULIN MTN21 /EAMA-LIKE TRANSPORTER FAMILY PROTEIN	SALK_037123C
SNP	Cab033878.1 Cab033876.1 Bo4g039860.1 Bo4g039880.1	AT2G34390.1	*NIP2;1*	NOD26-LIKE INTRINSIC PROTEIN 2;1	SAIL_439_A02 SALK_023890
SNP SNP/GEM	Cab021707.1 Cab038280.1 Bo9g014770.1 Bo2g161810.1	AT5G62000.3	*ARF2*	AUXIN RESPONSE FACTOR 2	SALK_035537

Firstly, Cab025460.1/ Bo4g026300.1 the orthologue of which in *A. thaliana* is USUALLY MULTIPLE ACIDS MOVE IN AND OUT TRANSPORTERS 14 (*UMAMIT14*; AT2G39510.1). A significant decrease in leaf B concentration between mutant and WT was observed on PiiMS. UMAMITs are amino acid transporters, with *AtUMAMIT14* mediated amino acid loading into seeds essential for seed development (Müller et al., 2015). Secondly, Cab033878.1/ Cab033876.1/ Bo4g039860.1/ Bo4g039880.1, the orthologue of which in *A. thaliana* is NOD26-LIKE INTRINSIC PROTEIN 2;1 (*NIP2;1*; AT2G34390.1). A significant increase in leaf B concentration between mutant and WT was observed on PiiMS. *AtNIP2;1* is a member of the NIP family of AQPs, which includes the known B transporters *AtNIP5;1, AtNIP6;*1 and *AtNIP7;1*. Furthermore, recent experimental evidence by Diehn et al., (2019) reported the B transport capacity of *BnaNIP2;1*s when expressed in *Xenopus* oocytes. Finally, Cab033878.1/ Cab033876.1/ Bo4g038360.1, the orthologue of which in *A. thaliana* is XYLOGLUCANASE 113 (*XEG113*; AT2G35610.1). *AtXEG113* is an arabinosyltransferase that modifies extensin proteins (EXTs) in root hair cells (Gille et al., 2009), with *Atxeg113* mutants exhibiting shorter root hair growth (Velasquez et al., 2011). The authors also discuss the potentiality of modification of EXTs within cell wall assembly of root hair as a means to increase plant biomass through both water and nutrient assimilation. The possible modulation of B concentrations through root hair proteins has been discussed in relation to controlling leaf B concentrations (Section 3.3.2.3) and as such this candidate was investigated further. Whilst possible that these candidate genes were in this associated region by chance, all markers were within estimated LD of the highly associated SNP markers (following guidelines outlined in Section 2.2.3).

4.2.4 Seed B candidate gene analysis: two candidate genes exhibit altered seed B concentrations

Of the *A. thaliana* T-DNA lines ordered for candidate analysis, three were also analysed in leaf (*BE3*; SALK_030954C, *PIP2;2*; SAIL_169_A03, *ARF2*; SALK_035537) as discussed in Section 3.3.1.1. No homozygous individuals were

identified for the *NIP2;1* line SALK_023890. Lines analysed via ICP-MS were compared individually with their respective Col-0 controls.

Once more, as the whole elemental profile was available via ICP-MS analyses, all essential elements whose percentage recovery was ≥ 85% were analysed. The following elements were included; B, Mg, P, S, K, Ca, Mn, Zn and Mo. As was also included within the analysis of the *NIP2;1* line. All lines saw significant differences in elemental concentrations in more than one element as compared to WT. For ease of comparisons between multiple elements, the subsequent figures show the percentage change from Col-0 for all elements studied. Specific statistical analyses (two-sided t-test with T-DNA line element vs WT control element) for all seed ionome candidates is provided throughout Appendix IV (*nip2;1* leaf is included in Appendix III).

A multi-element phenotype was observed within *nip2;1*, with a significant decrease in concentrations of six elements; B, P, K, Ca, Mn and Mo. A significant increase was observed in As (Figure 4.3). As was included within this analysis as additional *AtNIP*s facilitate arsenite transport (Bienert et al., 2008; Isayenkov and Maathuis, 2008). Both Mn and As seed AT analyses had clear association peaks on chromosome A5, with lead markers in LD of the *BnaA05NIP2;1* gene models (Cab033878.1/ Cab033876.1; lead markers Mn: Cab033880.2:639:G, As: Cab025693.1:180:C) (see Appendix II for Mn and As Manhattan Plots). Furthermore, the CDS model Cab033878.1 was the 9th most highly associated GEM within seed Mn AT and was significantly positively correlated with Mn concentrations ($R = 0.216$, $p < 0.001$, $n = 376$). Leaf *Atnip2;1* concentrations have been characterised on PiiMS, with a significant increase in B concentrations observed in the mutant line as compared to WT. However, within this analysis, leaf B concentrations were significantly decreased (Figure 4.4). A significant increase in seed B concentrations was observed in *xeg113b*, however no significant increase was observed within *xeg113a* (Figure 4.5). No significant variation in seed B concentrations was observed for the remaining four candidates (*be3, pip2;2 umamit14 and arf2*), with the elemental profiles provided in Figure 4.6 -Figure 4.9 respectively.

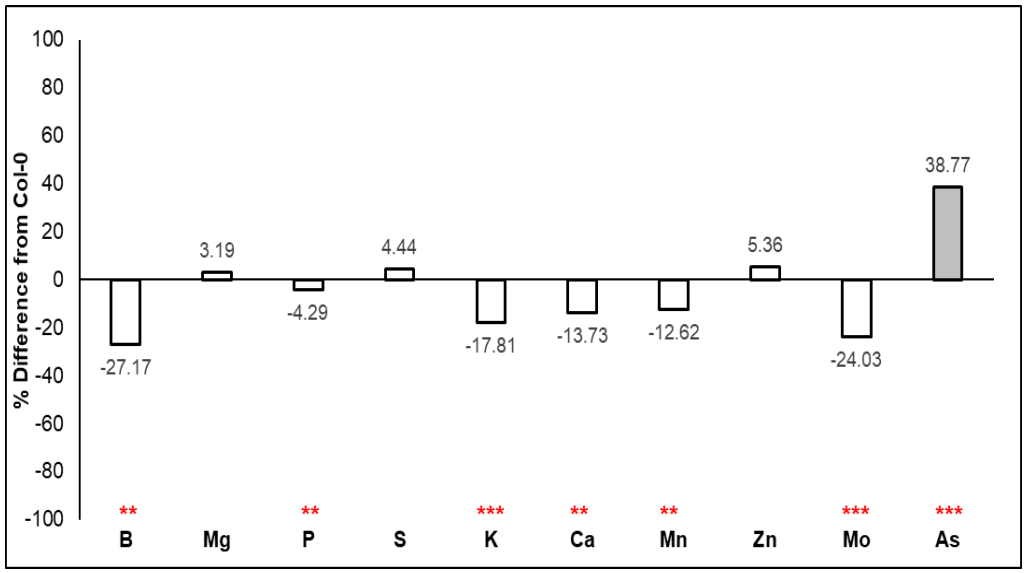


Figure 4.3 Percentage difference of elemental concentrations in seed nip2;1 compared to Col-0

Percentage difference of mean elemental concentrations (mg/kg) between control (Col-0) and arf2 T-DNA (SAIL_439_A02) A. thaliana seeds characterised via ICP-MS. Only essential elements whose percentage recovery was ≥ 85% were tested. As has also been included. Double and triple stars represent significance at p <0.01 and <0.001 respectively, as determined via a two-sided t-test for each element individually. A summary of statistical data is provided in Appendix 4.1.

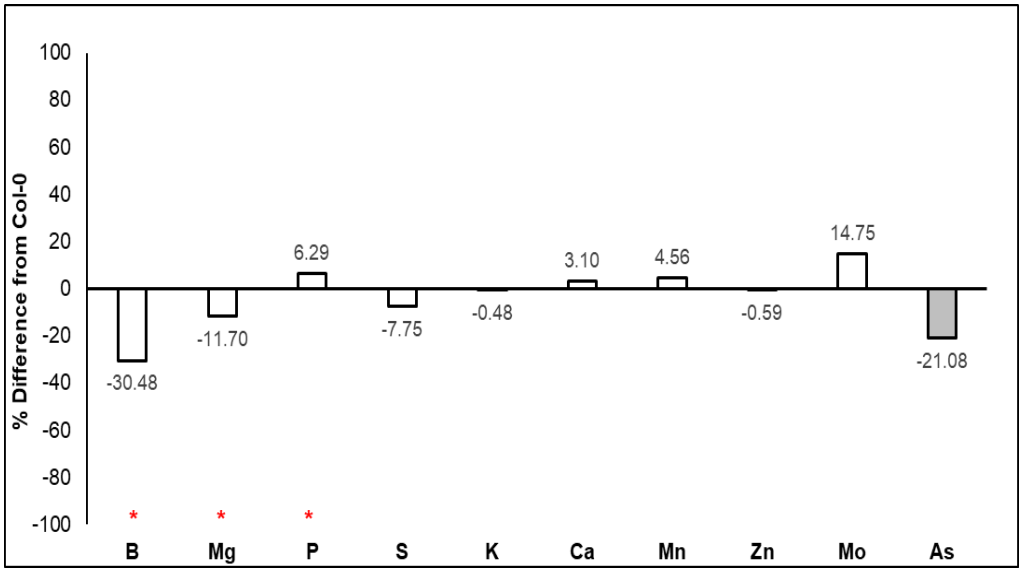


Figure 4.4 Percentage difference of elemental concentrations in leaf nip2;1 compared to Col-0

Percentage difference of mean elemental concentrations (mg/kg) between control (Col-0) and nip2;1 T-DNA (SAIL_439_A02) A. thaliana leaf characterised via ICP-MS. Only essential elements whose percentage recovery was ≥ 85% were tested. As has also been included. Single stars represent significance at p <0.05, as determined via a two-sided t-test for each element individually. A summary of statistical data is provided in Appendix 3.10.

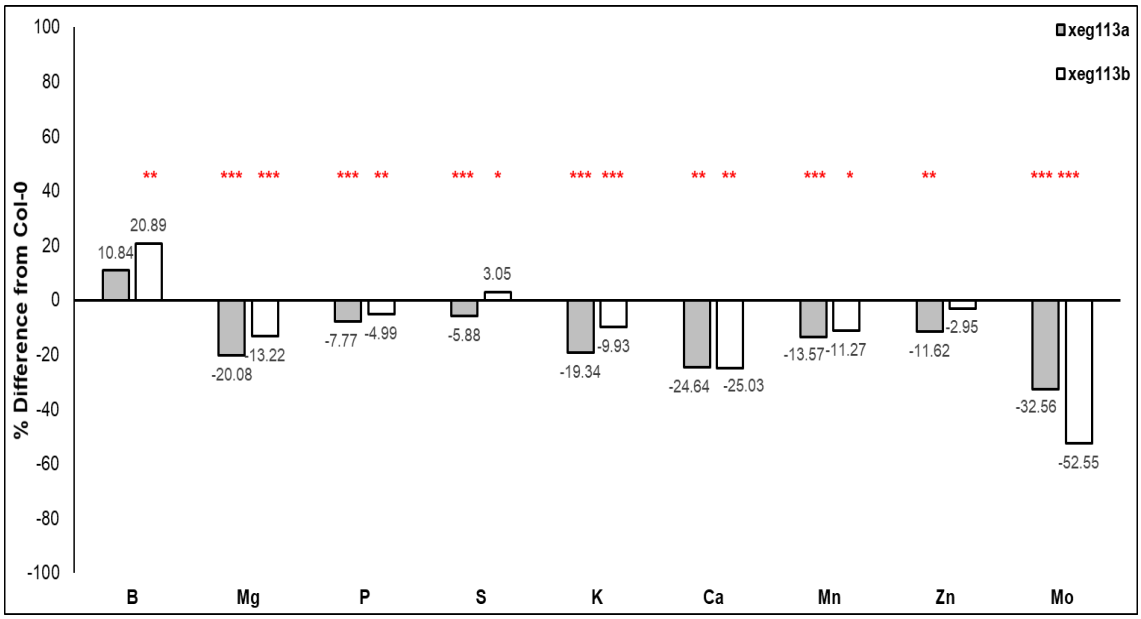


Figure 4.5 Percentage difference of elemental concentrations in seed xeg113a and xeg113b compared to Col-0

Percentage difference of mean elemental concentrations (mg/kg) between control (Col-0) and xeg113a T-DNA (SALK_007511C) and xeg113b T-DNA (SALK_066991C) A. thaliana seeds characterised via ICP-MS. Only essential elements whose percentage recovery was ≥ 85% were tested. Single, double and triple stars represent significance at p <0.05, p <0.01 and p <0.001 respectively, as determined via a two-sided t-test for each element individually. A summary of statistical data is provided in Appendix 4.6 and Appendix 4.7.

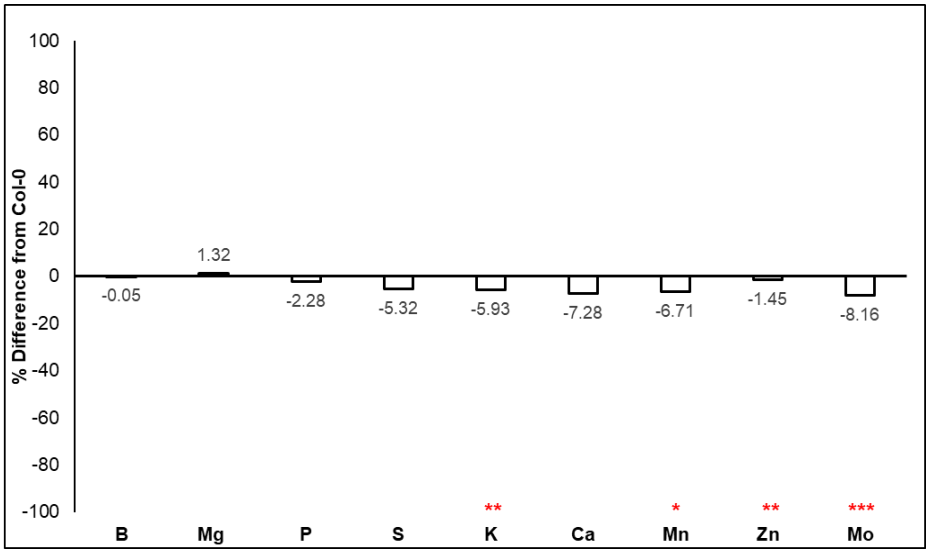


Figure 4.6 Percentage difference of elemental concentrations in seed be3 compared to Col-0

Percentage difference of mean elemental concentrations (mg/kg) between control (Col-0) and be3 T-DNA (SALK_030954C) A. thaliana seed characterised via ICP-MS. Only essential elements whose percentage recovery was ≥ 85% were tested. Single, double and triple stars represent significance at p <0.05, p <0.01 and p <0.001 respectively, as determined via a two-sided t-test for each element individually. A summary of statistical data is provided in Appendix 4.2.

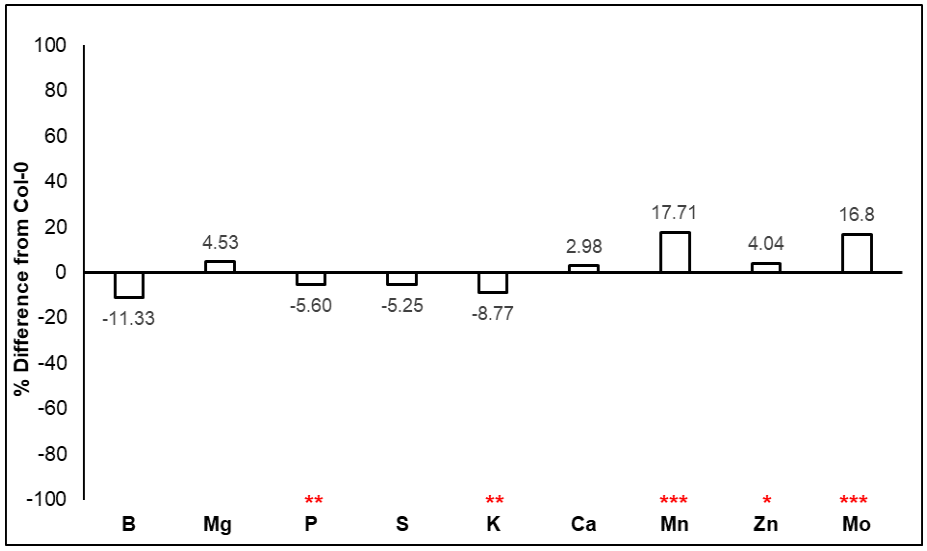


Figure 4.7 Percentage difference of elemental concentrations in seed pip2;2 compared to Col-0

Percentage difference of mean elemental concentrations (mg/kg) between control (Col-0) and pip2;2 T-DNA (SAIL_169_A03) A. thaliana seed characterised via ICP-MS. Only essential elements whose percentage recovery was ≥ 85% were tested. Single, double and triple stars represent significance at p <0.05, p <0.01 and p <0.001 respectively, as determined a via two-sided t-test for each element individually. A summary of statistical data is provided in Appendix 4.5

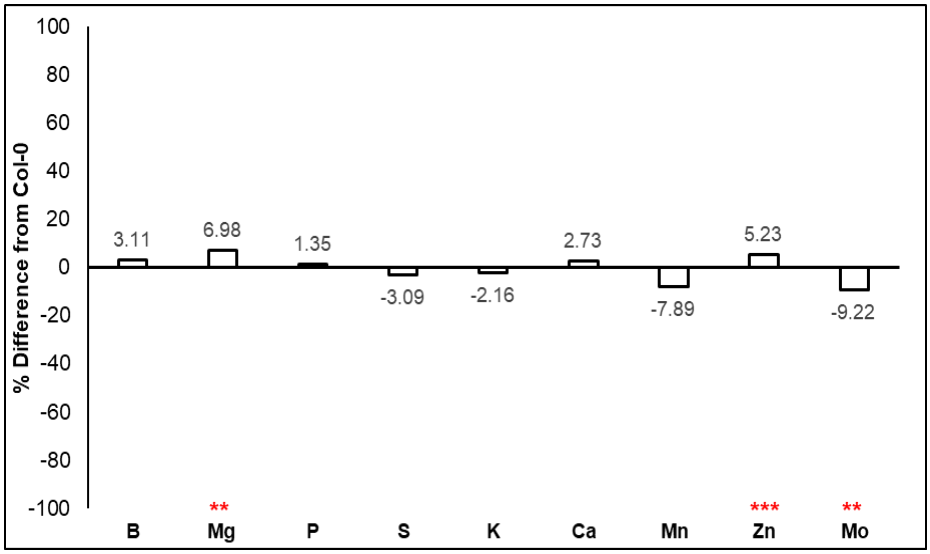


Figure 4.8 Percentage difference of elemental concentrations in seed umamit14 compared to Col-0

Percentage difference of mean elemental concentrations (mg/kg) between control (Col-0) and umamit14 T-DNA (SALK_037123C) A. thaliana seed characterised via ICP-MS. Only essential elements whose percentage recovery was ≥ 85% were tested. Double and triple stars represent significance at p <0.01 and p <0.001 respectively, as determined via a two-sided t-test for each element individually. A summary of statistical data is provided in Appendix 4.4.

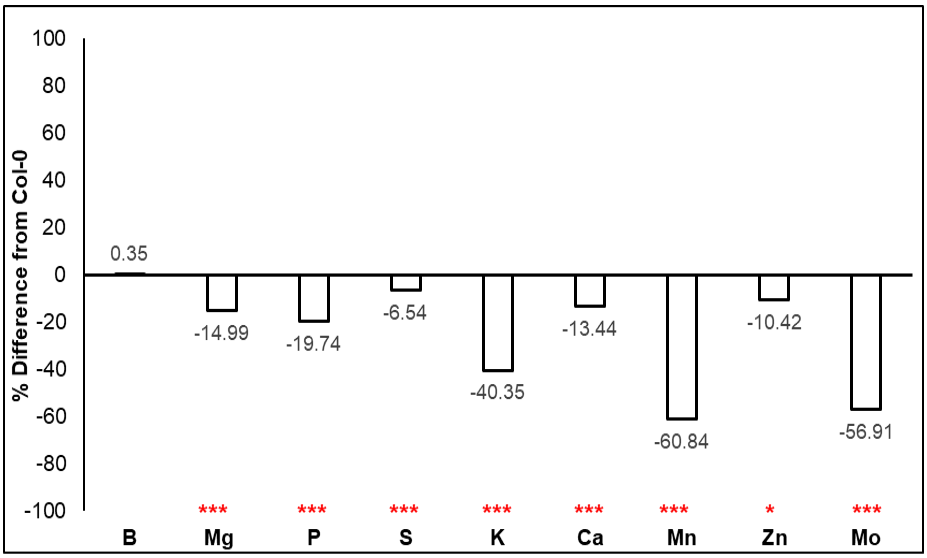


Figure 4.9 Percentage difference of elemental concentrations in seed arf2 compared to Col-0

Percentage difference of mean elemental concentrations (mg/kg) between control (Col-0) and arf2 T-DNA (SALK_035537) A. thaliana seed characterised via ICP-MS. Only essential elements whose percentage recovery was ≥ 85% were tested. Double and triple stars represent significance at p <0.01 and p <0.001 respectively, as determined via a two-sided t-test for each element individually. A summary of statistical data is provided in Appendix 4.3.

4.3 Discussion

4.3.1 Analysis of seed B concentrations in A. thaliana T-DNA insertion lines

4.3.1.1 Candidate gene analysis reveals A. thaliana mutants with altered seed B concentrations

In total 17 candidate genes, orthologues to 6 genes in *A. thaliana* were identified from both SNP and GEM AT analyses and further characterised using T-DNA insertion lines, with seed material (and *nip2;1* leaf material) analysed via ICP-MS. Any perturbation within the mutant line as compared to WT validated the candidate within the scope of this thesis' research. Two lines were found to vary significantly in seed B concentrations, with further information on these candidates provided hereafter.

Seed B concentrations in *nip2;1* (AT2G34390.1) was significantly reduced ($p < 0.01$), with a mean percentage difference of 27.17% (Figure 4.3). Four orthologues of *AtNIP2;1* were identified within AT analyses, which correspond to two tandem duplications on chromosome A5 and C4 (as also reported in Tao et al., 2014). It should be noted that Zhao et al., (2012) proposed that *BnaNIP2;1* was a candidate gene underlying QTLs for 'seed yield' and 'B-efficiency coefficient', as summarised in Table 1.3. However, these QTLs were mapped to chromosome C3 within this study, and not to chromosomes A5/C4. Like other members of the AQP protein family, *AtNIP2;1* exhibits water channel activity, although its permeability is relatively low (Choi and Roberts, 2007). *AtNIP2;1* is induced under anaerobic conditions and is able to transport lactic acid (Choi and Roberts, 2007). Interestingly, recent experimental evidence by Diehn et al., (2019) identified B transport capacity of one *BnaNIP2;1* isoform, *BnaA05.NIP2;1a*, when expressed in *Xenopus* oocytes.

Whilst the expression of both *BnaNIP2;1s* and *AtNIP2;1* are reportedly root-specific (Yuan *et al.*, 2017; Choi and Roberts, 2007), RPKM data used within this research identified significant expression of Cab033878.1 within the leaf transcriptome. Subsequent, analysis of expression profiles on Genevestigator has demonstrated that *AtNIP2;1* is expressed in many different tissues, including various inflorescence structures (Figure 4.10).

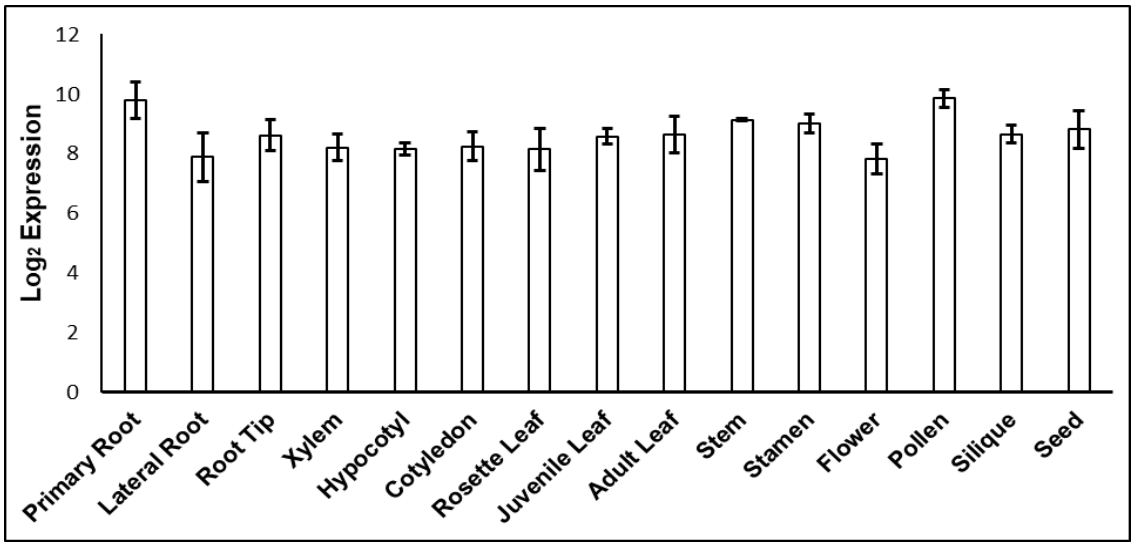


Figure 4.10 Expression of AtNIP2;1 in different plant tissues

AtNIP2;1 is expressed in various root, leaf and inflorescence tissues. Data was obtained from Genevestigator (Hruz et al., 2008), with data presented being the mean expression ± SD.

Protein structure is important for substrate selectivity of AQPs (Uzcategui et al., 2008). A study into the protein sequence of both *B. oleracea* and *B. rapa NIP2;1*s identified distinct amino acid differences compared to *A. thaliana,* including conserved sequences not found in the *A. thaliana* orthologue and varying protein lengths across the different gene copies. BolCNIP2;1.b, for example, is forty amino acids shorter than *AtNIP2;1* due to the introduction of a premature stop codon (Diehn et al., 2015). Furthermore in *B. napus*, the gene length of the four *BnaNIP2;1*s varies from 1111 bp to 1520 bp, although the protein length is consistently 286 amino acids. Differences in protein structure or gene length may be of importance when assessing the function of *BnaNIP2;1*s, for example, in case of sub-functionalisation of gene function, which can occur in polyploids (Adams et al., 2003). Yuan et al., (2017) also proposes this; stating that whilst gene structures were primarily conserved across *B. napus* AQPs, variation between homologous genes was observed (however the variation between *BnaNIP2;1*s was not specifically reported). This may be particularly relevant given that only one copy of *BnaNIP2;1* (Cab033878.1) was expressed in the leaf transcriptome, indicating differential expression profiles of the four *BnaNIP2;1*s. Resultantly, further clarification of the B transport capacity of each *BnaNIP2;1* is important and would be the next step within this research. Diehn et al., (2019),

for example, only reported transport capacity for a single isoform (the other three were seemingly not studied) and this research was conducted using *Xenopus* oocytes. Therefore assessing B transport capacity *in planta* is essential. Suitable methodology for this has been outlined in Diehn et al., (2019), who assessed the B transport ability of *BnaNIP4*, *BnaNIP5*, *BnaNIP6* and *BnaNIP7* isoforms transformed into a *Atnip5;1* knockout background under the control of the *AtNIP5;1* promoter. These lines were grown under B deficiency, with gene expression and shoot B concentrations subsequently quantified. Increased B concentrations in the transformed lines as compared to the *Atnip5;1* control provided quantitative evidence for functional B transport of these *BnaNIP*s. Therefore, the next step of this research would be to assess whether *BnaNIP2;1* can also facilitate B transport *in planta*, including assessing where B accumulates (i.e. shoot, inflorescence etc). As mentioned previously, this should be conducted for each of the four *BnaNIP2;1*s. Alternatively, mutational studies in *B. napus*, for example using TILLING lines, could be utilised. However, it would be essential that all *BnaNIP2;1*s were mutated in case of redundant gene function. This also wouldn't be a suitable option for assessing the functional B transport of each *BnaNIP2;1* individually (i.e. assessing which isoforms specifically facilitated B transport) unless multiple TILLING lines, with different gene copies knocked-out, were obtained.

Seed concentrations of P, K, Ca, Mn and Mo were also significantly reduced in *nip2;1* as compared to WT, whilst As was significantly increased. The significant increase in As concentrations was surprising given that additional *Atnip* mutants exhibited reduced total As content (Isayenkov and Maathuis, 2008). Furthermore, As transport capacity was reported for *BnaA05NIP2;1* when expressed in *Xenopus* oocytes (Diehn et al., 2019). It would therefore be expected a significant decrease in As concentrations would also be observed (as had been for B). As previously discussed, both Mn and As had association peaks clearly defined on chromosome A5 and had lead markers within LD of the *BnaNIP2;1* homoeologues (see Appendix IV for seed Mn and As Manhattan Plots). The *BnaA05NIP2;1* GEM, Cab033878.1, was also significantly positively correlated with seed Mn concentration ($p < 0.001$). This was a surprising result as, to the author's knowledge, no previous studies have proposed a role for *NIP*s in the control of Mn concentrations. Whist we can surmise that gene expression is

correlated with Mn concentrations, whether this gene is specifically controlling Mn concentrations is yet to be elucidated. Resultantly, this relationship warrants further investigation. Repeating ionomic profiling in *A. thaliana* with additional T-DNA insertion lines should initially be conducted before subsequent mutational studies in *B. napus* to assess the role *BnaNIP2;1* may play in controlling Mn concentrations. Assessing gene expression under different nutritional stresses (e.g. B or Mn deficiency/ toxicity) may also elucidate further mechanisms behind the control of these elements in *B. napus*.

The leaf ionome of *AtNIP2;1* was also studied, with a significant decrease ($p < 0.05$) in B concentrations in the mutant as compared to WT. A mean percentage difference of 30.48% was observed (Figure 4.4). This contrasts data obtained on PiiMS, from which a significant increase in B concentrations was observed within the mutant line. Further discussion on this has been provided in Section 4.3.1.2.

Two homozygous T-DNA insertion lines were obtained for *AtXEG113*. However, only one of the lines, *xeg113*b, showed significant variation; a significantly higher concentration of seed B ($p < 0.01$) was observed in the mutant line as compared to WT. Seed B concentrations in *xeg113a* were higher than WT, although this was not statistically significant ($p = 0.137$). Again, a multi-element phenotype was observed in both insertion lines; a significant decrease of Mg, P, K, Ca, Mn and Mo concentrations was observed in both T-DNA insertion lines. Comparatively, Zn was only significantly decreased in *xeg113a*, whereas S was significantly decreased in *xeg113a* and significantly increased in *xeg113b* (Figure 4.5). *AtXEG113* is involved within cell wall development and encodes an enzyme involved in the elongation of arabinose chains on EXTs (Gille et al., 2009). *Atxeg113* mutants exhibit shorter root hair growth (Velasquez et al., 2011) and increased lateral root development (Roycewicz and Malamy, 2014) however these phenotypes are not *Atxeg113* specific. Furthermore, the assembly of other wall types (such as pollen tubes), has been linked to such enzymes (Velasquez et al., 2011). The expression profile of *AtXEG113* is summarised in Figure 4.11.

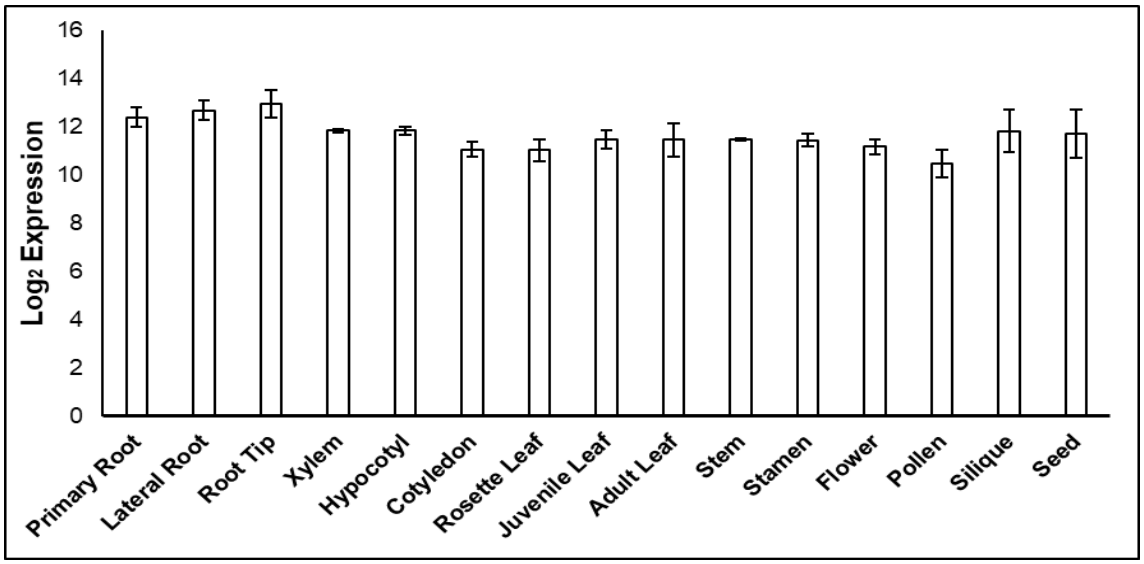


Figure 4.11 Expression of AtXEG113 in different plant tissues

AtXEG113 is expressed in various root, leaf and inflorescence tissues. Data was obtained from Genevestigator (Hruz et al., 2008), with data presented being the mean expression ± SD.

The alteration of root hair morphology within the mutant could explain the reduction in elemental concentrations, given that root hairs are important for mineral uptake (*c.f.* Tanaka *et al.*, 2014), although why an increased concentration of B was observed is unclear. However, analysis of RSA was not conducted and resultantly further research into the root hair phenotype of mutants should be conducted. Cell organisation within these mutants may also be of importance, given that the primary function of B *in planta* is within the formation of the cell wall. *Atxeg113* mutants exhibit greater cell elongation (Gille et al., 2009) which may resultantly affect B accumulation in different tissues (i.e. greater cell elongation may increase B integration into the cell wall). As such, further phenotyping in assessing the specific cell organisation is pertinent, which may elucidate why B concentrations are perturbed within the mutant.

4.3.1.2 Critical review of seed mutant characterisation

This research presents a method for validating candidate genes putatively involved in seed B accumulation, as identified through AT and subsequent testing in *A. thaliana* T-DNA insertions lines. Only one T-DNA insertion line was used in the ionomic analysis of leaf (Figure 4.4) and seed (Figure 4.3) material for *AtNIP2;1*. As discussed within Section 3.3.1.2, whilst acceptable for an initial

screen of candidate gene function, characterisation of additional T-DNA insertion lines is required. This is particularly relevant given that the two T-DNA lines analysed for *AtXEG113* did not vary significantly in both lines (both lines had increased seed B as compared to WT but only *xeg113b* was significant) (Figure 4.5). These two lines were grown at the same time, which should have minimised variation in both the environmental conditions within the growth room and the media used for plant growth (i.e. soil was from the same batch). The *xeg113a* T-DNA line could have been a knock-down (as opposed to knock-out) which may have affected the ionomic profile, although as transcript quantification was not conducted this cannot be assumed. However, the majority of other elements studied showed significant changes in both T-DNA lines, therefore negating this argument. Interestingly Baxter, (2010) described the difficulties in reproducibility when analysing seed elemental concentrations. This could be due to the age of the material analysed; for example seed harvested from different shoots would have developed at different time-points and subsequently dried on different days, which may ultimately impact the ionomic profile. Pooling seed from multiple plants, as was done within this research, should account for some of this variability. However, this should be considered for future seed ionome experiments.

Leaf B concentrations of *AtNIP2;1* were significantly reduced within the mutant line as compared to WT. This is in contrast with data obtained from PiiMS, where a significant increase in leaf B concentrations was reported. Data for only one T-DNA insertion line was available on PiiMS (the other T-DNA lines reported for *Atnip2;1* had insertions in two genes and not a single locus), with this the primary reason leaf material was also analysed. The discrepancy between results presented here and on PiiMS could be a result of many factors, including the specific T-DNA (line on PiiMS was a GABI-Kat line rather than SAIL), leaf material analysed (e.g. developmental stage) or growth media (e.g. specific soil used). Again, this highlights the importance of repeating ionomic experiments (both within this research and across the wider literature) and that minimising experimental variation between studies is essential.

Three candidate genes identified within leaf AT, which subsequently exhibited significant alterations in leaf B concentrations when tested in *A. thaliana* (*BE3*,

ARF2 and *PIP2;2*), were also identified within seed AT. *BE3* and *PIP2;2* were identified within the A5/C4 association peaks whilst *ARF2* was in the A2/C2 and A9/C9 association peaks. The *ARF2* GEM on chromosome A2, Bo2g161810.1, was also tested for trait predictability (Table 4.3). As this marker was not predictive of seed B concentrations it is not surprising that no significant differences in seed B concentrations were observed (Figure 4.9). Furthermore, all of these markers could have been identified within seed B by chance. For example *ARF2* markers in seed were identified within regions associated with the *AtHAG1* orthologues, and therefore it is likely the markers have been co-selected due to the extensive *B. napus* breeding programmes to lower seed GSL (Allender and King, 2010).

4.3.2 Candidates within LD of seed B association peaks may be causative of variation in seed B concentrations

4.3.2.1 Candidates identified within the A5/C4 association peaks

As previously discussed, the remaining four candidates tested using *A. thaliana* T-DNA insertion lines (*be3, pip2;2, umamit14 and arf2*) did not show a significant variation in seed B concentrations (Figure 4.6 - Figure 4.9). However, there were a selection of additional candidates identified from AT which were not analysed in *A. thaliana*. These candidates were identified within LD of highly associated SNPs within the homoelogous A5/C4 regions. Candidates have been discussed hereafter and summarised in Table 4.6.

Firstly, Cab046487.1/ Bo4g035060.1 the orthologue of which in *A. thaliana* is CASPARIAN STRIP MEMBRANE DOMAIN PROTEIN 1 (*CASP1*; AT2G36100.1). *AtCAPS1* is essential for Casparian strip formation (Roppolo et al., 2011) and is positively regulated by *AtMYB36* (Kamiya et al., 2015), a transcription factor identified with a minor association peak observed on chromosome A10 within leaf B AT. Secondly, Cab025664.1/ Bo4g031090.1 the orthologue of which in *A. thaliana* is METHIONINE ADENOSYLTRANSFERASE 3 (*SAMS3*; AT2G36880.1). *AtSAM3* is involved in the biosynthesis of S-adenosylmethionine (SAM) from ATP and L-Methionine (Ugalde et al., 2011). SAM is a methyl-group donor, essential for methylation reactions (Chen, Zou and

McCormick, 2016). *Atsams3* mutants have impaired pollen tube growth, reduced seed set (Chen, Zou and McCormick, 2016) and a reduced lignin concentration (Shen, Li and Tarczynski, 2002). Furthermore, a study in *B. napus* found that *BnaSAMS3* was highly upregulated under B deficiency, with the hypothesis that methylation activation under B deprivation is a mechanism to compensate for impaired cell wall structure (Wang et al., 2011). This candidate was also identified within QTL analyses for plant height traits (Zhao et al., 2012) (Table 1.3). Finally Cab025636.1/ Bo4g030720.1, the orthologue of which in *A. thaliana* is TRANSPARENT TESTA GLABRA 2 (*TTG2*/*WRKY55*; AT2G37260.1). *TTG2* is a WRKY transcription factor involved within the development of the seed coat (Johnson, Kolevski and Smyth, 2002), with this thought to be the primary site of B storage in the seed (Eggert and von Wiren, 2013).

Table 4.6 Candidate genes within linkage disequilibrium of highly associated SNP markers from seed B AT analyses

Candidates within LD of the most highly associated SNP makers within an association peak are summarised; the lead marker within the association peak and the candidate gene are given, as is the A. thaliana orthologue and annotated function on TAIR (Lamesch et al., 2012). If significnant variaition in leaf B concentrations between mutant and WT A. thaliana lines has been previoulsy reported on PiiMS (Baxter et al., 2007), this is also stated.

Association Peak	Lead SNP Marker	Candidate Locus	*A. thaliana* orthologue	Given Name	Annotated Function	PiiMS Leaf B variation
A5	Cab046474.1:417:G	Cab046461.3	AT2G36390.1	*BE3*	STARCH BRANCHING ENZYME 2.1	-
A5	Cab046474.1:417:G	Cab025641.1	AT2G37170.1	*PIP2;2*	PLASMA MEMBRANE INTRINSIC PROTEIN 2	-
A5	Cab046474.1:417:G	Cab025460.1	AT2G39510.1	*UMAMIT14*	NODULIN MTN21 /EAMA-LIKE TRANSPORTER FAMILY PROTEIN	Yes
A5	Cab046474.1:417:G	Cab033878.1 Cab033876.1	AT2G34390.1	*NIP2;1*	NOD26-LIKE INTRINSIC PROTEIN 2;1	Yes
A5	Cab046474.1:417:G	Cab033998.1 Cab033997.1	AT2G35610.1	*XEG113*	XYLOGLUCANASE 113	No
A5	Cab046474.1:417:G	Cab046487.1	AT2G36100.1	*CASP1*	CASPARIAN STRIP MEMBRANE DOMAIN PROTEIN 1	-
A5	Cab046474.1:417:G	Cab025664.1	AT2G36880.2	*SAMS3*	METHIONINE ADENOSYLTRANSFERASE 3	-
A5	Cab046474.1:417:G	Cab025636.1	AT2G37260.1	*WRKY44/ TTG2*	TRANSPARENT TESTA GLABRA 2	-
C4	Bo4g034820.1:975:A	Bo4g034640.1	AT2G36390.1	*BE3*	STARCH BRANCHING ENZYME 2.1	-
C4	Bo4g034820.1:975:A	Bo4g030810.1	AT2G37170.1	*PIP2;2*	PLASMA MEMBRANE INTRINSIC PROTEIN 2	-
C4	Bo4g034820.1:975:A	Bo4g026300.1	AT2G39510.1	*UMAMIT14*	NODULIN MTN21 /EAMA-LIKE TRANSPORTER FAMILY PROTEIN	Yes

Association Peak	Lead SNP Marker	Candidate Locus	*A. thaliana* orthologue	Given Name	Annotated Function	PiiMS Leaf B variation
C4	Bo4g034820.1:975:A	Bo4g039860.1 Bo4g039880.1	AT2G34390.1	*NIP2;1*	NOD26-LIKE INTRINSIC PROTEIN 2;1	Yes
C4	Bo4g034820.1:975:A	Bo4g038360.1	AT2G35610.1	*XEG113*	XYLOGLUCANASE 113	No
C4	Bo4g034820.1:975:A	Bo4g035060.1	AT2G36100.1	*CASP1*	CASPARIAN STRIP MEMBRANE DOMAIN PROTEIN 1	-
C4	Bo4g034820.1:975:A	Cab025664.1	AT2G36880.2	*SAMS3*	METHIONINE ADENOSYLTRANSFERASE 3	-
C4	Bo4g034820.1:975:A	Bo4g030720.1	AT2G37260.1	*ARKT44/ TTG2*	TRANSPARENT TESTA GLABRA 2	-

4.3.2.2 Seed B shares association peaks with seed S

The association peaks seen within both SNP and GEM AT on chromosomes A2, A9, C2 and C9 correspond to regions of genome deletion containing orthologues of *AtHAG1* (At5g61420), a regulator of aliphatic GSL production (Hirai et al., 2007). Breeding for low GSL accessions has been conducted (*c.f.* Wang et al., 2017) due to the anti-nutritional properties of seed GSL (Halkier and Gershenzon, 2006), with the deletion of *HAG1* homoeologues in low-GSL *B. napus* accessions found to be causative of these association peaks (Harper et al., (2012); reviewed in Section 1.3.4.2). GSL are S-rich compounds (Halkier and Gershenzon, 2006) and there is a clear relationship between S and GSL concentrations, with the former used as an estimate for GSL content (Kühn-institut and Kühn-institut, 2005). It is therefore unsurprising that the four association peaks (A2/C2, A9/C9) are seen in seed sulphur AT (Appendix II). There is also a strong negative correlation between seed S and B concentrations ($p < 0.001$) (Table 4.1), with this being the second most correlated pair of elements found within seed from the RIPR diversity panel (Thomas et al., 2016). However, why the association peaks corresponding to *AtHAG1* are observed in seed B, and whether this is a direct or indirect relationship, is as yet unclear. There is no suggestion that B plays a role within GSL biosynthesis and the observed associations may be a result of the extensive breeding programmes to lower seed-GLS, which may have indirectly perturbed additional seed traits. However, a correlation between B and GSL concentrations has been reported, with B fertilisation found to lower overall seed GSL (Yang et al., 2009; Jankowski et al., 2016), although specific mechanisms behind this have not yet been put forward. The GEM Cab038298.3 (*HAG1* homoeologue within A9 association peak) was also predictive of seed B phenotype, with its expression negatively correlated with seed B concentration ($p < 0.001$) (Table 4.4). The relationship between *HAG1* and the wider seed ionome was not directly investigated as part of this thesis and was a focus of another member of the Bancroft research group; however further characterisation would be pertinent. Analysis of B concentration in low and high GSL backgrounds, or analysing GSL in plants grown under differing B concentrations, could further elucidate this relationship.

4.3.3 Utilising AT to understand the genetic control of seed B concentrations

Given that transcriptome sequencing for AT analyses was conducted at the 2nd true leaf stage, genetic variation affecting seed B concentrations may not have been available for association analyses. For example, this dataset could have limited the discovery of genes expressed only during reproductive development. Limited detection of genes previously identified from QTL studies for seed B traits (Table 1.3) was also observed in seed AT (although this was also the case for leaf B traits). Resultantly, it could be argued that having a more representative transcriptome, to include genes expressed during seed development (e.g. sampling RNA from whole seeds), would have been beneficial. However, as discussed in detail within Section 3.3.3, this would have vastly increased the work involved in sampling the RIPR diversity panel.

No SNP or GEM marker passed the Bonferroni or FDR for seed B AT. The lack of SNPs passing the significance thresholds may be a result of the complexity of the trait, which can cause issues when conducting association analyses. For example, if the trait is controlled by many rare allelic variants which have a large effect on the phenotype, or conversely if many common variants have a small phenotypic effect; their detection may be limited (Korte and Farlow, 2013). However, the significance thresholds used within this AT analysis were fairly stringent; the 5% false discovery rate was maintained across all analyses to set the significance thresholds to $p < 0.05$. Two recent studies in *B. napus* implemented different FDRs; Wang et al., (2016) reports true marker-trait associations should return an FDR of < 0.1, whilst Cai et al., (2014) used an FDR of 0.2. Furthermore, Hejna et al., (2019) outlined their own significance thresholds as -log10P = 5 and -log10P = 4, rather than using the FDR and Bonferroni. Whilst significance thresholds could have been lowered (or decided on a case by case basis), this was decided against in order to maintain consistency between other AT studies conducted using the RIPR diversity panel (*c.f.* Alcock et al., 2017, 2018; Havlickova et al., 2018).

4.4 Summary and conclusions

The specific focus of the research presented in this chapter was assessing whether natural variation in seed B concentrations across the *B. napus* diversity panel is a result of underlying genetic variation. SNP and GEM AT outputs were successfully analysed for candidate genes linked to the control of B concentrations in seed.

No SNP or GEM markers passed the FDR or Bonferroni corrected significance thresholds. However, discernible association peaks were observed within both SNP and GEM analyses, particularly on chromosomes A2/C2 and A9/C9 and also on chromosome A5/C4. The A2/C2 and A9/C9 association peaks corresponded to regions of genome deletion containing orthologues of the GSL regulator *AtHAG1*. It is possible that the observed associations are a result of the extensive breeding programs over the previous decades to reduce seed GSL content (Allender and King, 2010), which may have indirectly resulted in additional seed traits being perturbed. However, further research into the relationship between B and GSL was not investigated directly as part of this current research.

Lead SNP markers were tested for trait predictability, with two found to be predictive of seed B concentration (Table 4.2). The effect of allelic variation was also tested, with B concentrations in accessions with the reference allele and the most frequent allelic variant compared. A significant variation between reference and the most common allelic variant was observed for four markers ($p < 0.001$). The effects were lower than those seen for leaf B concentrations, with up to a 1.19-fold change between allelic variants (compared to 1.80-fold for leaf; Table 3.3). However, this is still indicative of the scope to improve seed B accumulation across the diversity panel. GEMs were also tested for trait predictability, with five markers predictive of seed B concentrations (Table 4.3). Expression for three GEMs also varied significantly between winter and spring crop types (Cab008178.3 $p < 0.001$; Bo2g161850.1 and Bo7g110710.1 $p < 0.05$) (Table 4.4).

Two candidates identified within LD of the lead SNP markers within association peaks on chromosome A5 and C4, exhibited perturbed B concentration when tested in *A. thaliana*. Undoubtedly *NIP2;1* was the most interesting candidate identified within seed B AT analyses, which subsequently exhibited perturbed B

concentrations in the *A. thaliana* T-DNA line. Further research is required to assess whether *BnaNIP2;1*s are also able to facilitate B transport *in planta*.

Chapter 5: Investigating the B deficiency disorder Brown Heart

5.1 Introduction

Swedes are a *B. napus* variety cultivated as a root vegetable. Technically speaking the swede itself is not a true root, but a swollen hypocotyl - however for ease 'root' shall be used herein. BH is a physiological disorder found within the marketable swede root and characterised through internal brown discolouration. The disorder was first reported in the 1930s (Hurst and MacLeod, 1936) and is attributed to a localised B-deficiency within the developing root. Limited research focusing on BH has currently been published to date. A review of the current literature is presented herein.

A field trial by Sanderson, Sanderson and Gupta, (2002) was conducted across three years looking at the control of BH in two swede accessions. Plants were sown with one of four pre-plant broadcast (fertiliser applied directly to the soil) B treatments (0.00, 2.24, 4.48 and 6.72 kg ha^{-1}), and then supplied with either 0.00 or 2.24 kg B ha^{-1} as a foliar spray post seeding (with these given as two 1.12 kg B ha^{-1} treatments 28 and 42 days after seeding). Roots were harvested after approximately four month's growth, with yields (root weight) and BH severity rated at three time points (upon trial harvesting and 4 or 7 months of storage at 2°C). Both pre-plant broadcast and foliar B treatments decreased overall swede yields, with this being in concordance with field trial results from Gupta and Cutcliffe, (1978) whom similarly saw yield reductions with increasing broadcast and foliar sprays; surmised to be due to B toxicity with increasing B application. Whilst yields were reduced, BH severity was similarly reduced with increasing broadcast and foliar sprays in both studies (Gupta and Cutcliffe, 1978; Sanderson, Sanderson and Gupta, 2002). This was thought to be an acceptable compromise given roots need to be clear of BH to be deemed fit for human consumption. Sanderson, Sanderson and Gupta, (2002) also assessed leaf B concentrations under the differing treatments, however surprisingly root B concentrations in BH clear and affected tissue was not assessed. Leaf B concentrations increased linearly with both broadcast and foliar applications, however whether this was

correlated with BH symptoms (e.g. whether roots with higher leaf B concentration exhibited less instances of BH) was also unfortunately not assessed. Of the two accessions used within the trial, one was found to be more resistant to BH. Shelp and Shattuck, (1987) also assessed the severity of BH in two (different) accessions, similarly finding one to be more resistant. Broadcast and foliar B applications also decreased overall BH incidents.

More recently another collaborator of Elsoms Seeds Ltd (Spalding, Lincolnshire, UK) used twelve of their parent and hybrid lines to assess how genetic background and B supplementation affected the presence and severity of BH (Fadhel et al., 2013). Within this field trial, swedes were treated with differing concentrations of B-ethanolamine (0.00, 1.35, 1.80 and 2.70 kg B/ha) at three time points (pre sowing and two post-emergence sprays). Like past trials (*c.f.* Gupta and Cutcliffe, 1978; Sanderson, Sanderson and Gupta, 2002), certain accessions exhibited increased susceptibilities to BH, further implicating a genetic link to the disorder (some accessions were also used within this research and have been summarised in Table 5.1). B concentrations were measured via ICP-OES, with roots exhibiting BH symptoms generally showing lower root B concentrations as compared to clear roots. However, whether these reductions were statistically significant was surprisingly not reported. Additionally, significant negative correlations between root B content and the severity of BH was provided. One could argue that replicating the field trial on consecutive years (as done by Sanderson, Sanderson and Gupta, (2002)) would increase the validity of these results; given the multitude of changing environmental factors that cannot easily be controlled in field trial experiments (which may have affected BH incidents). Furthermore, such environmental factors (e.g. average temperature, rainfall) should be assessed and reported, which was not the case here. For example, Sanderson, Sanderson and Gupta, (2002) reported rainfaill within their study, and postulated that increased BH incidents in one particular year could have been a result of reduced rainfall (and resultantly reduced soil moisture) which may have affected B availability and susbequent uptake from the soil.

As there are no pre-harvest detection methods to identify BH affected roots, growers have to assess the marketability of their crop post-harvest. Furthermore, as roots affected with BH are deemed inedible, significant yields are lost each

year. Resultantly, the application of fertilisers is a common recommendation (Sanderson, Sanderson and Gupta, 2002). These are primarily in the form of fertiliser drilled into the soil prior to sowing, or post-emergence foliar spraying of crops. However, although some swede accessions are thought to be 'BH tolerant', farmers are still reporting incidences of BH regardless of B supplementation. Beauchamp and Hussain, (1974), for example, noted that application of B (through Borax fertiliser) did not completely mitigate BH symptoms. Incidences of BH can also occur to different severities within differing years, or even within different locations (e.g. different fields within a specific growing region). This fact was noted in early work by Dermott and Trinder, (1947) who were looking to assess the incidents of BH in Cumbria, UK. Whilst BH was widespread, the number and severity of roots affected varied widely, including field to field. Rudimentary B extraction (hot water) and quantification (assessing the colour change of the red compound 'alizarin' when exposed to B) were conducted in this trial. However, significant differences in root B concentrations were not reported.

Irrespective of past research, field trials conducted by Elsoms have often identified no significant differences in root B concentrations in BH affected and clear tissue (unpublished data). Furthermore, as previously mentioned, incidents are reported in roots thought to be BH tolerant. As such, the purpose of this research was to address two key questions:

1. *Does swede accession affect the presence and severity of BH symptoms within the root?*
2. *Does supplementation with B affect the presence and severity of BH symptoms within the root?*

Rather than using a field trial design, in which various additional factors could be affecting BH presence; such as sub-localised differences in soil pH, nutrient availability and moisture, a polytunnel experiment was designed (as discussed in the methods below). Briefly, sixteen swede accessions, comprising Elsoms parent, hybrid and RIPR lines, were sown into the soil substitute medium terra-green (Oil-Dri UK Ltd, Cambridgeshire, UK). These plants were then treated with different concentrations of boric acid. The trial was initially conducted within the glasshouses at the University of York and subsequently moved to a polytunnel at

Elsoms, where plants were transplanted into large individual pots. The rationale behind this experimental design was that in using terra-green, which as an inert medium requires supplementation with a comprehensive nutrient supply, the availability of B could be more tightly controlled. Additionally, the plants would be grown in uniform conditions in a controlled environment. The aim of this research was to gain further insight into whether swede accession and B availability are the sole factors affecting BH presence, or whether additional factors are at play.

5.2 Methods

5.2.1 Swede lines chosen

A total of 16 swede accessions were used in this case study. Accessions were sourced from both Elsoms (n = 11) and the RIPR diversity panel (n = 5). Accessions from the RIPR panel were ordered based on their leaf B concentration (mg/kg), with the three highest and two lowest accessions chosen (Figure 5.1). Information regarding accessions chosen for the trial is documented in Table 5.1. The BH susceptibility of some lines has been reported by Elsoms and published by Fadhel et al., (2013).

Table 5.1 Swede accessions used for BH trial

The anonymised code for each accession used with the trial is provided, as is the source (either Elsoms Seeds Ltd or RIPR panel) of the line. Whether the accession is a parent, hybrid or double haploid is stated, as is the susceptibility of the accession under investigation as reported in Fadhel et al., (2013).

Accession Code	Source	Accession Type	Reported Susceptibility?
Mg	Elsoms	Parent	Resistant
Lz	Elsoms	Parent	Resistant
He	Elsoms	Parent	Resistant
Ac	Elsoms	Parent	Susceptible
Ag	Elsoms	Parent	Susceptible

Accession Code	Source	Accession Type	Reported Susceptibility?
Sl	Elsoms	Parent	-
Tk	Elsoms	Parent	-
Mg x He	Elsoms	Hybrid	Resistant
He x Ac	Elsoms	Hybrid	Susceptible
Ag x He	Elsoms	Hybrid	Susceptible
G1*He	Elsoms	Double Haploid	-
Bh	RIPR	-	-
Hg	RIPR	-	-
Tn	RIPR	-	-
Kf	RIPR	-	-
BT	RIPR	-	-

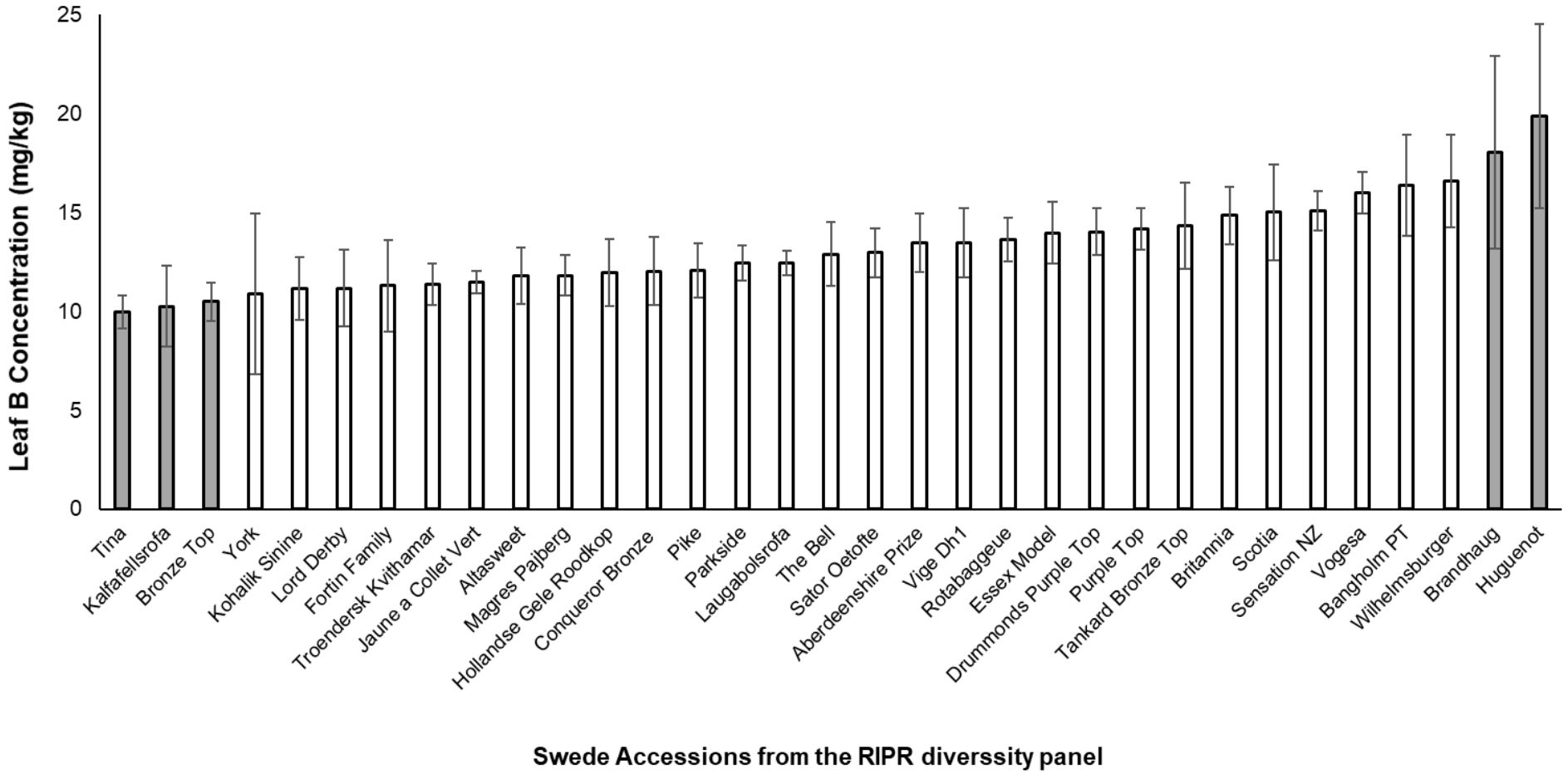


Figure 5.1 Leaf B concentrations for swede accessions in RIPR panel

Variation observed in leaf B (mg/kg) concentration from 33 swede accessions. Mean and ± standard error are reported for up to 5 replicates, dependent upon the number of data points available. The five accessions used within the trial are shown in grey and represent accessions with the three lowest and two highest leaf B concentrations.

5.2.2 Boric acid formulation

Three boric acid (H_3BO_3) treatments were used in this study; 0.25 µM, 10 µM and 40 µM. These treatments were determined by previous studies (Han et al., 2007; Liu et al., 2009) and represented low, sufficient and high treatments respectively (low, mid and high hereafter). The nutrient solution utilised was based on Murashige-Skoog (MS) (Murashige and Skoog, 1962) basal salt medium and was formulated by Grotech Productions Ltd (Goole, UK). Whilst not all of the components of MS basal salt medium are required for plant growth (aside from when conducting tissue culture), this formulation was successfully used to grow OSR seedlings in terra-green by another member of the Bancroft research group and therefore composition was kept constant for this trial. Medium composition is shown in Table 5.2.

Table 5.2 Formulation of plant nutrient solution used in BH trial

Component	MS Concentration (mg/L)	Novel MS Concentration (mg/L)	% of Solution
Potassium nitrate	1900.0	1900.0	42.91
Ammonium nitrate	1650.0	1650.0	37.27
Calcium chloride	332.20	332.20	7.50
Magnesium sulfate	180.70	180.70	4.08
Potassium phosphate monobasic	170.0	170.0	3.84
myo-Inositol	100.0	100.0	2.26
Disodium EDTA dihydrate	37.260	37.260	0.84
Ferrous sulfate heptahydrate	27.80	27.80	0.63
Manganese sulfate monohydrate	16.90	16.90	0.38
Zinc sulfate heptahydrate	8.60	8.60	0.19
Boric acid	*6.20*	*0.00*	*0.00*
Glycine	2.0	2.0	0.045

Component	MS Concentration (mg/L)	Novel MS Concentration (mg/L)	% of Solution
Potassium iodide	0.830	0.830	0.019
Nicotinic acid	0.50	0.50	0.011
Pyridoxine hydrochloride	0.50	0.50	0.011
Sodium molybdate dihydrate	0.250	0.250	0.0056
Thiamine hydrochloride	0.10	0.10	0.0023
Cobalt chloride hexahydrate	0.0250	0.0250	0.00056
Cupric sulfate pentahydrate	0.0250	0.0250	0.00056

2.2 g of nutrient medium was required for every 1 L of solution prepared; as a large quantity of nutrient media was needed for supplementation, generally 10 L of each of the three treatments was prepared at any given time. For this, 22 g of nutrient medium was weighed and subsequently diluted with 10 L ddH_2O. A stock solution of 0.1 M H_3BO_3 (Grotech Production Ltd, Goole, UK) was supplied to the media to give the three desired treatments (low 2.5 µL 0.1 M H_3BO_3 per 1 L of media; mid 100 µL/L; high 400 µL/L). pH was adjusted to 6.0 using 1 M hydrogen chloride (HCl). Soil pH is one of the key factors affecting B uptake (Shorrocks, 1997) and therefore pH 6.0 was chosen as an optimal pH so as not to inhibit B availability (although conventional soil itself was not used within this study). Medium was subsequently left on a magnetic stirrer for at least 12 hours to allow H_3BO_3 to fully diffuse into the nutrient solution.

5.2.3 Growth conditions

Sixteen swede accessions, three nutrient treatments with four biological replicates were used for this trial. Three seeds for each accession were sown directly into individual 25 cm pots containing the soil substitute terra-green (Oil-Dri UK Ltd, Cambridgeshire, UK). Sowing took place on the 13th July 2016. Plants

were randomised within trays, as determined by treatment and replicate. For this, within each treatment, plants from each biological replicate (comprised of 16 accessions within each replicate) were randomly split across 2 trays, with 8 plants in each tray. Trays were arranged so that they were randomised within each treatment, but blocked by treatment.

Plants were grown in a controlled glasshouse under sixteen hours of light (5am - 9pm), with supplementary lighting provided when light intensity fell below 250 W. On average, day temperatures were 25˚C with night lows of 20˚C. One-week post sowing, seedlings were thinned randomly to one plant per pot, giving a total of 192 plants. Plants were watered twice daily, with 500 mL nutrient solution supplied to each tray weekly. Seven weeks post sowing (31st August 2016), plants were moved into a polytunnel at Elsoms seeds (Figure 5.2) and transplanted into individual 10 L pots. Plants were randomised within each treatment and blocked by treatment. No supplementary lighting or heating was provided. An automatic drip irrigation system was used for plant watering. Whilst ddH2O was used to prepare the nutrient solution, this option was not feasible for daily watering due to the facilities available. The maximum permissible availability of B in UK ground water is 1.0 mg/L (as detailed in the UK water supply regulations; Drinking Water Inspectorate, 2018) and therefore the B supplied to the plants may have been greater than that of the desired treatments (see Discussion). 180 mL of nutrient solution was supplied to each plant every 10 days. Thiacolprid and prothioconazole was used for aphid and mildew control respectively.

Figure 5.2 Growth of plants in a polytunnel at Elsoms Seeds

Plants were moved from the University of York, transplanted in to 10 L pots and placed in a polytunnel at Elsoms Seeds on the 31st August 2016. Plants were randomised across the polytunnel, but blocked by H_3BO_3 treatment.

5.2.4 Tissue collection and analysis

The newest emerged leaf from each plant was sampled at three time-points; three (3-4 true leaf stage), six and eleven weeks post sowing. The newest emerged leaf was chosen for sampling (and subsequent ionome analyses) as immature leaves are more sensitive to B-deficient conditions and therefore are a better indicator of deficiency symptoms (Huang, Ye and Bell, 1996). Leaves were oven-dried at 90°C for 48 hours, with dried material stored under air-tight conditions with silica gel included to restrict moisture damage.

Harvesting of the trial occurred on the 14th December 2016, following approximately five months growth. At this time, one final leaf was sampled (again this was the newest emerged leaf which was subsequently oven-dried), with the roots brought back to the University of York for analysis. Roots were sliced transversely and the presence of BH assessed with a 0-3 scoring system rating symptom severity (0 = clean; 1 = slight; 2 = moderate; 3 = severe), as shown in Figure 5.3. Root sub-sections (~ 3 cm across and whole length of root) containing both BH affected and clean tissue were then sampled. These root samples were then stored at -20°C before being freeze-dried for 48 hours to constant weight.

All of the root and leaf material sampled during the trial was homogenised before 100 mg of each was digested in $HNO_{3,}$ (following methods outlined in Section 2.4). Samples were analysed in five ICP-MS runs at the University of Nottingham. Results have been provided throughout and summarised in Appendix V.

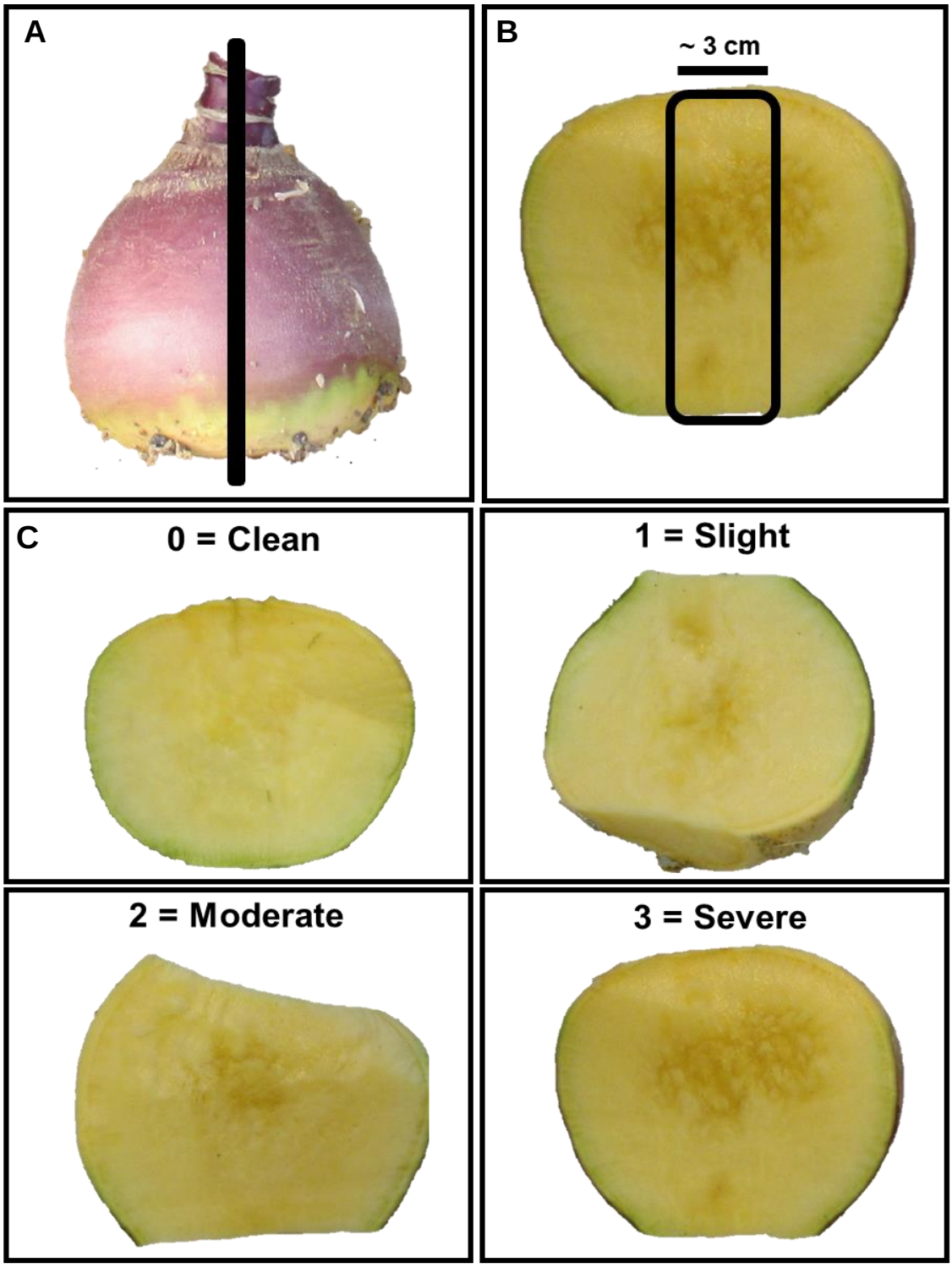


Figure 5.3 Sampling swede roots and scoring for BH symptoms

a. *Image of swede root with line depicting roughly where the root would be transversely sliced for BH scoring* ***b.*** *Depiction of how root was sub-sampled for elemental analysis; this would include the whole length of the root by approximately 3cm across* ***c.*** *scale representing the 0-3 scoring system for BH symptoms.*

5.2.5 Statistical Analyses

All data from the BH trial was analysed using GENSTAT 19th edition (VSN International, Hemel Hempstead, UK). Chi-Squared analyses was used to asses for significant differences in instances of BH between treatments, when the total number of expected incidents was > 5 for 80% of the data-points. If this assumption was not met, then the Freeman-Halton extension of Fisher's exact test was used to analyse the differences between counts. ANOVA with post-hoc Tukey honest significant difference (HSD) testing was used to assess for significant differences in B concentrations in root and leaf material between H_3BO_3 treatments. Two-sided t-test was used to compare the means of both root and leaf elemental concentrations in accessions exhibiting BH with those clear of symptoms. Pearson correlation coefficients (r) were calculated for element v element correlations (e.g. leaf B V root B), whereas Spearman's correlation (r_s) analysis was performed between elemental concentrations and BH severity. A threshold of $p < 0.05$ was used to assess statistical significance.

5.3 Results

5.3.1 Presence and severity of BH in swede roots

Growing swedes in terra-green, which to the authors knowledge has not been attempted before, enabled more control over factors which could be affecting plant growth (e.g. nutrient availability). However, upon starting this research it was initially unclear how well the plants would grow (and whether the swede root would fully develop) within this medium. Of the sixteen accessions used within the trial, five were obtained from the RIPR diversity panel. Of this, no swede root growth was seen in two accessions (Hg and BT) and as such these accessions were omitted from further analysis. Whilst swede yield (e.g. root weight) was not assessed as part of this trial, no visual differences in plant growth were observed between the three H_3BO_3 treatments for the remainder of the accessions.

The total number of roots analysed within each treatment and the number of roots exhibiting each symptom, is summarised in Table 5.3. 24 and 20 roots were affected with BH in low and mid treated plants respectively. However, under high H_3BO_3 the total number of affected roots decreased from 24 in low, to 11 in high

treated (30 and 44 clean roots in low and high respectively). The number of roots exhibiting severe symptoms was low across all treatments, with no roots exhibiting severe symptoms observed in high treated plants. Despite this, no significant difference in the number of roots affected with BH between treatments was observed, as determined by chi-squared (χ^2) ($p = 0.1004$) (Table 5.3).

Table 5.3 Two-way contingency table for number of roots affected with brown heart

160 swede roots were assessed for the presence of BH across the three H_3BO_3 treatments. Severity of symptoms were scored on a 0-3 rating system, with the number of roots exhibiting each of the symptoms shown. The total number of roots examined for each treatment (far right) and the total number of root exhibiting each system (bottom) is also provided. The Chi-squared statistic was calculated to determine if there was a significant difference in number of roots affected with BH between treatments, of which no significant difference was observed ($p = 0.1004$).

	BH Severity				**Total Roots**
	Clean (0)	**Slight (1)**	**Moderate (2)**	**Severe (3)**	
Low	30	12	9	3	54
Mid	31	11	5	4	51
High	44	7	4	0	55
Total Roots	105	30	18	7	**160**

The BH severity of each individual accession is shown in Figure 5.4. For this, the BH severity exhibited in each root was scored as a proportion (of each root within an accession and treatment), with data then split according to accession. Only one accession, Tk, exhibited no BH symptoms across all replicates and treatments.

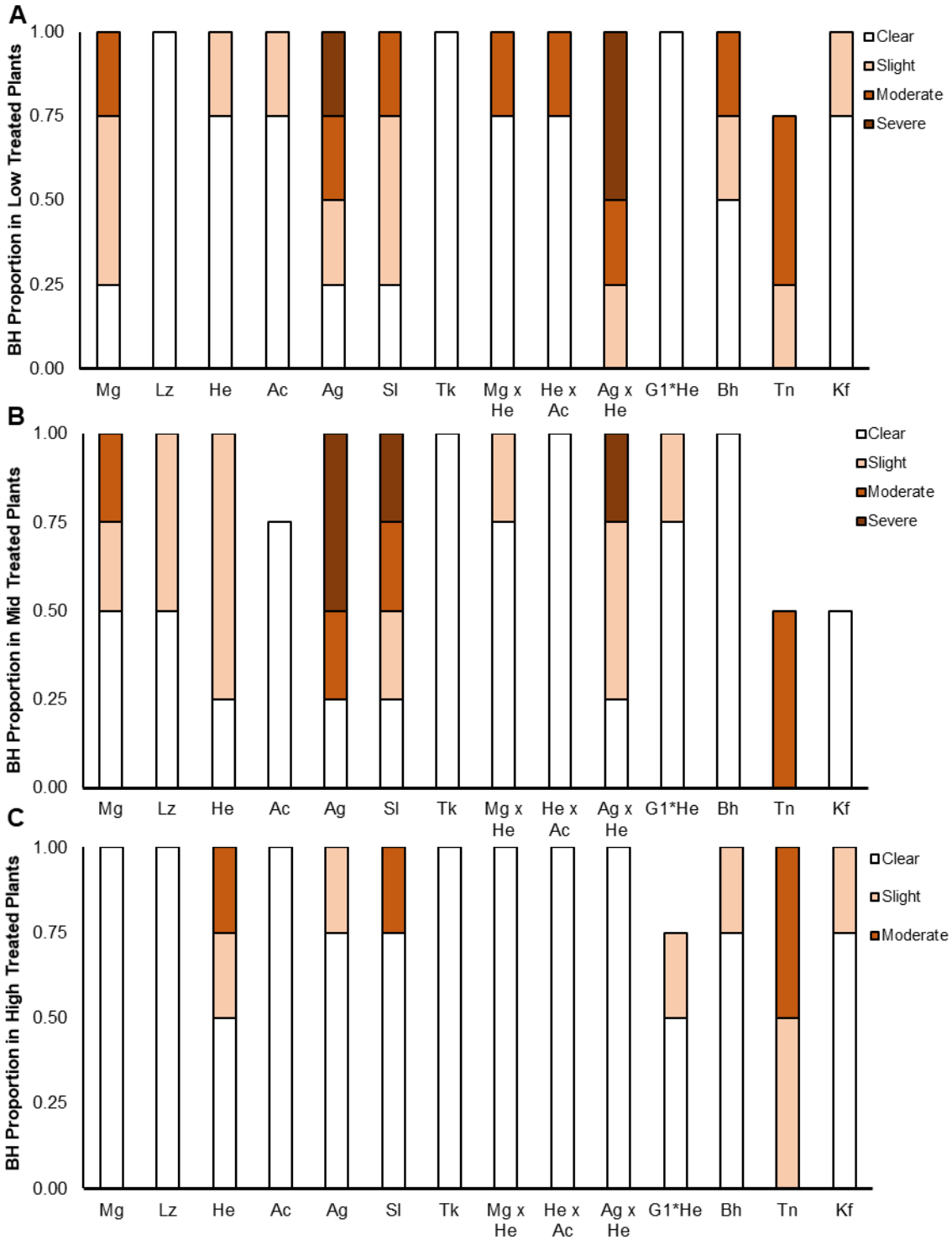


Figure 5.4 Proportion of BH severity observed in different swede accessions treated with (a) Low (b) Mid (c) High concentrations of H_3BO_3

The severity of BH for each accession treated with low, mid and high H_3BO_3 from four replicates was scored based on a 4-point system: 0 = clean, 1 = slight, 2 = moderate 3 = severe. Some roots showed an array of symptoms across the four replicates, e.g. Ag in (a), whereby in each a differing severity was seen; therefore each root appears as 0.25 of the total proportion for this accession.

The susceptibility to BH of certain accessions is summarised in Table 5.1, with four thought to be more four more resistant to BH (Mg, Lz, He and Mg x He) and four more susceptible (Ac, Ag, He x Ac and Ag x He). However, there was no significant difference in the total number of BH affected roots in the resistant as compared to the susceptible accessions ($p = 0.650$) (Appendix 5.1).

5.3.2 B concentrations within root

5.3.2.1 Effect of accession and treatment on root B concentrations

There was no significant interaction between accession and treatment on swede B concentrations ($p = 0.215$) (Appendix 5.2). However, when assessing accession and treatment individually, both independently significantly affected swede B concentrations. There was a significantly higher concentration of B within the high treated roots ($p < 0.001$) as determined by post-hoc Tukey testing, however B concentrations in low and mid treated roots were not significantly different (Appendix 5.3). Post-hoc Tukey testing of B concentrations in the 14 accessions also identified significant ($p < 0.001$) differences between accessions when not splitting by treatment (Appendix 5.4)

To further assess how accession affected B concentrations (i.e. which lines had significantly higher B within each treatment), data was then split according to accession. Significant differences in B concentrations were observed between accessions when focusing on each treatment individually, as identified through post-hoc Tukey testing (Table 5.4). However, when analysing B concentrations between treatments (specifically looking at whether B varied between treatments within each individual accession); limited significant variation was observed (Figure 5.5). To be specific, variation between treatments for each accession was only observed for Mg, He and Ag x He accessions. ANOVA results have been summarised in Appendix 5.5.

Table 5.4 ANOVA results of B concentrations in swede accessions across H_3BO_3 treatments

The F statistic (F), degrees of freedom and p value (p) are given, alongside the significant differences in mean B concentrations from up to four replicates, as highlighted with post hoc Tukey testing (p <0.05) of H_3BO_3 treatments across swede accessions. Where the means are significantly different the letters are not shared. Red and blue colouration shows relative high and low B concentrations respectively.

Accession	Low: Mean B Concentration (mg/kg)	Mid: Mean B Concentration (mg/kg)	High: Mean B Concentration (mg/kg)
	$F_{(13,41)} = 5.13$ **p= <0.001**	$F_{(13,38)} = 2.99$ **p= 0.004**	$F_{(13,41)} = 2.21$ **p= 0.027**
Mg	11.42^{abcd}	12.96^{b}	16.75^{b}
Lz	9.45^{abc}	8.71^{a}	9.04^{a}
He	10.27^{abcd}	8.93^{ab}	13.10^{ab}
Ac	8.69^{ab}	9.79^{ab}	12.60^{ab}
Ag	8.01^{a}	9.37^{ab}	11.15^{ab}
Sl	13.29^{bcd}	11.91^{ab}	12.40^{ab}
Tn	14.36^{cd}	12.85^{ab}	14.15^{ab}
Mg x He	11.15^{abcd}	10.60^{ab}	10.52^{ab}
He x Ac	8.48^{ab}	8.64^{a}	10.38^{ab}
Ag x He	9.46^{abc}	10.33^{ab}	13.58^{ab}
G1*He	14.83^{d}	11.98^{ab}	14.13^{ab}
Bh	9.81^{abcd}	9.64^{ab}	10.46^{ab}
Tk	9.25^{abc}	10.47^{ab}	11.63^{ab}
Kf	13.87^{cd}	9.86^{ab}	13.27^{ab}

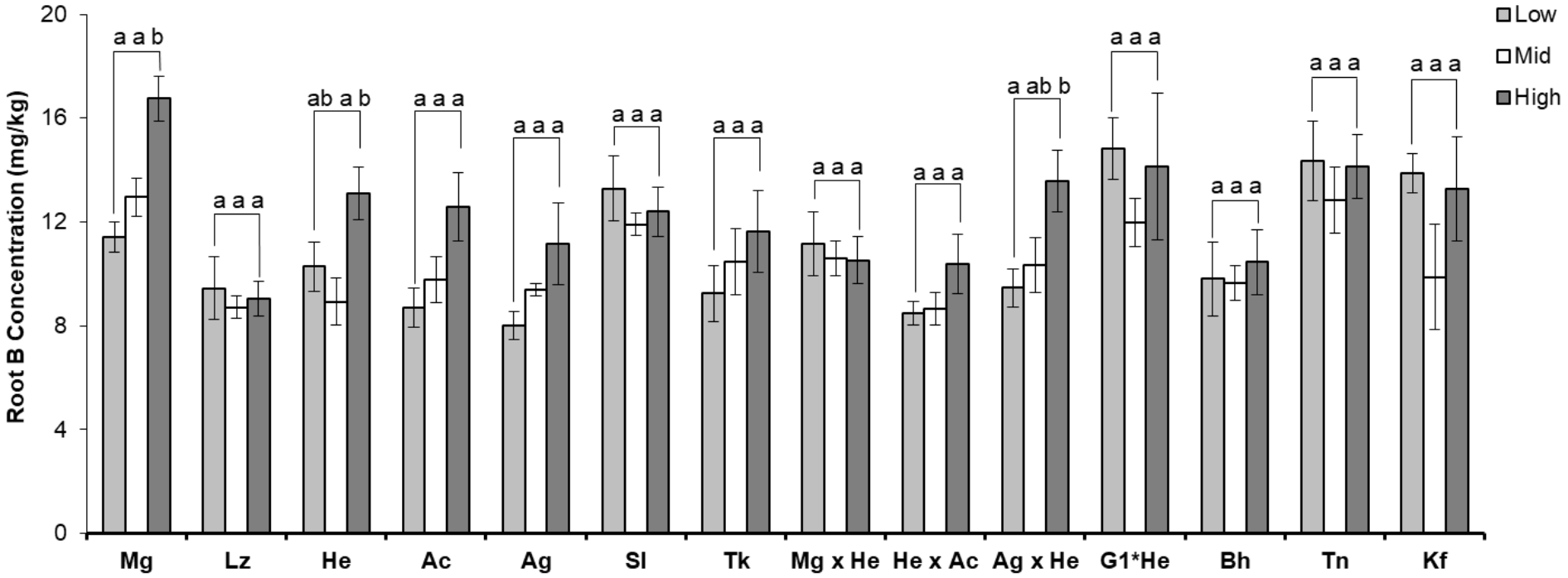


Figure 5.5 B (mg/kg) concentrations within swede roots split by accession and H_3BO_3 treatment

Fourteen accessions treated with three H_3BO_3 concentrations from up to four replicates were analysed for their root B concentration (mg/kg). Mean B concentration is presented with ± one standard error as error bars. Where the means of root B concentrations are significantly different the letters are not shared, as identified through post hoc Tukey testing of ANOVA results. See Appendix 5.4.

To summarise, B concentrations within swede roots exhibited significant variation when looking at all accessions within a treatment, indicating that certain accessions had increased uptake capacities (Table 5.4). Furthermore, mean B concentrations within high treated roots were significantly higher than low and mid treated roots ($p < 0.001$) (Appendix 5.3). However, when assessing each accession individually (i.e. B concentrations within roots across the three treatments) limited significant differences were observed (Figure 5.5).

5.3.2.2 Effect of root B concentration on presence or severity of BH

The effect of B concentrations within root and the presence of BH was assessed. For this, B concentrations within roots exhibiting any BH symptoms (scores 1-3 combined) was compared with B concentrations in clean roots. No significant differences in mean swede B concentrations between clean and BH affected roots was observed when assessing each of the three treatments individually, as summarised in Table 5.5. Furthermore, no correlation between BH severity and root B concentrations for accessions exhibiting BH was observed ($r_s = 0.013$, $p = 0.924$). These results were not necessarily surprising, given that overall instances of BH were not significantly reduced between treatments (Table 5.3).

Table 5.5 T-Test results for root B concentrations in BH affected and clean swede roots across H_3BO_3 treatments

Mean B concentrations in roots exhibiting BH (combined 1-3 score) and no BH symptoms between H_3BO_3 treatments are given, as is the number (n) for each. The T value (T), probability (p) and degrees of freedom (df) are then provided for each treatment.

Treatment	No BH: average swede B concentration	n	BH: average swede B concentration	n	T	p	df
Low	10.758	31	10.830	24	0.09	0.926	21,27
Mid	10.232	31	10.837	20	1.06	0.296	18,30
High	12.055	44	13.460	11	1.07	0.174	10,42

5.3.3 B concentrations within leaf

Leaf material was sampled at four developmental time points (TP1; 3-4 true leaf stage, TP2; 6-weeks post sowing, TP3; 11-weeks post sowing, TP4; newest emerged leaf upon root harvesting). Mean B concentrations in TP1, TP2 and TP3 were significantly higher in high treated plants as compared to low treated plants (TP1 $p < 0.001$, TP2 and TP3 $p < 0.01$). Interestingly, within TP4, an inverse pattern was observed, with B concentrations in high treated plants significantly lower than the low treated plants ($p < 0.05$) (Appendix 5.6). Leaf B concentrations within TP1 and TP4 were negatively correlated ($r = -0.487$, $p < 0.01$). In concordance with results demonstrated within the root, limited variation in leaf B concentrations between accessions was observed (Appendix 5.7 - Appendix 5.10).

Although no significant difference in B concentrations between roots exhibiting BH and clean roots was observed (Table 5.5); leaf B concentrations were also analysed. For this, corresponding leaf B concentrations for roots exhibiting BH (scores 1-3 combined) and clean roots were compared. No significant difference in mean leaf B concentrations between clean and BH affected roots was observed when assessing each of the three treatments individually for all leaf time points (Table 5.6).

Table 5.6 T-Test results for leaf B concentrations in BH affected and clean roots for four time points across H_3BO_3 treatments

Mean B concentrations in leaves for plants whose roots exhibited BH (combined 1-3 score) and no BH symptoms between H_3BO_3 treatments are given for TP1 - TP4, as is the number (n) for each. The T value (T), probability (p) and degrees of freedom (df) are then given for each treatment.

Time Point	Treatment	No BH: average leaf B concentration	n	BH: average leaf B concentration	n	T	p	df
TP1	Low	26.18	30	27.46	24	1.03	0.303	21,27
	Mid	38.04	30	36.00	20	0.82	0.414	18,30
	High	35.17	44	39.04	11	1.77	0.082	10,42
TP2	Low	33.43	30	30.43	24	1.68	0.097	21,27
	Mid	36.25	30	35.52	20	0.35	0.729	18,30
	High	37.79	44	37.80	11	0.00	0.998	10,42
TP3	Low	25.66	30	26.58	24	0.67	0.506	21,27
	Mid	24.03	30	23.86	20	0.90	0.121	18,30
	High	27.84	44	29.59	11	0.86	0.392	10,42
TP4	Low	23.63	28	25.27	22	0.93	0.218	21,27
	Mid	22.08	31	20.98	19	-0.89	0.379	18,30
	High	21.61	11	21.71	43	0.09	0.926	10,42

5.3.4 BH and additional elemental interactions

There was no effect of B availability (i.e. H_3BO_3 treatment), or B concentrations within the root, that were associated with BH symptoms. As such other elemental interactions were investigated. BH severity symptoms (scores 0-3; n = 158) was significantly correlated with Mg, Ca and Mn concentrations (Table 5.7). Furthermore, when assessing the relationship between elemental concentrations in BH affected and clean roots, both Ca and Mn concentrations were significantly higher in roots affected with BH than those clear of symptoms (Table 5.8).

Table 5.7 Correlation between elemental concentrations and BH symptoms

The Spearman correlation coefficient and significance for the correlation between elemental concentrations and BH symptoms (scores 0-3); n = 158. Significant values are in bold and underlined.

Element	Spearman Correlation Coefficient (r_s)	Significance (p)
B	0.050	0.496
Mg	0.162	**0.042**
P	0.141	0.079
S	-0.116	0.148
K	0.038	0.629
Ca	0.288	**<0.001**
Mn	0.221	**0.005**
Fe	0.118	0.140
Mo	0.043	0.574
Cu	0.088	0.271

Table 5.8 T-Test results for swede elemental concentrations in BH affected and clean roots across H_3BO_3 treatments

Mean elemental concentrations in roots exhibiting BH (combined 1-3 score) and no BH symptoms across all three treatments are given. The T value (T), probability (p) and degrees of freedom (df) are also provided. Significant values are in bold and underlined.

Element	BH: element average	No BH: element average	T	P	df
B	11.379	11.142	0.502	0.616	53,103
Mg	1274.657	1218.480	1.656	0.099	53,103
P	3387.339	3245.473	1.322	0.188	52,103
S	3580.927	3856.432	1.828	0.067	52,102
K	21194.011	20833.829	0.555	0.579	52,103
Ca	2810.023	2549.372	3.263	**0.001**	53,103
Mn	14.042	12.826	2.472	**0.015**	53,103
Fe	23.714	21.717	1.686	0.094	52,103
Cu	1.536	1.261	0.737	0.463	53,103
Mo	0.059	0.060	0.014	0.988	53,103

5.4 Discussion

5.4.1 Effect of accession and B supplementation on BH incidents

The aim of this trial was to address two key questions posed by both swede breeders and growers, relating to potential control mechanisms that can be put in place to limit the incidents of BH. Firstly, this trial assessed whether swede accession played a role in the propensity for roots to exhibit BH symptoms, as has been reported in past studies investigating this disorder (Shelp and Shattuck, 1987; Sanderson, Sanderson and Gupta, 2002; Fadhel et al., 2013). As provided in Figure 5.4 and summarised by accession (total roots across all treatments) in Table 5.9, the incidents of BH did vary across accessions. Tn was the only accession where symptoms were observed in all roots, regardless of H_3BO_3 treatment, although it should be noted that number of roots analysed for this accession was nine (as opposed to the total n of twelve). Accessions with Ag in their background were also more susceptible to BH (Ag and Ag x He hybrid), as was previously reported by Fadhel et al (2013). The previously determined resistant accessions (Hg, Mg, Lz and Mg x He) did show an overall proportional incidence less than that observed for the susceptible accessions (Ac, Ag, He x Ac and Ag x He). This was particularly the case with the Mg x He hybrid, with a incidence of 0.17, equating to two affected roots in total. However, the number of BH affected roots was not significantly reduced in resistant lines as compared to susceptible (Appendix 5.1). This may be due to the fact that a relatively small number of roots were available for analysis within each treatment, which may have limited the power of any subsequent statistical analyses. In hindsight, scoring fewer accessions for the presence of BH would have been beneficial and would have allowed for more replicates to be studied. In specifically focusing on those accessions determined as either resistant or susceptible, there would be further clarity as to whether incidents of BH differed significantly between those accessions. This is particularly important given that some of the accessions studied within this research are Elsoms breeding lines and elite hybrids (with the hybrids currently provided to farmers for growth in the UK market). It is therefore important that there is empirical evidence to show whether certain accessions exhibit reduced BH incidents.

Table 5.9 Total proportion of BH affected roots

The severity of BH within accessions was scored based on a 4-point system: 0 = clear, 1 = slight, 2 = moderate 3 = severe. For each replicate, the symptom incidence for each individual root was scored as 0.25 (out of 3.0) for each accession and treatment individually. Total symptom incidence was then calculated by adding the proportion of each symptom observed and subsequently dividing this value by the total number of symptoms. Red - blue colour scale represents relative high and low proportional incidence respectively.

Accession	Proportion of total BH affected roots
Mg	0.42
Lz	0.27
He	0.50
Ac	0.08
Ag	0.58
Sl	0.58
Tk	0.00
Mg x He	0.17
He x Ac	0.08
Ag x He	0.58
G1*He	0.18
Bh	0.25
Tn	1.00
Kf	0.17

The second aim of this trial was to assess how supplementation with B affected the presence and severity of BH symptoms. Whilst total incidents of BH did decrease with increasing H_3BO_3 supplementation, this reduction was not found to be significant ($p = 0.1004$) (Table 5.3). Furthermore, there was no significant difference in root B concentrations in BH affected and clean roots (Table 5.5). For this analysis, ideally data should have been split according to severity (and not just whether BH was present); however there was not a sufficient proportion of roots affected for this to be viable. Again, analysing fewer accessions would have

been beneficial here and would have allowed for both the severity of symptoms and root B concentrations to be characterised across each accession individually (rather than grouping all roots affected with BH and comparing with clean roots). Leaf B concentrations were also analysed at four time-points, with this looking to assess whether leaf B concentrations could be used as an indicator of BH presence. Utilising leaf B concentrations as a potential indicator of BH would be highly advantageous given the destructive nature of assessing symptom presence within the roots and that assessment can only be conducted post-harvest. However, there was no significant difference in leaf B concentrations of the respective roots affected with BH as compared to clean roots at any of the four time points (Table 5.6).

Data obtained from this research disputes results identified by Fadhel et al (2013), who saw reduced B concentrations in roots affected with BH. However, as discussed within this chapters' introduction, specific statistical analysis was unfortunately not reported in this study. Furthermore, Sanderson, Sanderson and Gupta, (2002) did not analyse B concentrations in roots and consequently did not report whether B concentrations were significantly lower in BH affected individuals. This was surprising given leaf B concentrations were quantified. However, both of these studies were conducted in the field, which makes direct comparisons with this research difficult. Furthermore, the method in which B was supplied to the plants differed in these studies (pre-plant B treatments and subsequent foliar spraying) as compared to this research (nutrient solution directly to terra-green), which again limits direct comparisons. The mode of B delivery to the plant is of importance and warrants further consideration. For example, it is highly likely that the availability of B within terra-green was different to that normally exhibited in soil, which is impacted by various environmental factors (such as pH, soil structure and water availability; Goldberg, (1997)). Terra-green cannot hold water and therefore frequent watering (twice daily) was required, which may have affected nutrient uptake. Sanderson, Sanderson and Gupta, (2002), for example, report that lack of soil moisture during root development may impact upon B uptake and subsequent BH control. Given that the terra-green was kept constantly moist by frequent watering, B uptake should not have been limited. Subsequent attempts to assess the nutrient concentrations within terra-green (via ICP-MS) were unfortunately unsuccessful as the material

itself could not be digested in nitric acid. Given this, and the fact terra-green is marketed as an inert medium, one can assume that the only elements available came from the nutrient solution and any trace elemental concentrations in the water. The impact of trace elements in the water provided to the plants is highly significant. As discussed previously, whilst deionised water was used to dilute the nutrient solution, this option was not available for plant watering at Elsoms. Specific concentrations of H_3BO_3 are not reported by Anglian Water Services (the Lincolnshire water provider), however the maximum permissible availability of B in UK ground water is 1.0 mg/L (UK water supply regulations; Drinking Water Inspectorate, 2018) and resultantly the concentrations of B supplied to the plant are likely to have been greater than the desired treatments. This poses one of the key flaws within this experimental design and in hindsight specific concentrations of H_3BO_3 within the water should have been quantified (through ICP-MS analysis) and subsequently accounted for (e.g. by altering concentrations of B in the nutrient solution).

Additionally, the mechanisms of nutrient supply is important. For example, if B is applied as a foliar spray there is then the question of how it is subsequently transported to growing tissues. In a study by Shelp and Shattuck, (1987), foliar spraying increased B concentrations in the roots, with phloem mobility assessed through ratios of B in the root as compared to mature leaves. Of the two accessions investigated by Shelp and Shattuck, (1987), one was thought to exhibit a greater capacity for B retranslocation (via the phloem) to the roots, and was resultantly less sensitive to BH. This fact is highly interesting given that empirical evidence documenting phloem mobility of B in *Brassica* species is limited. Unfortunately mature leaves were not sampled within this study and therefore B concentrations and subsequent capacities for phloem mobility was not quantified. However, leaf B concentrations in low H_3BO_3 treated plants were significantly higher than high treated plants ($p < 0.05$), with this result potentially validating what Shelp and Shattuck (1987) observed. For example, B translocation from mature to young tissue, under B limitation, may potentially be a mechanism to minimise B deficiencies within growing tissues (i.e. young leaves). The capacity for phloem transport should be considered in future studies assessing the impact of B (or indeed concentrations of other elements) on BH symptoms. Ratios of elements in different tissues could be used to simply assess

this, or, conversely, following methods outlined in Huang, Bell and Dell., (2008) (and discussed in Section 1.3.1.6), utilising hydroponics with different elemental isotopes.

The specific root sampling methods is also of importance when looking to assess the impact of B on BH symptoms. Root sampling within this research followed methods outlined by Fadhel et al (2013), from which central sub-sections of the root (containing both BH affected and clean tissue) was sampled and subsequently homogenised. However, it would be of interest to see how B concentrations varied between BH affected and clean tissue taken from a single root. For this, the root would have to be carefully dissected to include either BH affected or clean tissue and subsequently homogenised before elemental analyses. This may further elucidate whether localised differences in B concentrations within the root are observed (and whether homogenised material containing both affected and unaffected tissue, as utilised within this research, resulted in no significant differences being observed). However, field trials conducted by Elsoms similarly identified no significant differences in B concentrations between BH affected and clean roots (unpublished data). This therefore validates this research's results and strengthens the argument that there are additional factors underlying BH.

5.5 Summary and conclusions

The aim of the research conducted for this trial was to gain further insight into the swede disorder Brown Heart. The experimental design for this trial enabled tighter control over factors which could affect BH formation. No significant differences in BH incidents were observed between the three H_3BO_3 treatments. Furthermore, no results link B concentrations within the root or leaves to the presence of BH. However, the number of plants analysed for each treatment may have limited the scope for identifying statistically significant differences. Focusing on a smaller number of accessions and scoring more roots for BH presence would be pertinent. Factoring in trace B concentrations in the water supplied to the plants is also essential, as it is acknowledged that, once included, this amount was likely greater than the desired B treatments. However, the fact that B concentrations did not differ between roots affected with BH and clean roots indicates that other

factors are affecting the development of this disorder. Further research is required to elucidate these factors, with discussion on potential areas for future study provided in this thesis' general discussion within Section 6.3.

Chapter 6: General Discussion

6.1 Utilising AT to further understand the genetic control of B concentrations in *B. napus* leaves and seeds

The approach adopted for this research was associative transcriptomics, a quantitative genomics technique which looked to associate how variation in gene sequence and expression, across a diverse *B. napus* panel, affected trait variation. Through AT the primary hypothesis was tested; whereby this research aimed to elucidate whether variation in B concentrations in *B. napus* leaves and seeds is a result of underlying genetic variation. A diversity panel of 383 *B. napus* accessions, assembled by the Renewable Industrial Products from Rapeseed consortium (Thomas et al., 2016; Havlickova et al., 2018), was utilised throughout this research, with AT subsequently employed to identify genomic regions controlling B concentrations.

Understanding the genetic basis of complex traits is highly advantageous, for example when looking to conduct breeding programmes to enhance desirable traits (see Section 1.3.3.2). Association mapping studies, in comparison with more traditional QTL techniques, allow for a much smaller genomic window to be searched for candidate loci which control such traits. However, this then requires a reference sequence in order to score sequence polymorphisms against. As such, AT was originally developed as a means to analyse the association between sequence and trait variation in crops, where the assembly of genome sequences is complex due to the prevalence polyploidy (Harper et al., 2012). AT utilises transcriptome sequencing (rather than conventional genomic data typically utilised in GWAS studies), based on the methods outlined in Trick et al., (2009b), with its implementation ideal in species lacking complete reference sequences. However, published genome sequences for the *B. napus* cultivar Darmor-*bzh* have subsequently become available (Chalhoub et al., 2014) and recently refined (Bayer et al., 2017). Bayer et al., (2017) also presented *de novo* assembly of a second *B. napus* cultivar; Tapidor. Comparisons of Darmor-*bzh* and Tapidor identified genes unique to each cultivar, highlighting that a single reference is insufficient to fully represent gene content within complex crop species (Bayer et al., 2017). However, whilst these reference sequences are

available for association mapping studies (for example Bus et al., (2014) used Darmor-*bzh* as a reference genome for SNP discovery), the AC pan-transcriptome utilised within this research actually identified a greater number of CDS models than those for the Darmor-*bzh* reference (He et al., 2015); 116,098 were in the AC pan-transcriptome, as compared to 80,927 reported by Chalhoub et al., (2014). Furthermore, the SNP density in the AC reference is much higher than the density of the commercially available 60K Brassica Infinium® SNP arrays (Havlickova et al., 2018); 355,536 were scored from this study as compared to 21,117 in by Xu et al., (2016) and 25,183 in Wang et al., (2014). However, given AT utilities transcriptome data (i.e. the CDS), sequence variation in non-coding regions is unavailable; presenting one of the main limitations of the AT technique. Conversely, using transcribed sequences brings a major benefit to AT as compared to GWAS, in that transcript variation (as GEMs) can be quantified. Significant expression for 53,889 CDS models was identified within the AC pan-transcriptome, representing 46.4% of all gene models (Havlickova et al., 2018). The AC pan-transcriptome database also included *A. thaliana* orthologue annotations for *B. napus* markers, with this fundamentally important when looking to identify candidate genes. For example, it enabled the use of the PiiMS database to see if any candidates had previously documented ionomic profiles in *A. thaliana* T-DNA insertion lines. Furthermore, functional annotations of candidates could be viewed, for example, through TAIR. However, not all *B. napus* markers had an *A. thaliana* annotation and therefore these genes could only be analysed *in silico* via trait predictions. Some of the most highly associated GEMs in both leaf and seed B AT were unannotated. For example, Cab004528.1, on chromosome A3, was one of the most highly associated markers and was significantly predictive for leaf B concentrations. Unannotated markers within SNP association peaks were also observed. Whilst the CDS of these markers could be compared against other species (through an in-house script and BLAST; see Section 2.2.3); in the majority of cases no *A. thaliana* orthologue could be identified. Therefore, it is possible that candidates controlling the trait under investigation were not identified. Consequently, it would be worthwhile to consider further study of a selection of unannotated candidates, with a particular focus on the highly associated GEMs. This could be conducted using TILLING or

other viable mutagenesis methods in *B. napus* (see Section 3.3.3) if no *A. thaliana* orthologue could be identified.

Both Harper et al., (2012) and Havlickova et al., (2018) were able to combine the power of both SNP and GEM analyses to identify loci controlling seed glucosinolate and tocopherol variation respectively. Alcock et al., (2018) similarly identified a potassium transporter in GEM analyses for leaf K concentrations. However, within this research, few candidates were identified which were obviously controlling B concentrations in either leaves or seeds (i.e. known B transporters). As discussed previously in Section 3.3.3, this may be a result of the experimental design used for the AT panel. For example, the leaf transcriptome was sampled at the 2nd true leaf stage, with the plants grown under nutrient sufficiency. Resultantly, known genes involved in the control of B accumulation (such as orthologues of *AtNIP5;1* which is primarily expressed during B-deficiency and in roots; *c.f.* Takano et al., (2006)) may not have been identified due to these aforementioned limitations in experimental design. This point is also pertinent when considering many of the candidate genes identified within past QTL studies, as reviewed in Section 1.3.1.9, identified orthologues of *AtNIP5;1* in their studies. For example, fine-mapping by Hua et al., (2016a) identified *BnaA03.NIP5;1* as a gene underlying a QTL for 'B-efficiency coefficient' (which was defined as the ratio of plant dry weight in low B to that in high B conditions). However, given that many of the past QTL studies have focused on identifying efficiency mechanism under B-deficiency, rather than looking for loci controlling B accumulation under nutrient-sufficiency, it is perhaps unsurprising these loci were not identified during AT analyses. Considering this, growing the panel under different nutritional statuses would be an asset, as discussed below.

As further discussed in Section 3.3.3, growing plants under B-deficiency and toxicity, and not just nutrient sufficiency, may have yielded more candidates from AT analyses. One way this could have been conducted is through hydroponics, which would have enabled tighter control over nutrient availability. Increased stringency over growth conditions may also be pertinent, given experimental design accounted for 42% and 40% of the variance in leaf and seed B concentrations respectively (Thomas et al., 2016), as determined by variance components analysis. This analysis also indicated that the effect of experimental

variance was generally higher for micronutrients than macronutrients. This may be due to the lower requirements of micronutrients (as compared to macronutrients) during plant growth, which may resultantly have a compounding affect if elemental concentrations are perturbed due to experimental design and conditions. Specifically, as the plants were grown in 5 L pots containing soil, it would be difficult to ensure that the availability of nutrients was highly consistent across individuals. Various factors could impact this, such as localised differences in soil structure or water availability. Utilising hydroponics would therefore also be beneficial here (in addition to allowing changes to nutrient concentrations as mentioned previously). This would, however, pose a highly complex task given the size of the diversity panel (2160 individual plants) and the high number of traits scored for the RIPR collaboration (i.e. during phenotyping there remains the potentiality of the ionomic profile indirectly affecting other traits under investigation). Terra-green (as used for this thesis' BH research) could also feasibly be utilised, which would allow for tighter control over elemental concentrations provided to the panel. Furthermore, both hydroponics and terra-green would also have enabled root traits (e.g. the root transcriptome) to be analysed more easily than when using soil as a growth medium. Sequence and transcript quantification from root tissue would be a major asset to this research, especially given many of the processes initially controlling B accumulation are applied in the root under both B-deficiency (e.g. *BnaNIP5;1* upregulation; Yuan et al., (2017)) and toxicity (e.g. repression of *HvNIP2;1*; Schnurbusch et al., (2010)), whilst more generally root traits are important for nutrient acquisition (White et al., 2013).

Multiple techniques were employed to identify loci controlling B concentrations in leaves and seeds. Trait predictions were performed using the split training and test AT panels, containing 274 and 109 accessions respectively. SNP predictions, for example, looked to assess how allelic variation affected B concentrations and subsequently if markers were predictive for B phenotype (i.e. if the allelic variant present in particular accessions was predicative of B concentrations in other accessions with that same allele). A maximal fold difference of up to 1.80 and 1.19 in B concentrations between allelic variants was observed for leaf and seed respectively, indicative of variation within the panel and the scope for uncovering genetic loci controlling B accumulation. However,

whilst predictive markers were observed from both leaf and seed, many predictions using the split training and test panels were not significant. This is potentially due to the technique of splitting the diversity panel into two smaller panels, which could reduce the likelihood of identifying rare allelic variants and therefore significant predictions (i.e. there may not be sufficient diversity to identify significant predictions if the particular allele is only present in a small number of accessions). A 'take one out' approach (Harper et al., 2016) would have been a suitable alternative and could have been conducted on the full 383 diversity panel rather than using the training and test panels. This technique looks to perform trait predictions by removing each accession in turn, with AT conducted for the trait of interest and the phenotype of the missing accession then estimated. This technique would also have been beneficial when performing GEM predictions, as AT analyses performed on the 383 and 274 panels did exhibit slight variations (i.e. some markers were more highly associated within the 274 relative to the same marker within the 383 panel or *vice versa*). Additionally, the lack of significant markers passing the Bonferroni and FDR corrected significance thresholds requires further consideration, as discussed in within Section 4.3.3.

Despite the aforementioned limitations of AT as a technique for this research, SNP and GEM outputs were successfully analysed for the identification of candidate genes controlling leaf and seed B concentrations. This research identified a number of previously characterised candidates, such as orthologues of MYB DOMAIN PROTEIN 36 (*MYB36*; AT5G57620.1), a regulator of Casparian strip development (Kamiya et al., 2015) and NOD26-LIKE INTRINSIC PROTEIN 6;1 (*NIP6;1*; AT1G80760.1) a B channel protein (Tanaka et al., 2008) (although both were observed in minor association peaks). Candidates with previous characterisation within the literature, or ionomic profiles available of PiiMS, were not further validated within the scope of this research. As such, only novel candidates were analysed further. Twenty-six candidate genes, orthologous to thirteen genes in *A. thaliana,* were characterised using T-DNA lines. Disruption in B concentrations in seven of these lines as compared to WT controls was observed. One of the most promising candidates identified from AT analysis, which subsequently exhibited disruptions in B concentrations in the T-DNA line as compared to the WT control, is *BnaNIP2;1* (Figure 4.3 and Figure 4.4).

Interestingly, recent experimental evidence has demonstrated that certain isoforms exhibit functional B transport when expressed in *Xenopus* oocytes (Diehn et al., 2019), whilst other NIP AQPs act as B transporters (Takano et al., 2006; Tanaka et al., 2008; Li et al., 2011). As discussed in detail within Section 4.3.1.1, further characterisation of this candidate should be conducted, particularly in assessing whether the different isoforms have functional B transport capacities *in planta.* Furthermore, clarity on expression profiles of *BnaNIP2;1* is essential. For example, past research has shown root-specific expression of *AtNIP2;1/ BnaNIP2;1* (Mizutani et al., 2006; Yuan et al., 2017), however expression profiles obtained for *AtNIP2;1* from Genevestigator (Hruz et al., 2008) show that it is actually expressed in many tissues, not just within roots (Figure 4.10). Additionally, significant expression for one *BnaNIP2;1* copy, Cab033878.1, was identified within the leaf pan-transcriptome (no significant expression for the other three copies was detected), which could indicate divergent gene functions between the homoeologues (such as differences in relative expression in different tissues), as is observed in species exhibiting polyploidy (Adams et al., 2003). Further insight into how this gene is affecting B accumulation could also be obtained by assessing difference in expression (for the four gene copies) under B deficiency or toxicity (i.e. is there upregulation or repression under different nutritional statuses, and in which tissues does this occur). For example, Diehn et al., (2019) reported expression for different *BnaNIP3;1* isoforms was tissue dependent, with upregulation and downregulation observed in different tissues under B-deficiency (Section 1.3.1.6). However, prior to conducting any further research in *B. napus*, screening of additional T-DNA lines is essential. This should be conducted for all candidates showing B disruption identified within this research. This is particularly pertinent given the majority of candidates were only analysed using a single T-DNA line. Furthermore discrepancies between T-DNA lines for *xeg113a* and *xeg113b* were observed (B was not significantly perturbed in both lines), whilst *nip2;1* (analysed in leaf for this research) showed a significant decrease in B concentrations, whereas on PiiMS a significant increase was reported (see Section 4.3.1.2). Transcript quantification for all T-DNA lines should also be assessed to ensure knock-out of the particular gene under investigation. Further discussion on the use of T-DNA

lines for ionome analysis, and the need for repetition of experiments (both within this research and the wider literature), is provided in Sections 3.3.1.2 and 4.3.1.2.

6.2 Analysis of leaf and seed nutrient concentrations within the RIPR diversity panel: boron and beyond

Variation in elemental concentrations can occur between cellular locations, accessions, crop habits and species under investigation. Furthermore, environment and growth conditions can have a large impact on the ionomic profile of the organisms in question (Baxter, 2015). Resultantly, species wide variation in elemental concentrations, as was available for this research as part of the RIPR collaboration (Thomas et al., 2016), was a great asset. Not only could B concentrations be utilised for AT analyses, but relationships between B and other elements could be further interrogated.

Whilst the focus of this research was assessing factors controlling B concentrations, co-localisations of B association peaks with additional elements was observed. Past research has also indicated this, for example Liu et al., (2009) found co-localisations of QTLs for B/Cu and B/P concentrations in a *B. napus* mapping population. Within this research, loci for B and Mn concentrations co-localised in leaf (Section 3.3.2.4), with one association peak passing the Bonferroni corrected significance threshold containing a known Mn transporter (*MTP8*; Eroglu et al., (2016)). Furthermore in seed, shared association peaks for B, Mn and As concentrations were observed on chromosome A5, which contained orthologues of *AtNIP2;1*. Subsequent characterisation of this candidate using T-DNA lines identified significant perturbations of each of these elements (Figure 4.3). Shared association peaks for B and S concentrations in seed was also observed. The seed B/S association peaks contained orthologues of *AtHAG1*; an aliphatic GSL regulator. Whilst it is not clear why seed B should exhibit these association peaks, the breeding for low-GSL accessions may have indirectly selected for other additional seed traits. However, the *HAG1* GEM Cab038298.3 was predictive for B concentrations, indicating some form of (direct or indirect) relationship. This relationship between GSL/*HAG1* and the wider ionome was investigated by another member of the Bancroft research group

(Sweeney, 2018; PhD thesis). Furthermore, the association between the flowering time regulators *FLC* and *SOC1* and concentrations of additional micronutrients was observed in this research by Sweeney, (2018), whilst the association between macronutrient concentrations and flowering time was also reported by Alcock et al., (2017), (2018). These points further highlight the relationship between pair or groups of elements (and potential underlying factors affecting these relationships). This interdependence indicates that studying each element independently is not sufficient when looking to elucidate mechanisms underlying the control of the plant ionome. As such, further detail on relationships between B and additional elements has been provided hereafter.

Seed B and leaf B concentrations were not significantly correlated ($p = 0.089$), with this in concordance with elemental correlations between tissue types across the RIPR diversity panel (Thomas et al., 2016). Baxter et al., (2012) also identified limited leaf/seed (or leaf/root) elemental correlations in 96 *A. thaliana* accessions, surmising that when studying a particular ionomic trait (e.g. the seed ionome), that specific tissue should be used for future genetic screening. Furthermore, if looking to assess nutrient deficiencies in a particular tissue, that specific tissue should be assessed. This is particularly relevant when considering this thesis' BH research, which looked to asses if leaf B concentrations could be used as an indicator of BH incidents (see Sections 5.3.3 and 6.3). Other studies, such as Ghandilyan et al., (2009), also report limited elemental correlations between pairs of tissues.

Within tissues, B was significantly correlated with a number of other essential elements from the RIPR ionomic datasets. In leaf, B was significantly correlated with nine out of the eleven essential elements analysed via ICP-MS; Ca, Cu, Fe, K, Mg, Mn, Mo, P, and S were all positively correlated (Table 3.2). In seed, B was also significantly correlated with nine of the essential elements; Ca, Cu, Fe, Mn, Ni, P, S, Zn were negatively correlated, whereas Mo was positively correlated (Table 4.1). B was also significantly positively correlated ($p < 0.001$) with As in both leaf and seed, which is unsurprising given these elements can share transport mechanisms (Bienert et al., 2008; Isayenkov and Maathuis, 2008). Multiple studies in *B. napus* have reported significant correlations between elements. Bus et al., (2014), for example, studied ionomic traits in a diversity

panel containing 509 accessions, with B concentrations significantly positively correlated with the majority of the elements also studied for this research (only P was not significantly correlated and Ni not reported). However, it should be noted that whole shoot material was analysed from this panel, not just leaf material. Baxter et al., (2012) also identified multiple significant leaf elemental correlations between 96 *A. thaliana* accessions grown in two soil experiments. However, correlations did differ between experiments; for example in *experiment 1* B was positively correlated with Li and negatively correlated with Cu, whilst in *experiment 2* B was positively correlated with Li and Mg, and negatively correlated with Fe and P (Li correlations cannot be compared as Li was not quantified for the RIPR panel). However, correlations were only reported with R^2 >0.32 and p <0.001, and therefore additional weaker correlations may have been observed. Within this same study, seed correlations were investigated, however B was not significantly correlated with any other element, although again only R^2 >0.32 was reported. This again further highlights the difficulties in analysing the ionome and that repetition of experiments is essential in order to further elucidate patterns within the data.

B concentrations across the RIPR panel varied widely, with 5.40- fold variation in leaf (Figure 3.1), and 3.38-fold in seed (Figure 4.1). The higher levels of variation in leaf B concentrations as compared to seed is not necessarily surprising given the higher estimates of heritability for leaf B; 22% of the variation in leaf B was attributed to genotype as compared to 13% for seed (Thomas et al., 2016). Variation in B concentrations in *A. thaliana* accessions, reported by Baxter et al., (2012), were comparable to this study (seed: 4.57- fold; leaf: 5.56- fold). However Liu et al., (2009) only identified 2.48- fold variation in B concentrations in *B. napus* leaves (although this was a QTL mapping population containing 162 lines formed from a cross between two accessions). Shoot B concentrations in the *B. napus* panel collated by Bus et al., (2014) exhibited ~ 4.00-fold variation. Furthermore, the heritability of shoot B concentration was estimated to be 0.74, whilst the percentage of phenotypic variation that could be attributed to population structure (as determined by r^2), was 1.57%. To the authors knowledge, no species wide variation in B concentrations for seed in other *Brassica* diversity panels have been reported.

Levels of variation for other elements within the RIPR panel fluctuated widely (Table 6.1). In leaf, for example, Ni exhibited the least variation, with 1.58-fold difference between the lowest and highest concentrations, whilst Mo exhibited the highest variation at 14.18-fold. P variation was low in both leaf and seed, whilst Mn variation was high in both tissues. These trends follow fold-changes in the Baxter et al., (2012) *A. thaliana* diversity panel; for example P exhibited the lowest variability (1.46-fold) in seed concentrations in all 19 elements quantified.

Table 6.1 Fold change in leaf and seed elemental concentrations across the RIPR diversity panel

Fold change in elemental concentrations (mg/kg) for 12 essential elements within the RIPR diversity panel was calculated for leaf (n = 385) and seed (n = 380).

	B	Ca	Cu	Fe	K	Mg	Mn	Mo	Ni	P	S	Zn
Leaf	5.40	3.21	2.82	2.12	2.11	2.62	8.09	14.18	1.58	1.99	2.55	2.97
Seed	3.38	3.06	2.75	3.02	2.04	2.09	5.76	6.03	3.11	1.68	7.49	2.73

The net allocation of B in either leaves or seeds can be demonstrated by the ratio of the concentrations in different tissues; whereby seed B: leaf B > 1 indicates a net increase within the seed (< 1 indicates net increase in leaves). As shown in Figure 6.1, there was a net increase of B within seeds for all crop types. This is unsurprising given the higher requirement for B during reproductive growth (Dell and Huang, 1997) and that the inflorescence act as a B sink (Pommerrenig et al., 2018). The seed accumulation of B in winter OSR was higher than for spring OSR ($p < 0.01$), which may indicate mechanisms for improved allocation in this crop type. Furthermore, differences in nutrient translocation for essential and non-essential elements (and differences between crop types) can also be compared using this method. For this, ratios of essential nutrients in seed and leaves can be compared with ratios for non-essential elements which exhibit similar transport or assimilation pathways, as outlined in Thomas et al., (2016). This is pertinent for the relationship between leaf and seed B/As, which can share NIP transporters (Bienert et al., 2008; Isayenkov and Maathuis, 2008).

$[B]_{seed}/[As]_{seed}$)/ $[B]_{leaf}/[As]_{leaf}$ > 1 indicates seed accumulation of the essential B as compared to the non-essential As, with the relative net accumulation of B in seeds as compared to As observed for all crop types (Figure 6.2) (as was expected given this is the essential element under investigation).

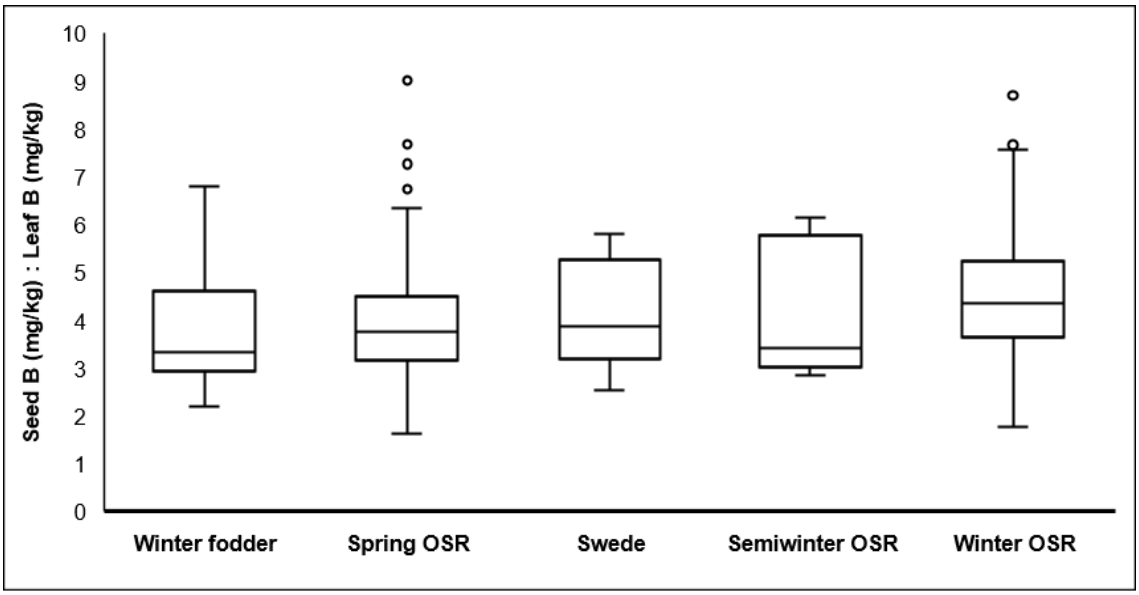


Figure 6.1 Seed: Leaf ratios of B concentrations for B. napus crop types

Seed: leaf ratios were calculated for 338 B. napus accessions within the RIPR diversity panel whose crop types was assigned; 15 winter fodder, 128 spring OSR, 31 swede, 7 semiwinter OSR and 159 winter OSR. Boxes represent the first and third quartiles with the median drawn. The whiskers represent the lowest and highest data points < 1.5 times the upper and lower quartiles. Circles represent outliers > 1.5 times the upper and lower quartiles. Boxes are organised by mean ratio from lowest (left) to highest (right). Method adapted from Thomas et al., (2016).

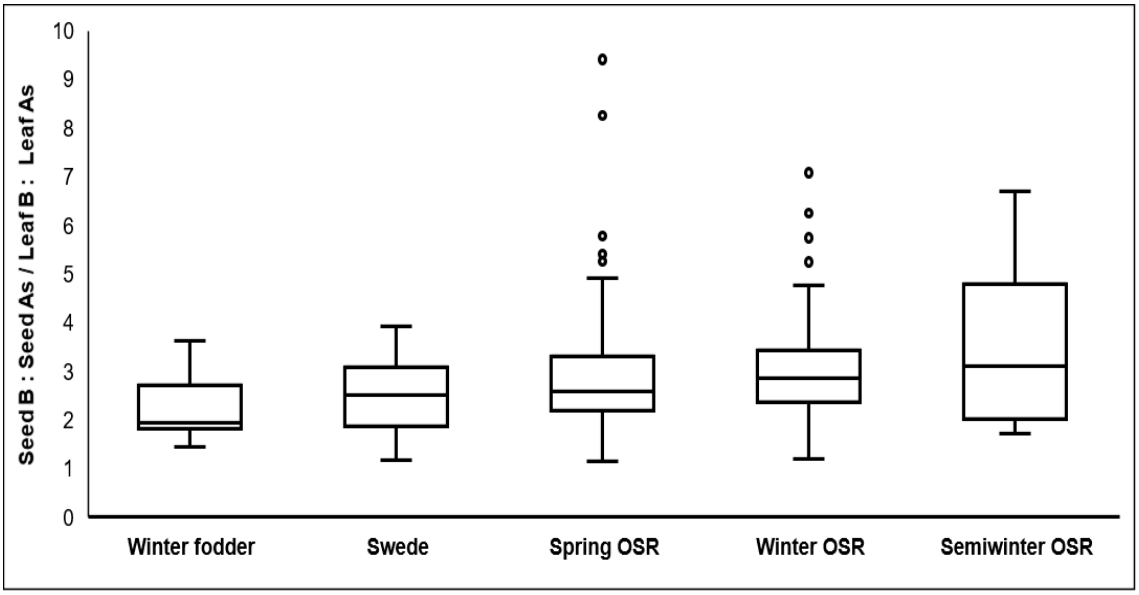


Figure 6.2 Leaf and seed ratios for B and As concentrations

([B] seed/[As] seed) / ([B] leaf/[As] leaf) ratios were calculated for 338 B. napus accessions within the RIPR diversity panel whose crop types was assigned; 15 winter fodder, 128 spring OSR, 31 swede, 7 semiwinter OSR and 159 winter OSR. Boxes represent the first and third quartiles with the median drawn. The whiskers represent the lowest and highest data points < 1.5 times the upper and lower quartiles. Circles represent outliers > 1.5 times the upper and lower quartiles. Boxes are organised by mean ratio from lowest (left) to highest (right). Method adapted from Thomas et al., (2016).

6.3 Understanding the control of Brown Heart in swedes: future research directions

The industrial collaborators for this research, Elsoms Seeds Ltd, conduct the only swede breeding programme in the UK. Resultantly, an intention to assess potential factors controlling the disorder known as Brown Heart, which is currently attributed to localised B-deficiency within the swede root, remains a key area of investigation. Research by another collaborator of Elsoms Seeds identified that incidence of BH within their parent lines was associated with swede accession (Fadhel et al., 2013). In concordance with previous research, this seemingly indicated that certain accession had increased BH susceptibilities (Shelp and Shattuck, 1987; Sanderson, Sanderson and Gupta, 2002). These studies also demonstrate, to varying extents, that supplementation with B can mitigate BH symptoms. However, anecdotal evidence from farmers still report that BH is

present within their crops, despite using ‘BH tolerant’ accessions and extensive B fertilisation. As such, to gain further insight into the basis of the disorder, this research designed a trial to specifically assess the impact of swede accession and B supplementation on the presence and severity of BH. As discussed within Section 5.4.1, whilst the proportion of roots affected with BH did exhibit some variation between accessions, there was no significant difference in the number of BH affected roots between previously determined susceptible or resistant accessions (Appendix 5.1). Furthermore, whilst plants treated with high concentrations of H_3BO_3 did exhibit fewer incidents of BH than plants treated with low H_3BO_3, this reduction was not statistically significant (Table 5.3). Perhaps most remarkably, there was no significant difference in B concentrations between BH affected and clean roots (Table 5.5).

Given the results obtained from this trial, it seems appropriate to consider other potential factors which could be underlying this disorder. Of the points put forward for consideration by other researchers, almost all relate to how they impact upon B uptake and utilisation, rather than novel causes. For example, Sanderson, Sanderson and Gupta, (2002) suggest that certain soil K concentrations can negatively impact B uptake. Specific K concentrations were not explicitly stated in this study, however Hadas and Hagin, (1972) report that increasing K concentrations in soils with low B can exacerbate B-deficiency symptoms. Within this thesis’ research, K concentrations were constant across treatments, whilst no correlation between root B and K concentrations was observed (Appendix 5.11). As reported by Gupta, (1993), soil moisture has also been suggested to be an important factor for consideration; with Sanderson, Sanderson and Gupta, (2002) proposing increased BH incidents within years of low rainfall was a result of reduced B uptake. As discussed in greater detail within Section 5.4.1, as terra-green was used within this research, frequent watering was a necessity. Although B uptake itself was not specifically assessed for this research (for uptake to be feasibly quantified hydroponics should be utilised; see below), however if assuming that B availability is impacting BH incidents, then this could explain why relatively few roots exhibited BH even in low H_3BO_3 treated plants (i.e. even though B concentrations were low they were readily available). However, this would not explain why B concentrations did not differ between BH affected and clean roots, nor why BH severity and B concentrations were not significantly

correlated. Interestingly Mg, Ca and Mn concentrations were all positively correlated with BH (Table 5.7). Furthermore, Ca and Mn were the only elements that differed significantly in roots affected with BH and clean roots (both elemental concentrations were higher in BH affected roots) (Table 5.8). What effect these elemental concentrations may play on BH development is currently unclear. From previous research conducted into BH, there is limited reference to additional elemental interactions. Shelp and Shattuck, (1987) observed that root concentrations of Mg, Mn and Zn accumulated under B deficiency. Whilst definitive reasons for this was not clear, it was hypothesised that changes to water flux or cell wall structure may have ultimately impacted elemental accumulation. However, there were no suggestions that concentrations of these elements affected BH symptoms. Shelp and Shattuck, (1987) also assessed if ratios of Ca and B concentrations could be used as an indicator of B status (i.e. deficiency or sufficiency), as was reported in earlier work by Beauchamp and Hussain, (1974), who found that Ca/B >170 mg/kg was associated with increased BH incidents. Shelp and Shattuck, (1987) however, did not observe any relationship between Ca/B ratios and BH incident. Within this thesis' research Ca/B ratios did not differ significantly between BH affected and clean roots ($p = 0.427$). However, it is interesting that Ca concentrations were significantly higher in BH affected roots ($p = 0.001$) (Table 5.8). It is possible that the concentrations of Ca are impacting on cell wall profiles in BH affected roots, given that one of the roles Ca plays *in planta* the cross-linkage of pectic residues in the cell wall (reviewed in Helper, (2005)). However, it is not immediately obvious why these higher concentrations of Ca are associated with roots affected with BH (and if the two are directly linked). As such, further research is required to elucidate potential mechanisms. It would therefore be useful to assess the specific cell wall structure and cell organisation in BH affected tissues. For example, one hypothesis might investigate if the cells are bursting, which could subsequently lead to oxidative damage resulting in tissues appearing brown. Imaging these cells through microscopy would be an appropriate successive method in assessing cell wall properties, which may further elucidate factors affecting BH. At the very least, this would assess what affect BH presence has on the structural integrity of the cell.

Of the results obtained from this research there are various opportunities for future research, some of which have previously been discussed. Whilst B

concentrations did not differ between BH affected and clean roots (indicating additional factors controlling the disorder), given the limitations of this research it would be pertinent to enact tighter control of B concentrations in precisely supplementing B, given the trace concentrations within ground water. This could be conducted using terra-green and deionised water, or alternatively hydroponics. Furthermore, focussing on a smaller number of accessions (specifically the hybrids supplied to farmers or some key breeding lines) should be considered. This may allow for greater potentiality in identifying significant differences between accessions, as the scope of this research was somewhat limited through only analysing four replicates per line. This would also control the scale of the experiment, which would be particularly important if hydroponics was utilised, given the complexities of the technique and the length of growing period prior to analysis (~ 5 months prior to root harvest). It would also be interesting to assess the impact of a selection of additional elements on BH incidents, particularly Ca concentrations, given this research identified positive correlations between BH symptoms and Mg, Ca and Mn root concentrations, and that Ca and Mn concentrations were significantly higher in BH affected roots. Once again, either terra-green or hydroponics could also be utilised here.

However, the scope for comparison of such research with subsequent field-trial assessments should be considered, given the multitude of environmental factors which can affect plant growth within the field (i.e. can such research be translated into actionable mechanisms to control BH in the field). Despite this, greater insight into factors affecting BH may allow for future recommendations of best-practice, which could positively impact BH control in the years to come.

6.4 Summary of thesis content

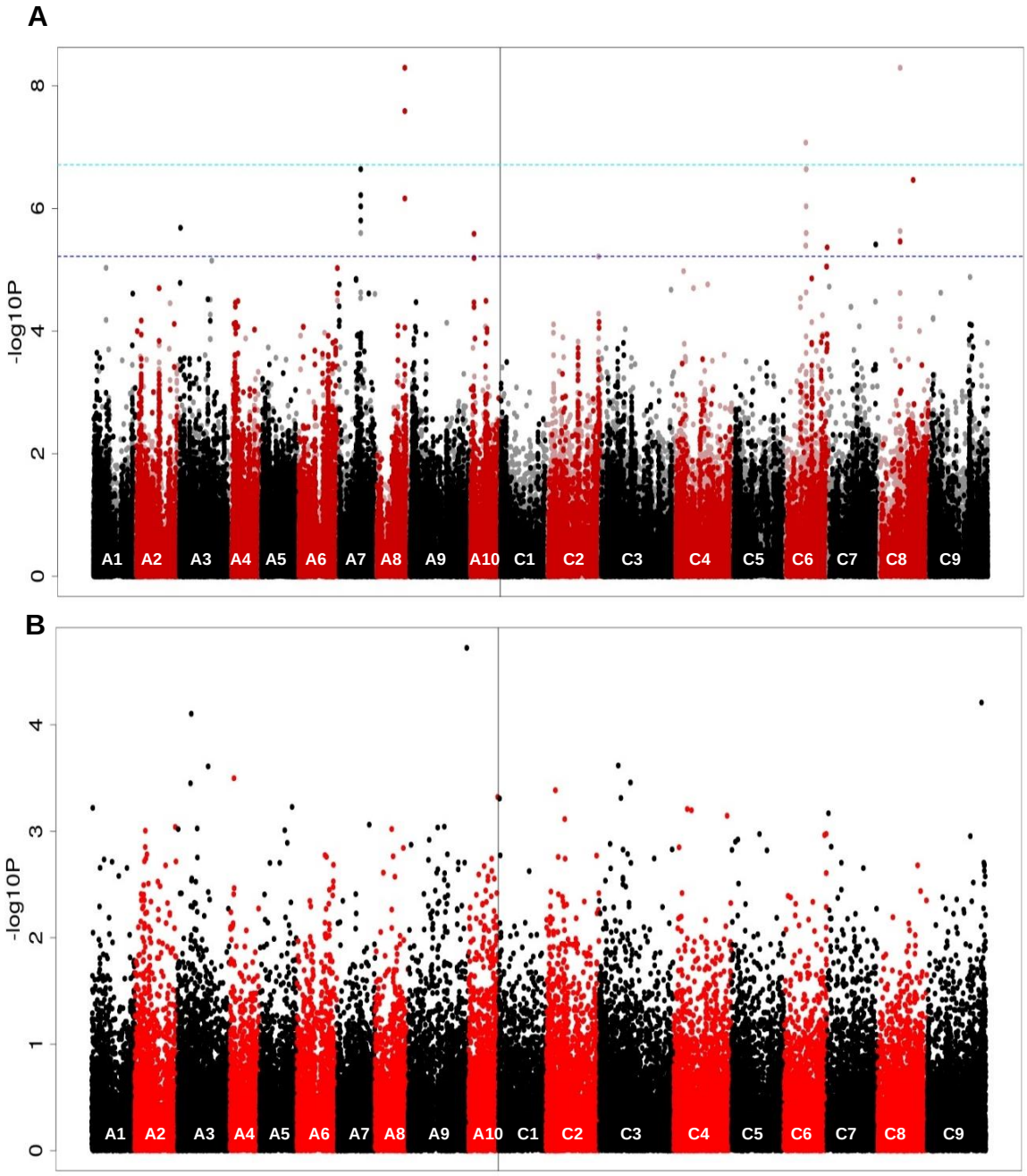
To summarise, this research has focused on two aspects relating to B concentrations in *B. napus*. The importance of B within swede roots was analysed as part of a trial into the physiological disorder Brown Heart. No relationship with B supplementation and the presence of BH was identified. Furthermore, B concentrations did not exhibit any significant variation in roots exhibiting BH with those clear of symptoms. Resultantly, research into additional factors attributed to this disorder is required. Secondly, associative transcriptomics was performed

to identify loci controlling B concentrations within a diverse panel of *B. napus* leaves and seeds. Allelic variation within the diversity panel was indicative of variation in B concentrations within both leaves and seeds, demonstrating significant scope for improving B acquisition. Candidate genes identified from AT were subsequently analysed using *A. thaliana* T-DNA insertion lines, with seven candidates exhibiting altered B concentrations relative to WT controls. Further validation of these genes, for example using TILLIING, is required to confirm the biological relevance of these candidates in *B. napus*. Markers identified from AT could be useful for marker-assisted selection strategies for improving B-usage efficiency, whilst minimising fertiliser requirements. This may be particularly pertinent for improving cultivation within B-deficient soils and therefore enhancing the arable land available. The ultimate aim would be for crop yields to be improved; an imperative initiative given the projected population growth and increased food demands within the coming decades.

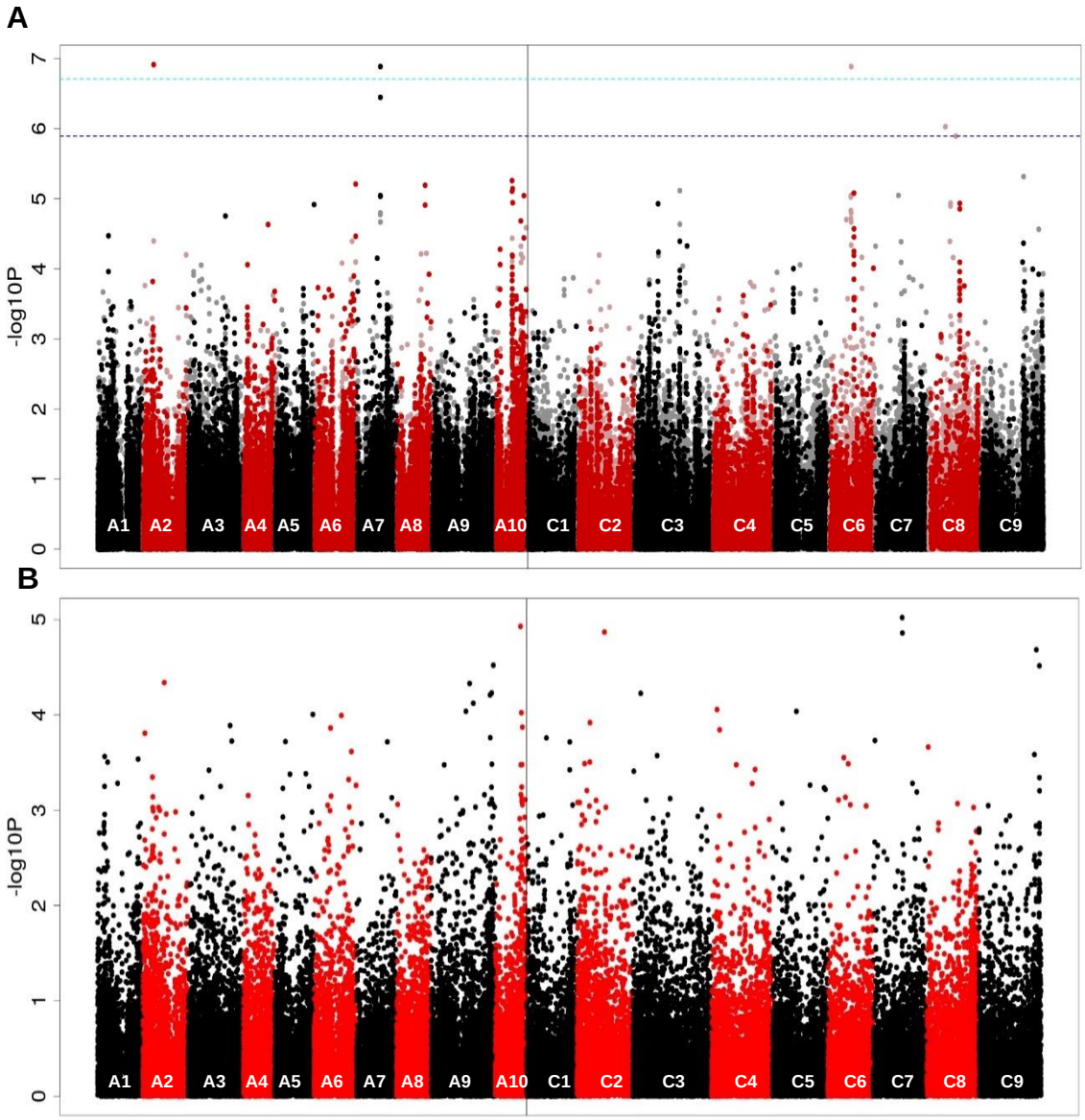
Appendices

Appendix I – Leaf AT Manhattan Plots

SNP associations were calculated with the R script GAPIT (Lipka et al., 2012) using a compressed mixed linear modelling with inference of population structure by Q-matrix produced using PSIKO (Popescu et al., 2014). GEM associations were performed using Regress, performing fixed-effect linear modelling, using PKISO-derived Q-matrix and RPKM values as explanatory variables, and leaf nutrient concentration as the response variable. The –log10P value of SNP and GEM associations were plotted against the pseudomolecule database (representative of the nineteen *B. napus* chromosomes) based of the CDS gene model order. In the SNP analyses, light red and grey points represent hemi-SNP markers that have not been linkage mapped to a genome, whereas black and red markers represent markers assigned with confidence to a genomic position. The grey and blue lines show the 5% FDR (Benjamini & Hochberg 1995) and the Bonferroni corrected significance threshold of 0.05 (Dunn, 1961) respectively.



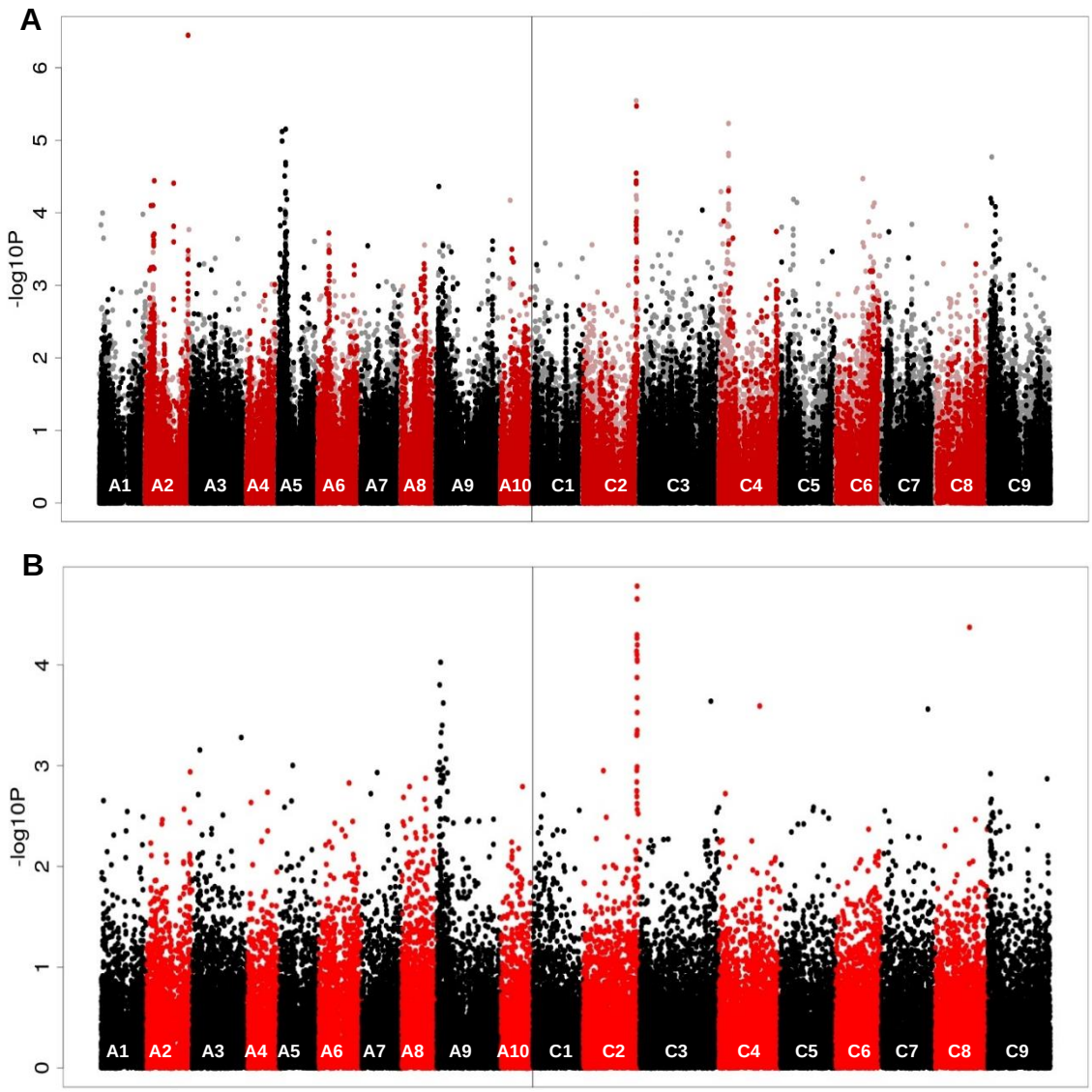
Appendix 1.1 Genome wide distribution of (a) SNP and (b) GEM markers associating with B concentration (mg/kg) in leaves from 274 accessions



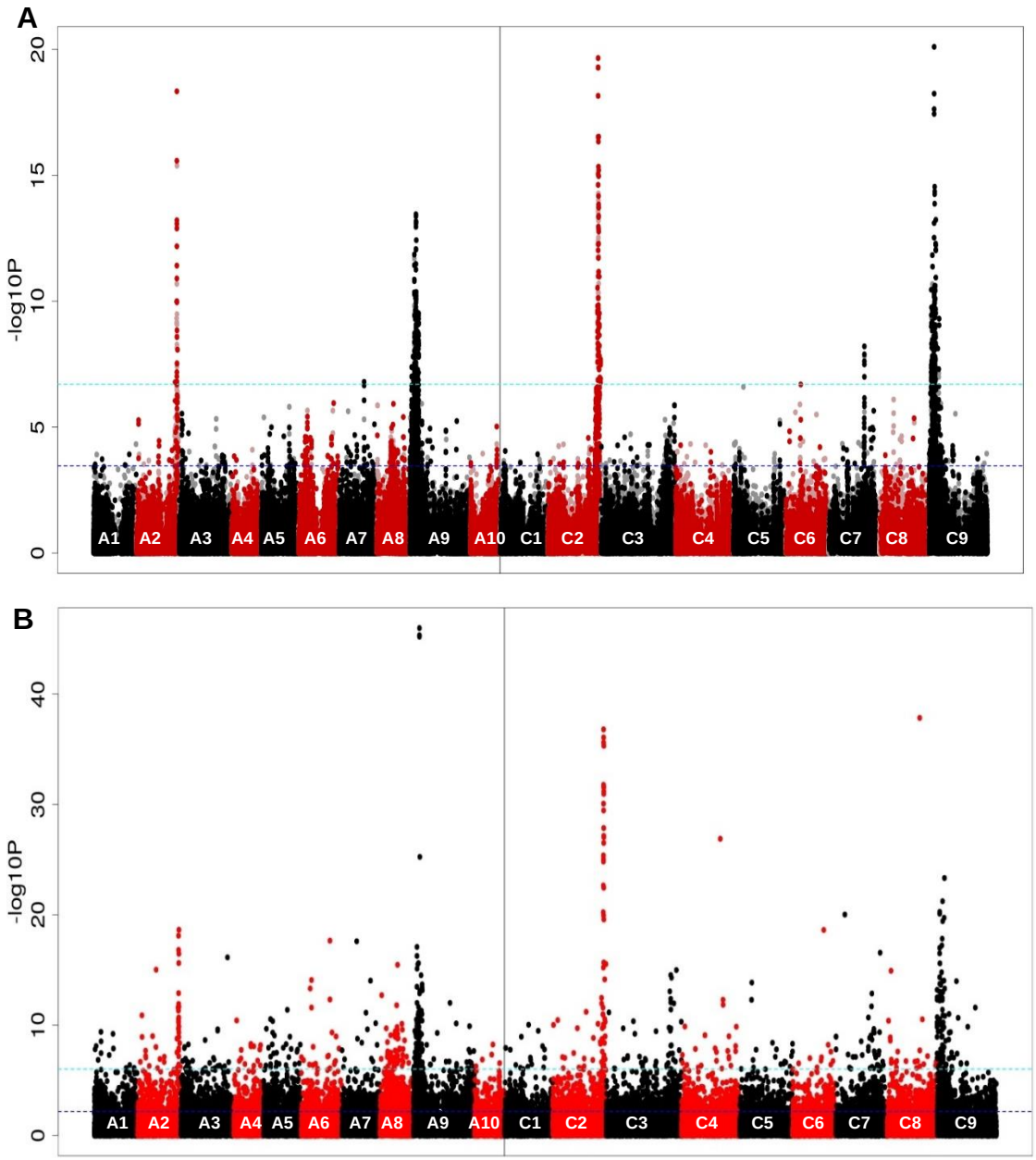
Appendix 1.3 Genome wide distribution of (a) SNP and (b) GEM markers associating with Mn concentration (mg/kg) in leaves from 383 accessions

Appendix II- Seed AT Manhattan Plots

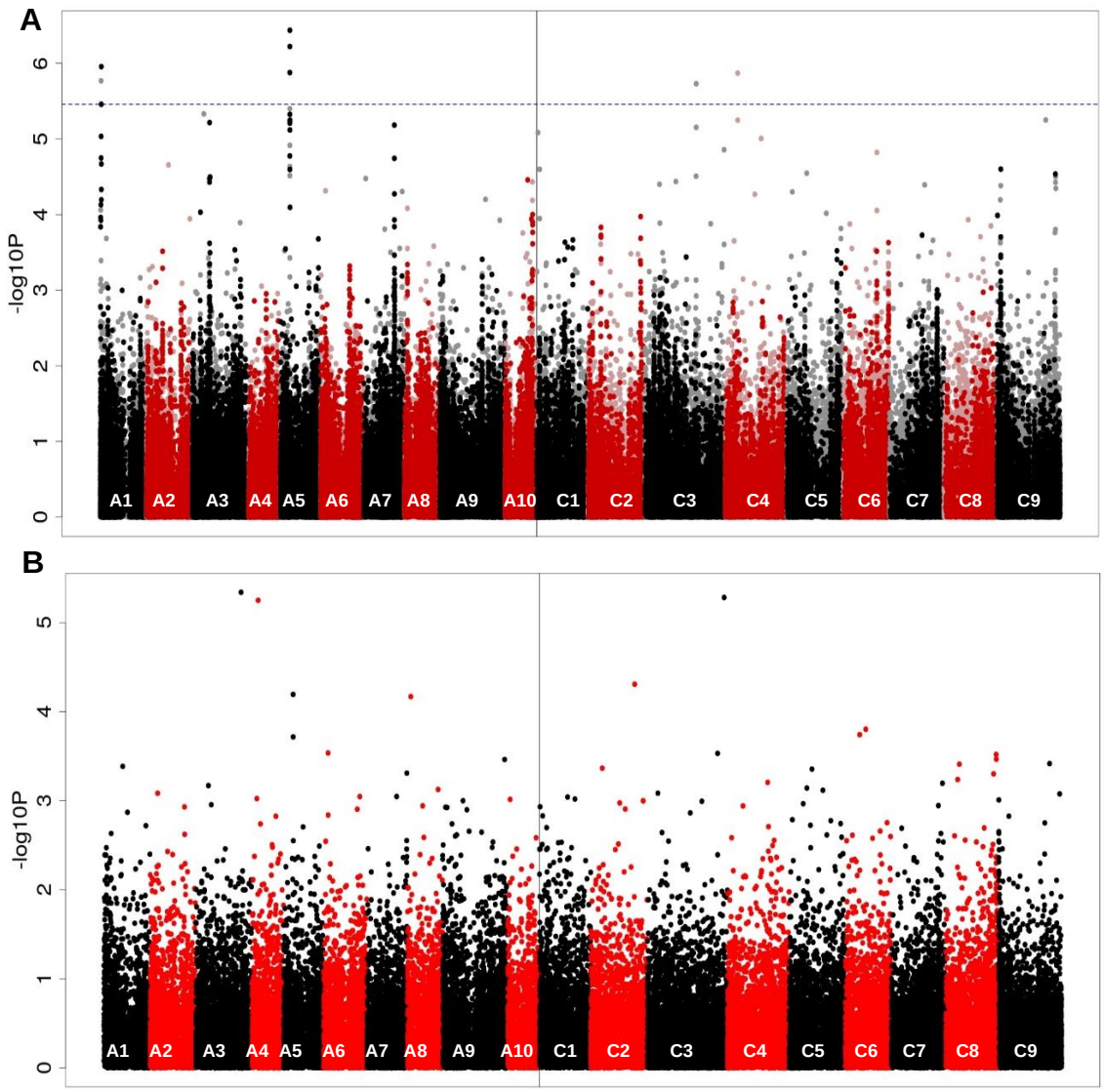
SNP associations were calculated with the R script GAPIT (Lipka et al., 2012) using a compressed mixed linear modelling with inference of population structure by Q-matrix produced using PSIKO (Popescu et al., 2014). GEM associations were performed using Regress, performing fixed-effect linear modelling, using PKISO-derived Q-matrix and RPKM values as explanatory variables, and seed nutrient concentrations as the response variable. The –log10P value of SNP and GEM associations were plotted against the pseudomolecule database (representative of the nineteen *B. napus* chromosomes) based of the CDS gene model order. In the SNP analyses, light red and grey points represent hemi-SNP markers that have not been linkage mapped to a genome, whereas black and red markers represent markers assigned with confidence to a genomic position. The grey and blue lines show the 5% FDR (Benjamini & Hochberg 1995) and the Bonferroni corrected significance threshold of 0.05 (Dunn, 1961) respectively.



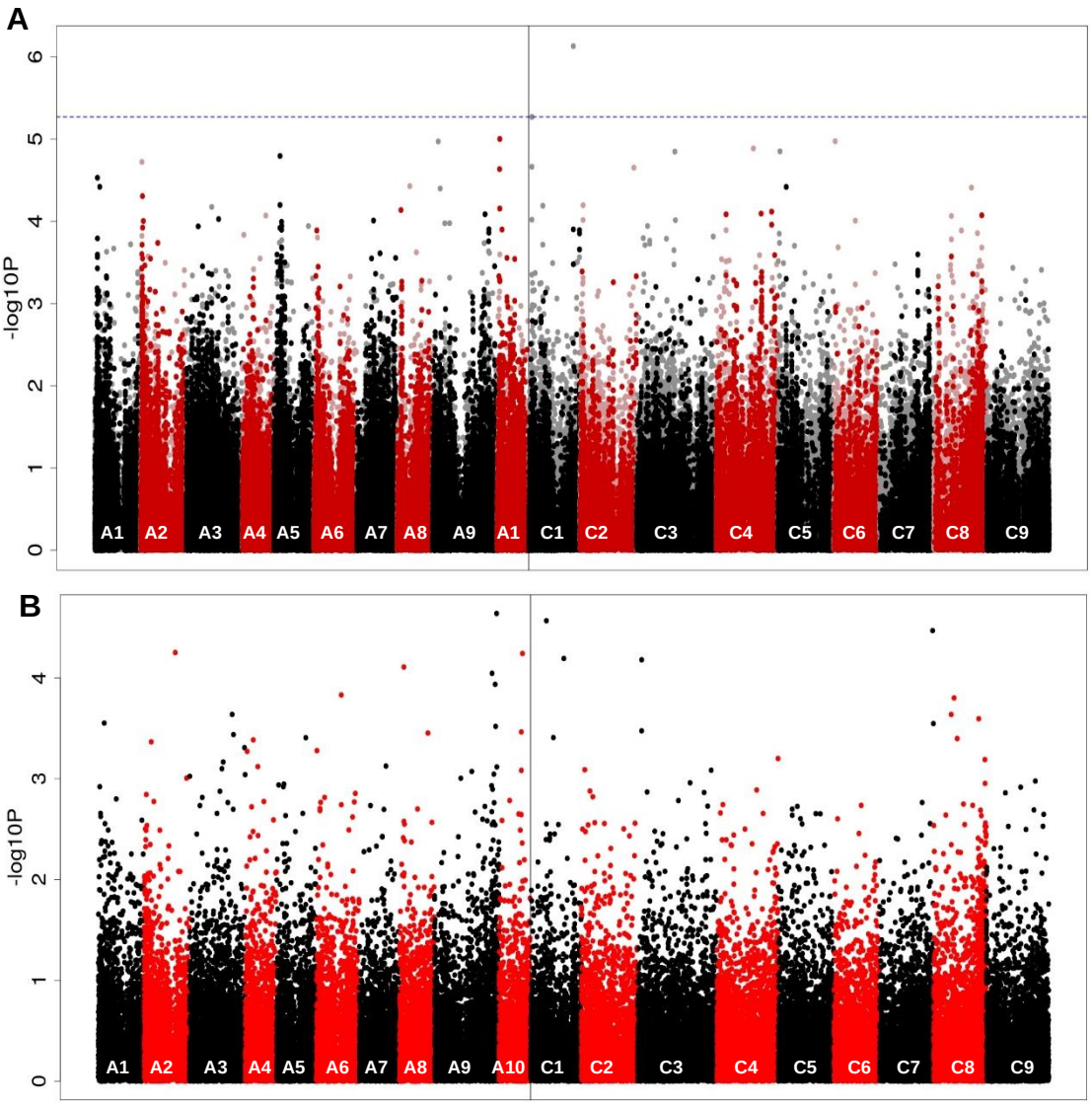
Appendix 2.2 Genome wide distribution of (a) SNP and (b) GEM markers associating with B concentration (mg/kg) in seeds from 274 accessions



Appendix 2.3 Genome wide distribution of (a) SNP and (b) GEM markers associating with S concentration (mg/kg) in seeds from 383 accessions



Appendix 2.4 Genome wide distribution of (a) SNP and (b) GEM markers associating with Mn concentration (mg/kg) in seeds from 383 accessions



Appendix 2.5 Genome wide distribution of (a) SNP and (b) GEM markers associating with As concentration (mg/kg) in seeds from 383 accessions

Appendix III- Genotyping and statistical analyses of leaf ionome candidates

Appendix 3.1 T-Test output for arf2 (SALK_035537) in comparison to Col-0 wild type in leaf.
T value (T), probability (p) and degrees of freedom (df) are given for each element.

	B	Mg	P	S	K	Ca	Mn	Zn	Mo
t	6.56	-2.56	5.54	2.71	-1.13	-3.91	1.31	1.27	-1.05
p	<0.001	0.028	<0.001	0.033	0.285	0.007	0.219	0.233	0.316
df	9	10	10	10	10	10	10	10	10

Appendix 3.2 T-Test output dur (SALK_000741C) in comparison to Col-0 wild type in leaf.
T value (T), probability (p) and degrees of freedom (df) are given for each element.

	B	Mg	P	S	K	Ca	Mn	Zn	Mo
t	0.99	0.85	-1.2	-3.09	0.42	2.19	-3.73	0.37	-8.51
p	0.341	0.408	0.249	0.008	0.679	0.046	0.002	0.719	<0.001
df	13	14	14	14	14	14	14	14	14

Appendix 3.3 T-Test output wnk (SALK_201692C) in comparison to Col-0 wild type in leaf.
T value (T), probability (p) and degrees of freedom (df) are given for each element.

	B	Mg	P	S	K	Ca	Mn	Zn	Mo
t	2.20	2.50	1.78	-0.43	-2.46	1.86	1.70	-0.25	0.01
p	0.047	0.025	0.096	0.671	0.028	0.084	0.110	0.809	0.993
df	13	14	14	14	14	14	14	14	14

Appendix 3.4 T-Test output pip2;2 (SAIL_169_A03) in comparison to Col-0 wild type in leaf.
T value (T), probability (p) and degrees of freedom (df) are given for each element.

	B	Mg	P	S	K	Ca	Mn	Zn	Mo
t	3.54	1.1	-0.89	0.44	-0.104	0.83	-4.06	-0.45	-0.87
p	0.005	0.295	0.391	0.671	0.319	0.420	0.002	0.661	0.401
df	11	11	12	12	12	12	11	12	12

Appendix 3.5 T-Test output be3 (SALK_030954C) in comparison to Col-0 wild type in leaf.
T value (T), probability (p) and degrees of freedom (df) are given for each element.

	B	Mg	P	S	K	Ca	Mn	Zn	Mo
t	-11.65	-1.89	1.62	2.86	-3.04	-2.12	-0.50	1.26	-1.36
p	<0.001	0.081	0.128	0.013	0.009	0.068	0.628	0.247	0.198
df	13	13	13	13	13	13	13	13	13

Appendix 3.6 T-Test output acr11 (SALK_206346C) in comparison to Col-0 wild type in leaf. *T value (T), probability (p) and degrees of freedom (df) are given for each element.*

	B	Mg	P	S	K	Ca	Mn	Zn	Mo
t	2.26	0.72	-8.35	-3.89	5.8	0.31	0.02	-1.29	-2.86
p	0.039	0.482	<0.001	0.001	<0.001	0.759	0.986	0.216	0.012
df	15	15	15	15	15	15	15	15	15

Appendix 3.7 T-Test output atp (SALK_078674C) in comparison to Col-0 wild type in leaf. *T value (T), probability (p) and degrees of freedom (df) are given for each element.*

	B	Mg	P	S	K	Ca	Mn	Zn	Mo
t	1.96	1.23	-1.44	-1.45	-0.72	1.1	-2.11	1.26	-2.39
p	0.074	0.250	0.172	0.198	0.486	0.292	0.054	0.231	0.031
df	12	9.21	13	6	13	13	14	13	14

Appendix 3.8 T-Test output ger (SALK_020145C) in comparison to Col-0 wild type in leaf. *T value (T), probability (p) and degrees of freedom (df) are given for each element.*

	B	Mg	P	S	K	Ca	Mn	Zn	Mo
t	-0.3	1.26	0.77	1.27	-1.03	1.37	2.72	2.57	0.33
p	0.772	0.235	0.431	0.233	0.325	0.200	0.022	0.028	0.745
df	10	10	10	10	10	10	10	10	10

Appendix 3.9 T-Test output hmt (SALK_060068C) in comparison to Col-0 wild type in leaf. *T value (T), probability (p) and degrees of freedom (df) are given for each element.*

	B	Mg	P	S	K	Ca	Mn	Zn	Mo
t	1.3	1.64	0.32	-0.45	-0.80	5.51	-0.56	-0.95	8.45
p	0.215	0.124	0.757	0.659	0.439	<0.001	0.583	0.367	<0.001
df	13	13	13	13	14	14	14	9.08	14

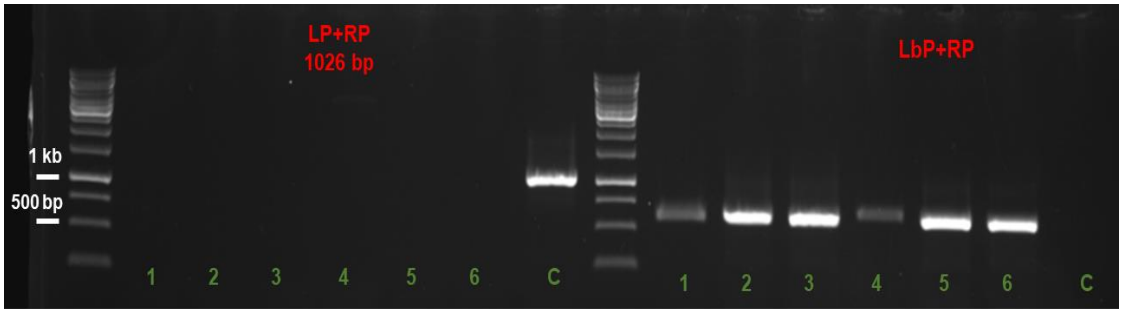
Appendix 3.10 T-Test output nip2;1 (SAIL_439_A02) in comparison to Col-0 wild type in leaf. *T value (T), probability (p) and degrees of freedom (df) are given for each element.*

	B	Mg	P	S	K	Ca	Mn	Zn	Mo	As
t	2.6	2.47	-2.30	0.67	0.15	-1.15	-2.04	0.38	-2.04	1.23
p	0.023	0.027	0.037	0.516	0.880	0.270	0.061	0.713	0.061	0.241
df	13	14	14	14	14	13	14	9.6	14	13

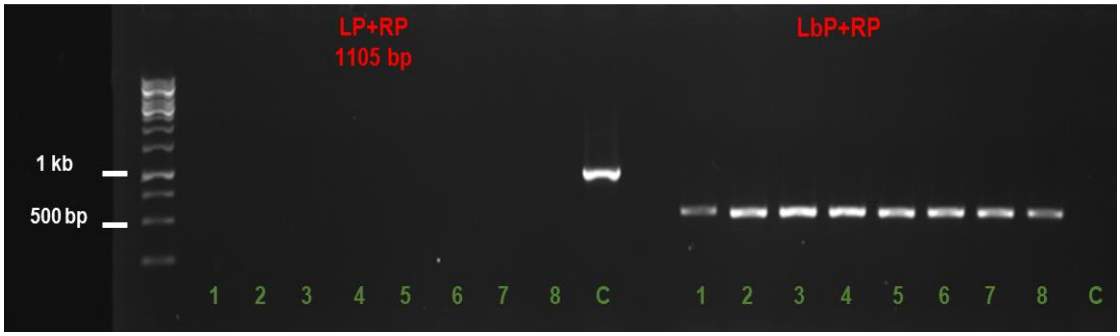
Appendix 3.11 Gel electrophoresis for the detection of homozygous T-DNA mutants for leaf ionome candidates tested via ICP-MS analyses

(a) arf2 (b) dur (c) wnk (d) pip2;2 (e) be3 (f) acr11 (g) atp (h) ger (i) hmt (j) nip2;1. All gels were electrophoresed on a 1% agarose gel with a GeneRuler 1 kb DNA ladder (ThermoFisher Scientific) included and visualised under UV. The 500 bp and 1 kb bands on the ladder have been highlighted for reference. The gels show either the left primer (LP) and right primer (RP) or the left border primer (LbP) and right primer reaction combination. The fragment length for LP+RP was obtained from the SALK T-DNA design tool (signal.salk.edu/tdnaprimers). Fragment lengths for the homozygous inserts vary from ~ 400-700 bp. The numbers represent the gDNA stocks extracted from each individual plant and used as the template for both reaction pairs. 'C' represent the Col-0 control used to test the wild type amplification in the LP+RP reaction, and that no template was amplified in the LbP+RP reaction. (-) shows non-amplified bands whereas 'WT' shows wild type individuals. Individuals homozygous for the T-DNA insert should only show a band in the LbP + RP reaction, which is shown either on the bottom or right hand side.

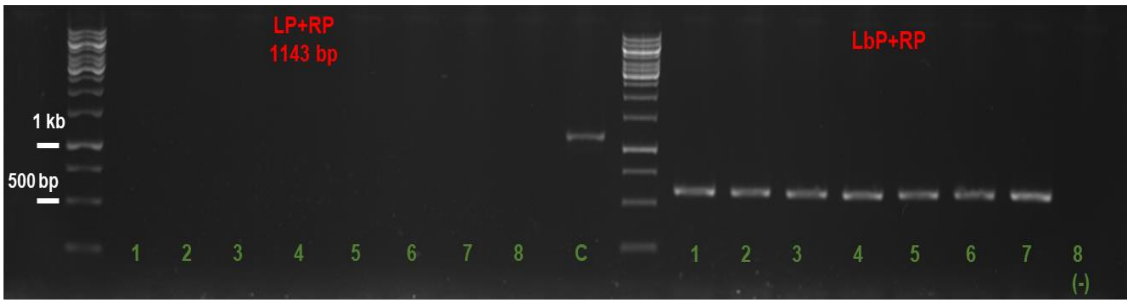
(a) *arf2*



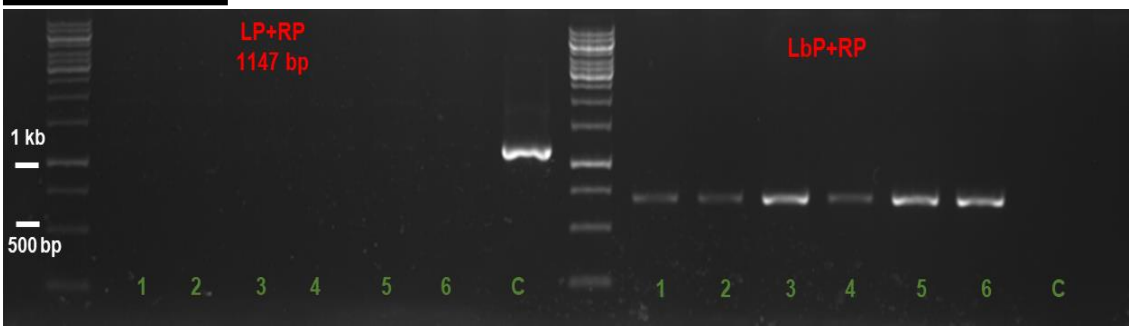
(b) *dur*



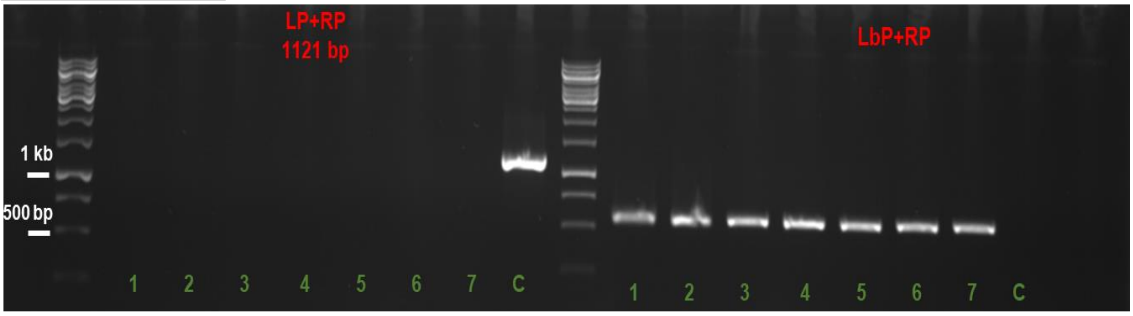
(c) *wnk*



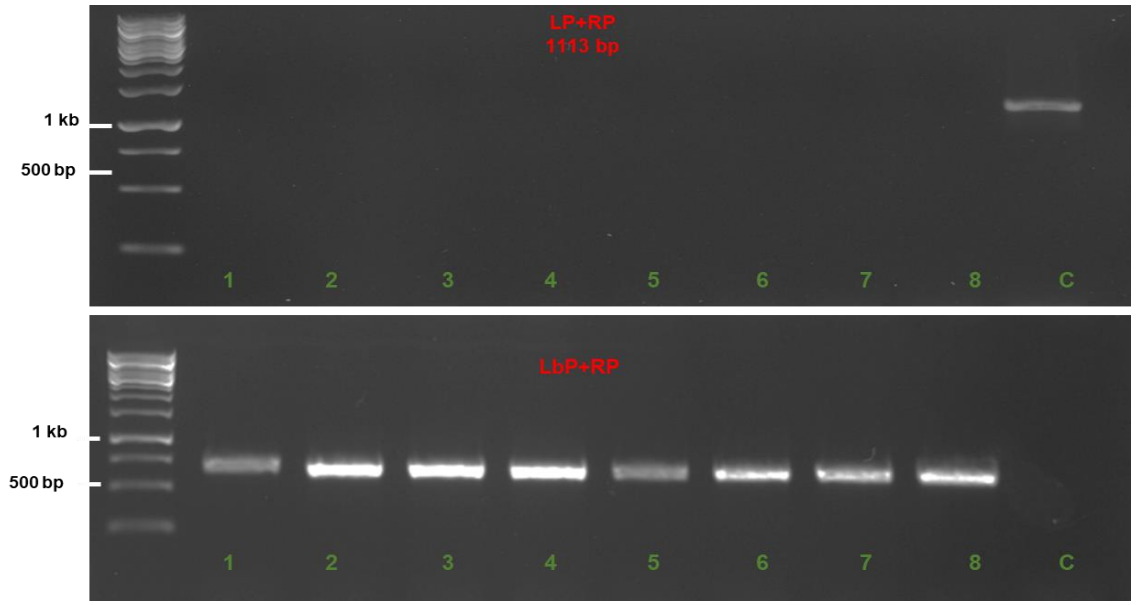
(d) *pip2;2*



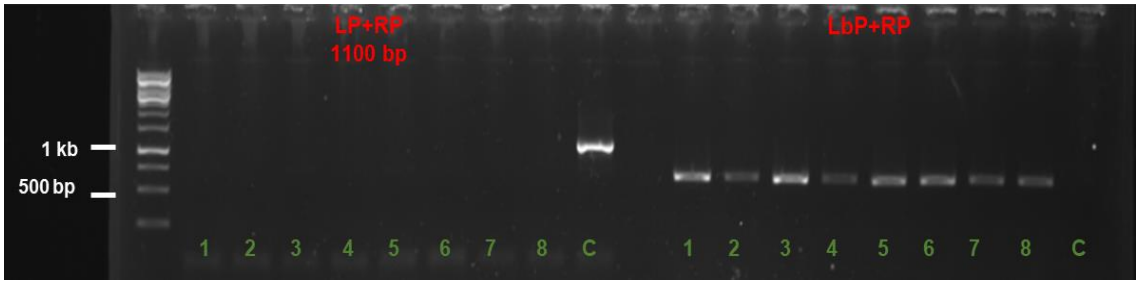
(e) *be3*



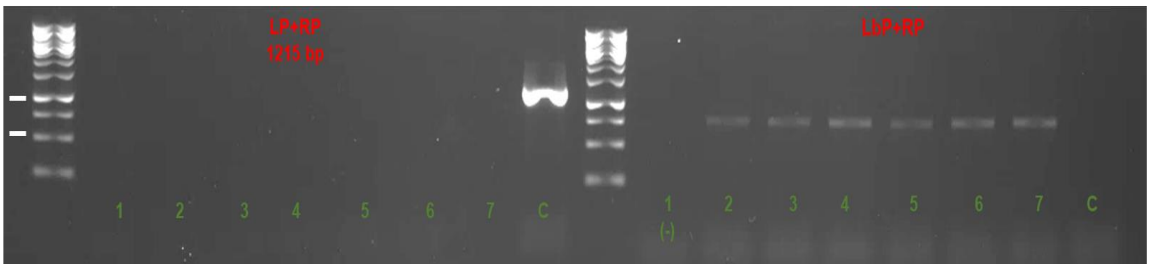
(f) a*cr11*



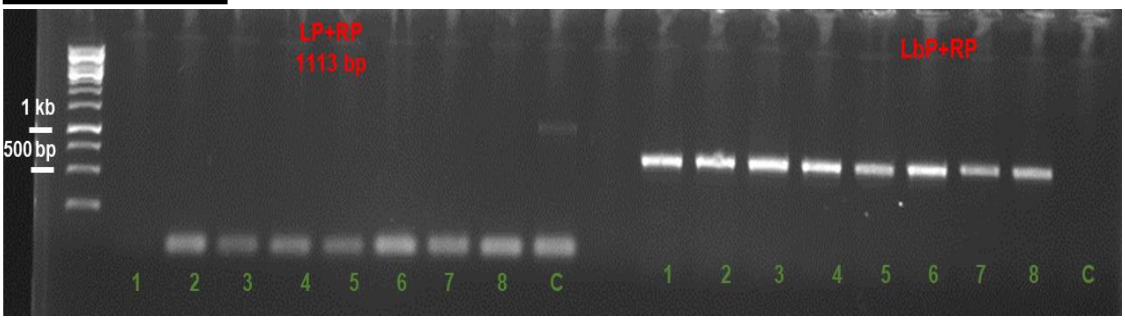
(g) *atp*



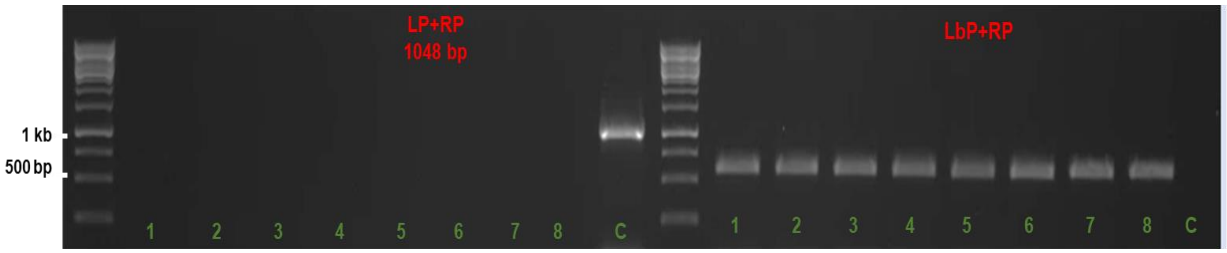
(h) *ger*



(i) *hmt*


LP+RP
1113 bp
LbP+RP
1 kb
500 bp
1 2 3 4 5 6 7 8 C
1 2 3 4 5 6 7 8 C

(j) *nip2;1*


LP+RP
1048 bp
LbP+RP
1 kb
500 bp
1 2 3 4 5 6 7 8 C
1 2 3 4 5 6 7 8 C

Appendix IV- Genotyping and statistical analyses of seed ionome candidates

Appendix 4.1 T-Test output nip2;1 (SAIL_439_A02) in comparison to Col-0 wild type in seed. *T value (T), probability (p) and degrees of freedom (df) are given for each element.*

	B	Mg	P	S	K	Ca	Mn	Zn	Mo	As
t	3.39	1.34	2.86	1.86	9.92	3.1	3.71	-1.69	10.68	-11.08
p	0.005	0.206	0.014	0.087	<0.001	0.009	0.008	0.138	<0.001	<0.001
df	12	12	12	12	12	12	6.68	12	6.45	12

Appendix 4.2 T-Test output be3 (SALK_030954C) in comparison to Col-0 wild type in seed. *T value (T), probability (p) and degrees of freedom (df) are given for each element.*

	B	Mg	P	S	K	Ca	Mn	Zn	Mo
t	0.01	-0.59	1.27	1.81	4.03	1.91	2.81	0.83	5.52
p	0.994	0.567	0.232	0.101	0.002	0.085	0.018	0.425	<0.001
df	10	10	10	10	10	10	10	10	10

Appendix 4.3 T-Test output arf2 (SALK_035537) in comparison to Col-0 wild type in seed. *T value (T), probability (p) and degrees of freedom (df) are given for each element.*

	B	Mg	P	S	K	Ca	Mn	Zn	Mo
t	0.09	8.39	21.68	6.68	46.44	2.66	22.22	2.64	17.00
p	0.932	<0.001	<0.001	<0.001	<0.001	0.019	<0.001	0.019	<0.001
df	14	14	14	14	14	14	14	14	14

Appendix 4.4 T-Test output umamit14 (SALK_037123C) in comparison to Col-0 wild type in seed. *T value (T), probability (p) and degrees of freedom (df) are given for each element.*

	B	Mg	P	S	K	Ca	Mn	Zn	Mo
t	-0.52	-3.28	-1.26	1.63	1.52	-0.66	2.32	-4.84	4.18
p	0.612	0.007	0.249	0.128	0.177	0.525	0.052	<0.001	0.003
df	12	12	12	12	12	12	12	12	12

Appendix 4.5 T-Test output pip2;2 (SAIL_169_A03) in comparison to Col-0 wild type in seed. *T value (T), probability (p) and degrees of freedom (df) are given for each element.*

	B	Mg	P	S	K	Ca	Mn	Zn	Mo
t	0.149	-1.66	3.21	1.91	3.59	-0.52	-4.61	-2.93	-4.81
p	0.149	0.126	0.008	0.082	0.004	0.571	<0.001	0.014	<0.001
df	11	11	11	11	11	11	11	11	11

Appendix 4.6 T-Test output xeg113a (SALK_007511C) in comparison to Col-0 wild type in seed. *T value (T), probability (p) and degrees of freedom (df) are given for each element.*

	B	**Mg**	**P**	**S**	**K**	**Ca**	**Mn**	**Zn**	**Mo**
t	1.6010	9.35	8.253	5.12	16.00	3.85	5.07	4.36	8.81
p	0.137	<0.001	<0.001	<0.001	<0.001	0.003	<0.001	0.001	<0.001
df	11	11	11	11	11	11	10	11	11

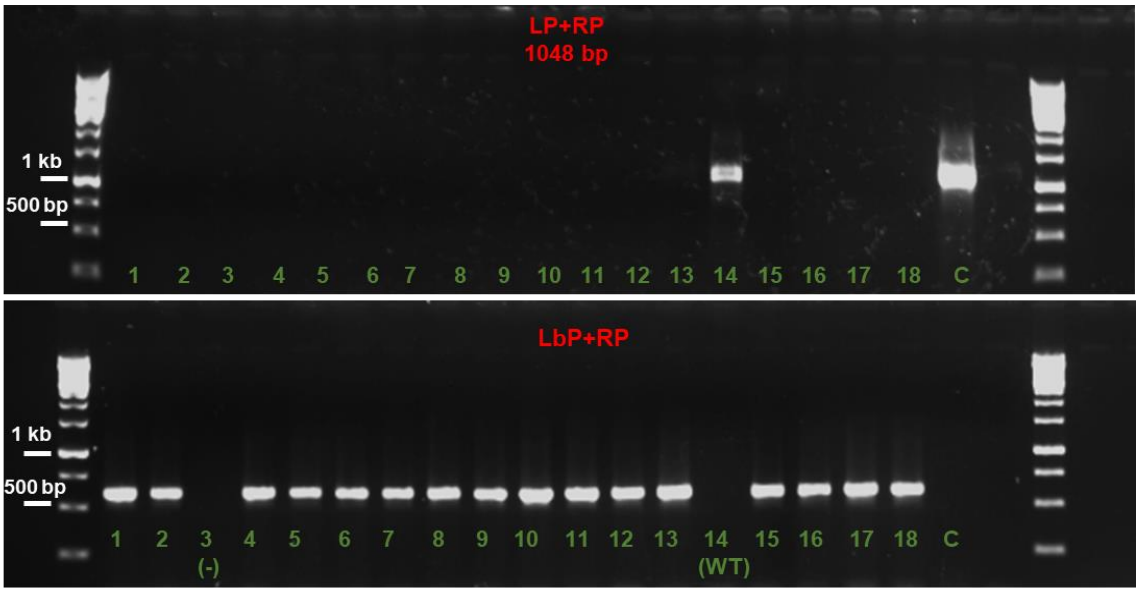
Appendix 4.7-Test output xeg113b (SALK_066991C) in comparison to Col-0 wild type in seed. *T value (T), probability (p) and degrees of freedom (df) are given for each element.*

	B	**Mg**	**P**	**S**	**K**	**Ca**	**Mn**	**Zn**	**Mo**
t	3.55	6.00	4.06	2.28	4.99	3.81	2.85	0.37	12.96
p	0.005	<0.001	0.002	0.048	<0.001	0.003	0.017	0.722	<0.001
df	10	10	10	10	10	10	10	10	10

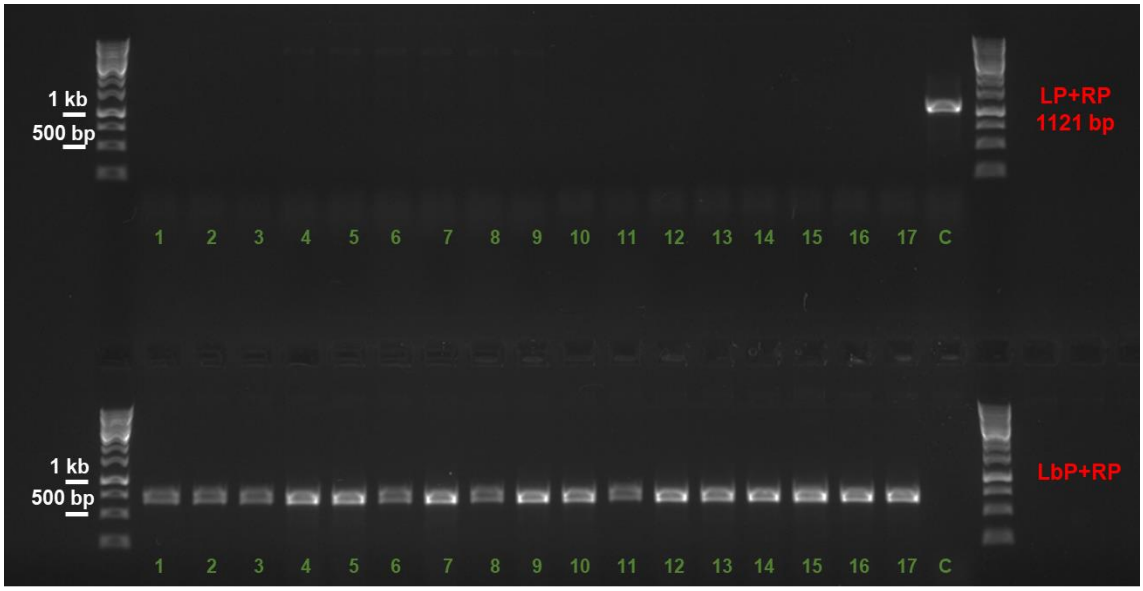
Appendix 4.8 Gel electrophoresis for the detection of homozygous T-DNA mutants for seed ionome candidates tested via ICP-MS analyses

(a) nip2;1, (b) be3 (c) arf2 (d) umamit14 (e) xeg113a (f) xeg113b (g) pip2;2. All gels were electrophoresed on a 1% agarose gel with a GeneRuler 1 kb DNA ladder (ThermoFisher Scientific) included and visualised under UV. The 500 bp and 1 kb bands on the ladder have been highlighted for reference. The gels show either the left primer (LP) and right primer (RP) or the left border primer (LbP) and right primer reaction combination. The fragment length for LP+RP was obtained from the SALK T-DNA design tool (signal.salk.edu/tdnaprimers). Fragment lengths for the homozygous inserts vary from ~ 400-700 bp. The numbers represent the gDNA stocks extracted from each individual plant and used as the template for both reaction pairs. 'C' represent the Col-0 control used to test the wild type amplification in the LP+RP reaction, and that no template was amplified in the LbP+RP reaction. (-) shows non-amplified bands whereas 'WT' shows wild type individuals. Individuals homozygous for the T-DNA insert should only show a band in the LbP + RP reaction, which is shown on the bottom panel.

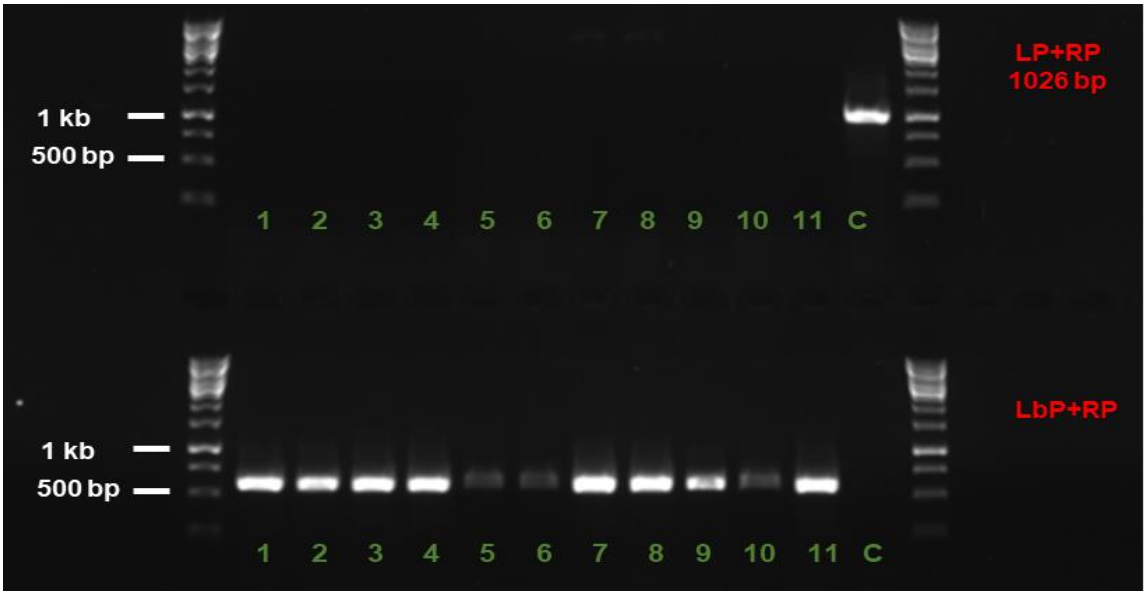
(a) *nip2;1*



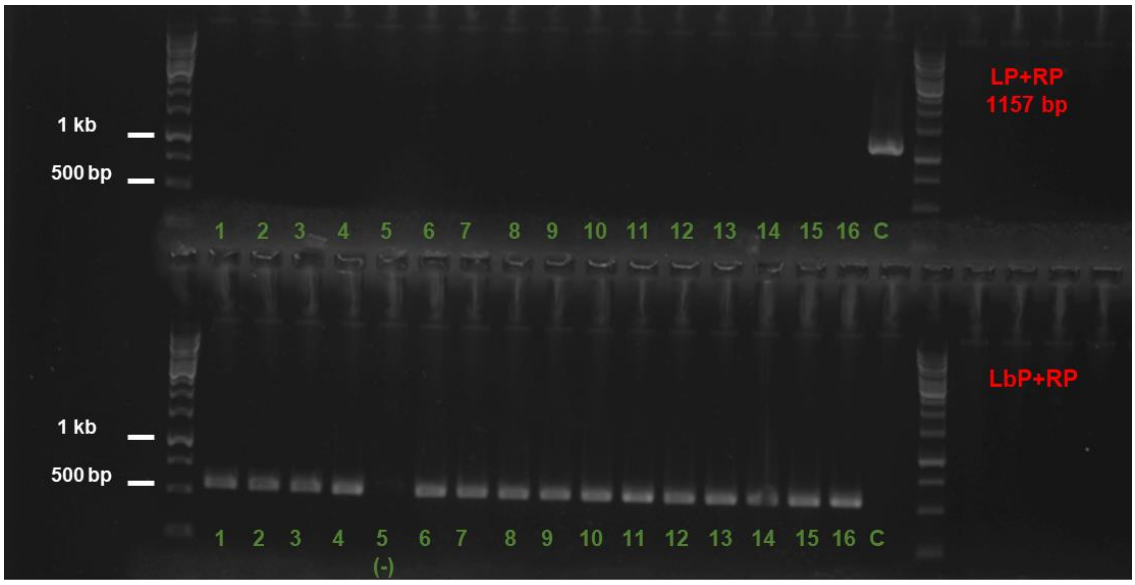
(b) *be3*


1 kb
500 bp
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 C
LP+RP
1121 bp
1 kb
500 bp
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 C
LbP+RP

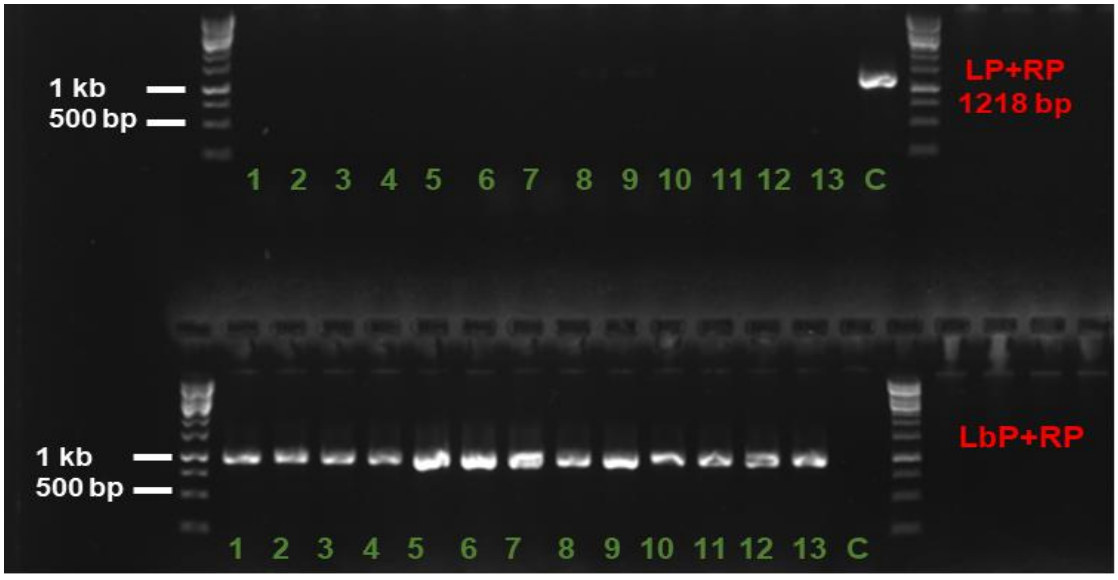
(c) *arf2*


1 kb
500 bp
1 2 3 4 5 6 7 8 9 10 11 C
LP+RP
1026 bp
1 kb
500 bp
1 2 3 4 5 6 7 8 9 10 11 C
LbP+RP

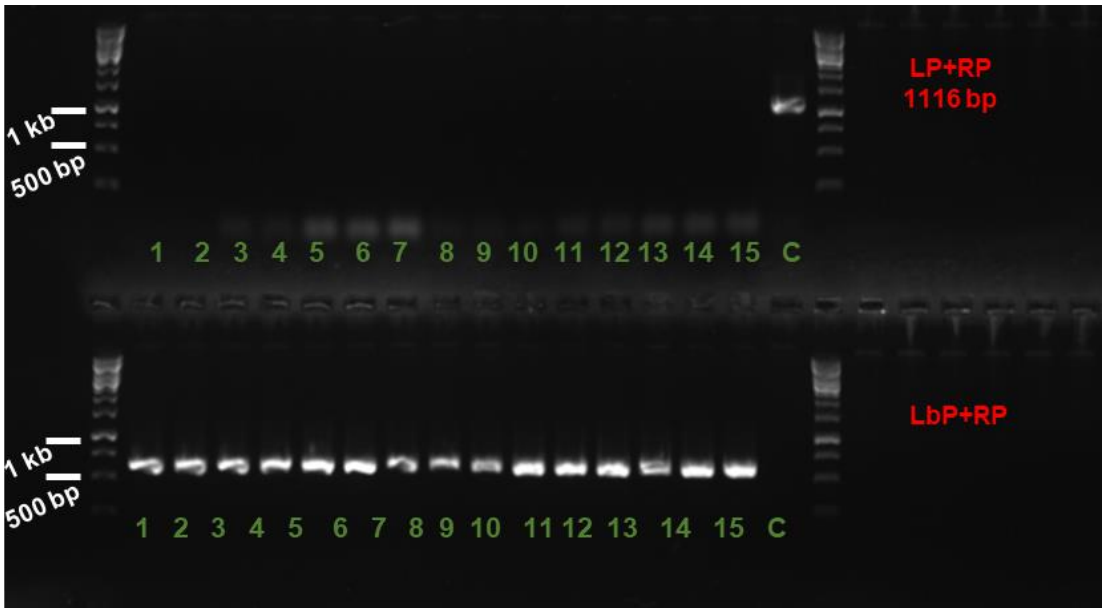
(d) *umamit14*



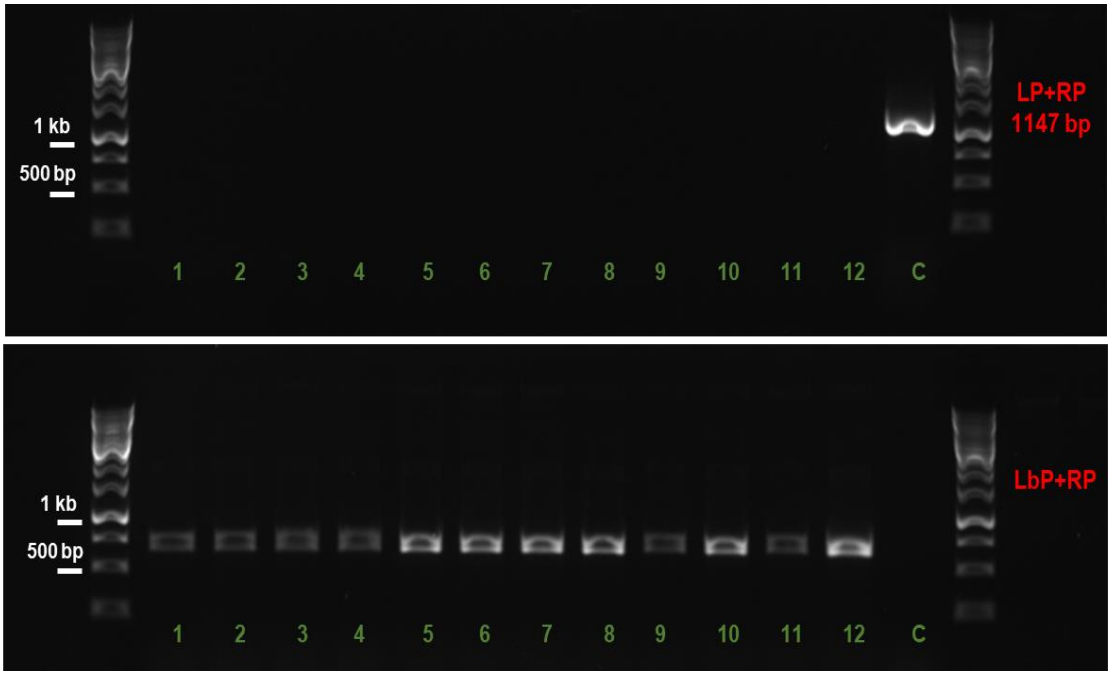
(e) *xeg113a*



(f) *xeg113b*



(g) *pip2;2*



Appendix V- Statistical analyses of swede data

Appendix 5.1 Number of BH affected roots in resistant and susceptible accessions

The number of BH affected roots in BH resistant (Mg, Lz, Mg x He and He) and susceptible (He x Ac, Ag x He, Ac and Ag) have been provided. No significant difference in the instances of BH was observed, as calculated by the Freeman-Halton extension of Fisher's exact test (p = 0.650).

Treatment	BH Susceptibility	Number of BH affected Roots
Low	Resistant	4
	Susceptible	6
Mid	Resistant	8
	Susceptible	6
High	Resistant	2
	Susceptible	1

Appendix 5.2 Two- way ANOVA output for the effect of treatment and accession on B concentration within swede roots.

F statistic (F), probability (p) and degrees of freedom (df) are given for each.

Swede	F	df	p
Accession	7.09	13,161	<0.001
Treatment	11.444	2,161	<0.001
Accession*Treatment	1.242	26,161	0.215

Appendix 5.3 ANOVA results of B concentrations in swede between H_3BO_3 treatments *The F statistic (F), degrees of freedom (df) and p value (p) are given, alongside the differences highlighted with post hoc Tukey analysis (p <0.05) between H_3BO_3 treatments across swede accessions. Where the means are significantly different the letters are not shared, with the letter a indicating the lowest B concentrations and b the highest.*

Treatment	F	df	p	Tukey (p <0.05)		
				Low	Mid	High
	7.73	2,159	<0.001	a	a	b

Appendix 5.4 ANOVA results of B concentrations between swede accessions

The F statistic (F), degrees of freedom (df) and p value (p) are given, alongside the differences highlighted with post hoc Tukey analysis (p <0.05) between swede accessions. Where the means are significantly different the letters are not shared, with the letter a indicating the lowest B concentrations and d the highest.

F	df	p	Tukey (p <0.05)						
5.99	13,148	<0.001	**Mg**	**Lz**	**He**	**Ac**	**Ag**	**Sl**	**Tk**
			d	a	abcd	abcd	abc	bdc	abcd
			Mg x He	**He x Ac**	**Ag x He**	**G1*He**	**Bh**	**Tn**	**Kf**
			abcd	ab	abcd	d	abc	d	cd

Appendix 5.5 ANOVA results of B concentrations in root within swede accessions

The F statistic (F), degrees of freedom (df) and p value (p) are given, alongside the significant differences in mean B concentrations, as highlighted with post hoc Tukey analysis (p <0.05) between H_3BO_3 treatments for individual accessions. Where the means are significantly different the letters are not shared, with the letter a indicating the lowest B concentrations and b the highest. As shown in figure 5.9.

Accession	F	df	p	Tukey (p <0.05)		
				Low	Mid	High
Mg	14.12	2,9	0.002	a	a	b
Lz	0.20	2,9	0.824	a	a	a
He	4.86	2,9	0.037	ab	a	b
Ac	4.01	2,8	0.062	a	a	a
Ag	2.58	2,9	0.130	a	a	a
Sl	0.56	2,9	0.592	a	a	a
Tk	0.81	2,9	0.476	a	a	a
Mg x He	0.13	2,9	0.880	a	a	a
He x Ac	0.07	2,9	0.930	a	a	a
Ag x He	4.69	2,9	0.040	a	ab	b
G1*He	0.91	2,9	0.440	a	a	a
Bh	0.14	2,9	0.871	a	a	a
Tn	0.35	2,7	0.719	a	a	a
Kf	1.26	2,7	0.342	a	a	a

Appendix 5.6 ANOVA results of B concentrations in leaf between H_3BO_3 treatments

The F statistic (F), degrees of freedom (df) and p value (p) are given, alongside the differences highlighted with post hoc Tukey analysis (p <0.05) between H_3BO_3 treatments across swede accessions for leaf material sampled at four time points (TP1-TP4). Where the means are significantly different the letters are not shared, with the letter a indicating the lowest B concentrations and b the highest.

Time Point	F	df	p	Tukey (p <0.05)		
				Low	Mid	High
TP1	24.07	2,39	<0.001	a	a	b
TP2	5.46	2,39	0.008	a	b	b
TP3	5.92	2,39	0.006	a	ab	b
TP4	4.58	2,39	0.016	b	a	a

Appendix 5.7 ANOVA results of B concentrations in leaf TP1 within swede accessions

The F statistic (F), degrees of freedom (df) and p value (p) are given, alongside the significant differences in mean B concentrations, as highlighted with post hoc Tukey analysis (p <0.05) between H_3BO_3 treatments for individual accessions. Where the means are significantly different the letters are not shared, with the letter a indicating the lowest B concentrations and b the highest.

Leaf One (TP1)	F	df	p	Tukey (p <0.05)		
				Low	Mid	High
Mg	0.25	2,9	0.768	a	a	a
Lz	2.30	2,9	0.156	a	a	a
He	7.71	2,9	0.011	a	a	b
Ac	2.00	2,9	0.191	a	a	a
Ag	1.59	2,9	0.256	a	a	a
Sl	21.12	2,9	<0.001	a	a	b
Tk	1.93	2,9	0.201	a	a	a
Mg x He	1.26	2,9	0.328	a	a	a
He x Ac	6.63	2,9	0.017	a	ab	b
Ag x He	11.96	2,9	0.003	a	a	b
G1*He	9.17	2,8	0.009	a	ab	b
Bh	1.94	2,9	0.199	a	a	a
Tn	2.81	2,7	0.127	a	a	a
Kf	9.40	2,7	0.010	a	ab	b

Appendix 5.8 ANOVA results of B concentrations in leaf TP2 within swede accessions

The F statistic (F), degrees of freedom (df) and p value (p) are given, alongside the significant differences in mean B concentrations, as highlighted with post hoc Tukey analysis (p <0.05) between H_3BO_3 treatments for individual accessions. Where the means are significantly different the letters are not shared, with the letter a indicating the lowest B concentrations and b the highest.

Leaf Two (TP2)	F	df	p	Tukey (p <0.05)		
				Low	Mid	High
Mg	1.03	2,9	0.393	a	a	a
Lz	3.45	2,9	0.077	a	a	a
He	0.36	2,9	0.701	a	a	a
Ac	0.42	2,9	0.674	a	a	a
Ag	1.50	2,9	0.274	a	a	a
Sl	1.66	2,9	0.250	a	a	a
Tk	3.39	2,9	0.079	a	a	a
Mg x He	4.87	2,9	0.037	a	b	b
He x Ac	1.42	2,8	0.297	a	a	a
Ag x He	3.04	2,9	0.09	a	a	b
G1*He	0.88	2,8	0.451	a	a	a
Bh	2.42	2,9	0.147	a	a	a
Tn	5.85	2,7	0.032	a	b	b
Kf	2.48	2,7	0.153	a	a	a

Appendix 5.9 ANOVA results of B concentrations in leaf TP3 within swede accessions

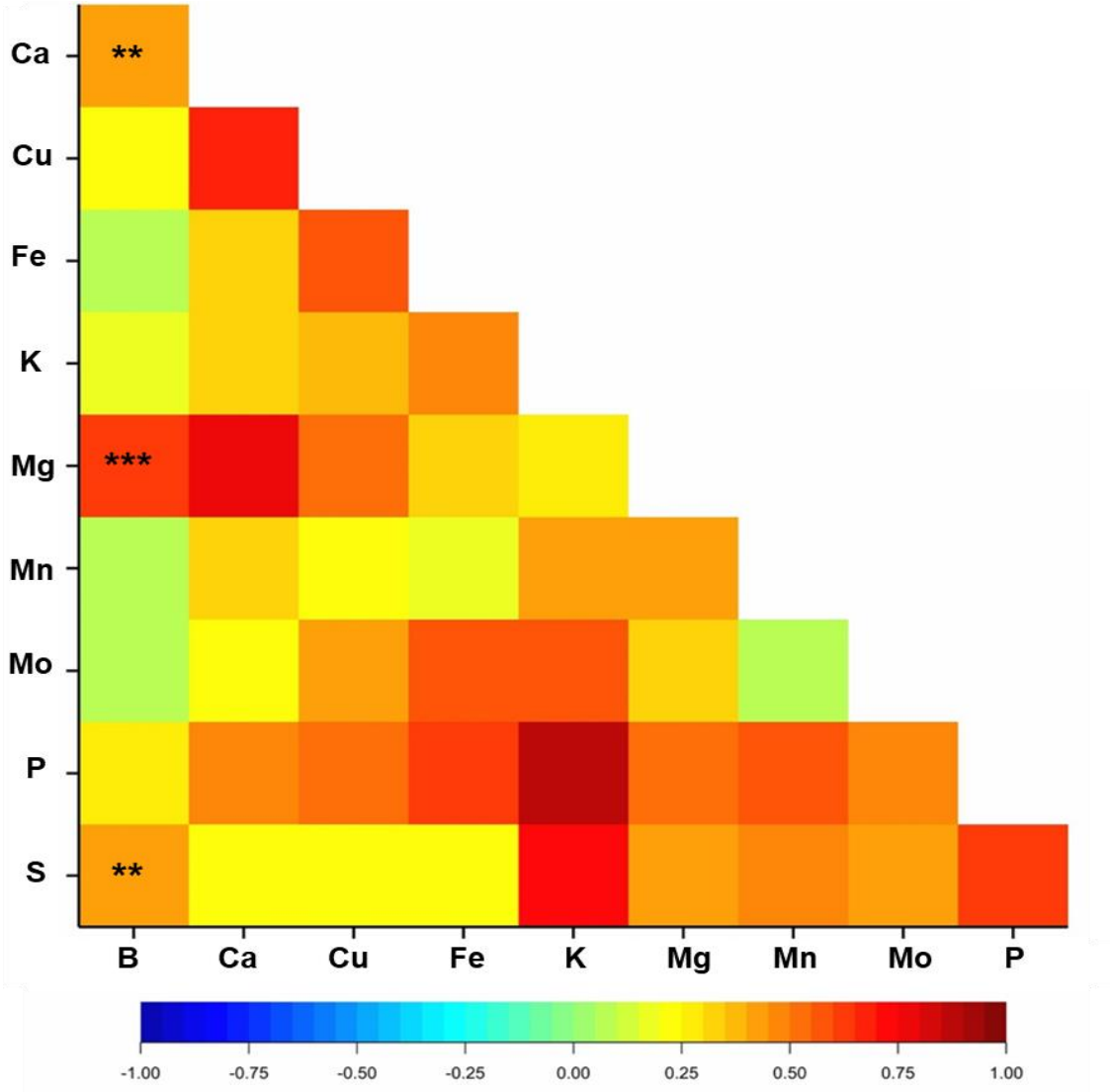
The F statistic (F), degrees of freedom (df) and p value (p) are given, alongside the significant differences in mean B concentrations, as highlighted with post hoc Tukey analysis (p <0.05) between H_3BO_3 treatments for individual accessions. Where the means are significantly different the letters are not shared, with the letter a indicating the lowest B concentrations and b the highest.

Leaf Three (TP3)	F	df	p	Tukey (p <0.05)		
				Low	Mid	High
Mg	1.38	2,9	0.301	a	a	a
Lz	0.25	2,9	0.784	a	a	a
He	1.79	2,8	0.226	a	a	a
Ac	0.81	2,9	0.480	a	a	a
Ag	1.93	2,9	0.200	a	a	a
Sl	1.46	2,9	0.252	a	a	a
Tk	0.06	2,9	0.942	a	a	a
Mg x He	0.73	2,9	0.510	a	a	a
He x Ac	2.47	2,9	0.140	a	a	a
Ag x He	1.82	2,9	0.216	a	a	a
G1*He	1.10	2,8	0.380	a	a	a
Bh	3.40	2,9	0.079	a	a	a
Tn	1.42	2,7	0.303	a	a	a
Kf	2.63	2,7	0.141	a	a	a

Appendix 5.10 ANOVA results of B concentrations in leaf TP4 within swede accessions

The F statistic (F), degrees of freedom (df) and p value (p) are given, alongside the significant differences in mean B concentrations, as highlighted with post hoc Tukey analysis (p <0.05) between H_3BO_3 treatments for individual accessions. Where the means are significantly different the letters are not shared, with the letter a indicating the lowest B concentrations and b the highest.

Leaf Four (TP4)	F	df	p	Tukey (p <0.05)		
				Low	Mid	High
Mg	0.12	2,9	0.886	a	a	a
Lz	0.08	2,9	0.920	a	a	a
He	2.63	2,9	0.126	a	a	a
Ac	0.30	2,9	0.747	a	a	a
Ag	10.62	2,7	0.008	a	ab	b
Sl	0.71	2,9	0.517	a	a	a
Tk	0.69	2,9	0.525	a	a	a
Mg x He	0.01	2,9	0.994	a	a	a
He x Ac	0.62	2,9	0.561	a	a	a
Ag x He	3.40	2,9	0.080	a	a	a
G1*He	4.10	2,8	0.062	a	a	a
Bh	2.98	2,9	0.101	a	a	a
Tn	4.74	2,6	0.058	a	a	a
Kf	3.81	2,6	0.076	a	a	a



Appendix 5.11 Pair-wise correlation of elemental concentrations in root

Correlation between ten essential elements analysed in roots across all H_3BO_3 treatments and accessions, whose percentage recovery was ≥ 85%. Double and triple starts represent 'B v element' significance at p <0.01 and <0.001 respectively.

Abbreviations

Abbreviation	Definition
˚C	degrees centigrade
A. thaliana	*Arabidopsis thaliana*
ABC	ATP-BINDING CASSETTE
AFLP	amplified fragment length polymorphism
AGI	*Arabidopsis* Genome Initiative
ANOVA	analysis of variance
AQPs	aquaporins
As	arsenic
As(III)	arsenite
AT	Associative Transcriptomics
B	boron
B(OH)4-	borate
B. napus	*Brassica napus*
B. napus var. *napobrassica*	swede
BH	Brown Heart
bp	base pair
BSA	bulk segregant analysis
Ca	calcium
cDNA	complementary DNA
CDS	coding DNA sequence
Cl	chlorine
cM	centimorgan
cMLM	compressed Mixed Linear Modelling
Col-0	Colombia-0 ecotype
CRM	certified reference material
CTAB	cetyl trimethylammonium bromide
Cu	copper
ddH2O	distilled deionised water
df	degrees of freedom
DGE	digital gene expression

EMS	ethyl methanesulphonate
EST	expressed sequence tags
EtOH	ethanol
Fe	iron
FST	fixation index
g	gram
GAPIT	Genome Association and Prediction Integrated tool
gDNA	genomic DNA
GEM	gene expression markers
GM	genetically modified
GWAS	genome wide association study
H_3BO_3	boric acid
Ha	hectare
HCl	hydrogen chloride
HNO_3	nitric acid
HWSB	hot-water soluble boron
IAA	isoamyl alcohol
ICP-MS	inductively coupled plasma mass spectrometry
ICP-OES	inductively coupled plasma optical emission spectroscopy
IDT	Integrated DNA Technologies
IPHs	inter-homoeologue polymorphisms
K	potassium
kb	kilo base
LA-ICP-MS	laser ablation inductively coupled plasma mass spectrometry
LbP	left border primer
LD	linkage disequilibrium
LP	left primer
MAS	marker assisted selection
Mb	megabases
mg	milligram
mg/kg	milligram per kilogram
mg/L	milligram per litre
MIPs	major intrinsic protein
mL	millilitre
Mn	manganese

Mo	molybdenum
mol/L	molar concentration
MS	Murashige-Skoog
mya	million years ago
n	number
N	nitrogen
NaAc	sodium acetate
NaCl	sodium chloride
NASC	Nottingham *Arabidopsis* Stock Centre
Ni	nickel
NIPs	NOD26-like intrinsic proteins
OSR	oilseed rape
P	phosphorus
PCA	principal component analyses
PCR	polymerase Chain Reaction
PIPs	plasma membrane intrinsic proteins
pKa	dissociation constant
PM	plasma membrane
PRKM	reads per kb per million aligned reads
PSIKO	Population Structure Inference Using Kernel-PCA and Optimization
QTL	quantitative trait loci
RFLP	restriction fragment length polymorphism
RG-I	rhamnogalacturonan-I
RG-II	rhamnogalacturonan-II
RIPR	renewable industrial products from rapeseed
RP	right primer
RT-PCR	reverse transcription- polymerase chain reaction
S	sulfur
SA	structured association
saf	second allele frequency
SD	standard deviation
Si	silicon
SIPs	small basic intrinsic proteins
SNP	single nucleotide polymorphism
SRAP	sequence related amplified polymorphisms

TALENs	transcription activator-like effector nucleases
T-DNA	transfer DNA
TIPs	tonoplast intrinsic proteins
TP1	time point one
TP2	time point two
TP3	time point three
TP4	time point four
UTR	untranslated region
UV	ultra violet
WHO	World Health Organisation
WT	wild type
ZFNs	zinc finger nucleases
Zn	zinc
µM	micro molar

References

Abbadi, A. and Leckband, G. (2011). Rapeseed breeding for oil content, quality and sustainability. *European Jorunal of Lipid Science and Technology*, 113, pp.1198–1206. [Online]. Available at: doi:10.1002/ejlt.201100063.

de Abreu-Neto, J. B. et al. (2013). Heavy metal-associated isoprenylated plant protein (HIPP): Characterization of a family of proteins exclusive to plants. *FEBS Journal*, 280 (7), pp.1604–1616. [Online]. Available at: doi:10.1111/febs.12159.

Adams, K. L. et al. (2003). Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proceedings of the National Academy of Sciences of the United States of America*, 100 (8), pp.4649–4654. [Online]. Available at: doi:10.1073/pnas. 0630618100.

Akamichi, N. N. et al. (2002). Compilation and characterization of a novel WNK family of protein kinases in *Arabiodpsis thaliana* with reference to circadian rhythms. *Bioscience, Biotechnology, and Biochemistry*, 66 (11), pp.2429–2436. [Online]. Available at: doi:10.1271/bbb.66.2429

Alcock, T. D. et al. (2017). Identification of candidate genes for calcium and magnesium accumulation in *Brassica napus* L. by association genetics. *Frontiers in Plant Science*, 8, 1968. [Online]. Available at: doi:10.3389/fpls.2017.01968.

Alcock, T. D. et al. (2018). Species-wide variation in shoot nitrate concentration, and genetic loci controlling nitrate, phosphorus and potassium accumulation in *Brassica napus L*. *Frontiers in Plant Science*, 9, 1487. [Online]. Available at: doi:10.3389/fpls.2018.01487.

Alexander, D. H., Novembre, J. and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19, pp.1655–1664. [Online]. Available at: doi:10.1101/gr.094052.109.vidual.

Alexandersson, E. et al. (2010). Transcriptional regulation of aquaporins in accessions of *Arabidopsis* in response to drought stress. *The Plant Journal*, 61 (4), pp.650–660. [Online]. Available at: doi:10.1111/j.1365-313X.2009.04087.x.

Allender, C. J. and King, G. J. (2010). Origins of the amphiploid species *Brassica napus L*. investigated by chloroplast and nuclear molecular markers. *BMC Plant Biology*, 10, 54. [Online]. Available at: doi:10.1186/1471-2229-10-54.

Alonso, J. M. et al. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science*, 301 (5633), pp.653–657. [Online]. Available at: doi:10.1126/science.1086391.

Altschup, S. F. et al. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215 (3), pp.403–410. [Online]. Available at: doi:10.1016/S0022-2836(05)80360-2.

Aquea, F. et al. (2012). A molecular framework for the inhibition of *Arabidopsis* root growth in response to boron toxicity. *Plant, Cell and Environment*, 35, pp.719–734. [Online]. Available at: doi:10.1111/j.1365-3040.2011.02446.x.

Armstrong, T et al. (2002). Long-term effects of boron supplementation on reproductive characteristics and bone mechanical properties in gilts. *Journal of Animal Science*, 80, pp.154–161.

Bancroft, I. et al. (2011). Dissecting the genome of the polyploid crop oilseed rape by transcriptome sequencing. *Nature Biotechnology*, 29 (8), pp.762–766. [Online]. Available at: doi:10.1038/nbt.1926.

Bao, F., Azhakanandam, S. and Franks, R. G. (2010). SEUSS and SEUSS-LIKE transcriptional adaptors regulate floral and embryonic development in *Arabidopsis*. *Plant Physiology*, 152 (2), pp.821–836. [Online]. Available at: doi:10.1016/0002-9416(81)90382-1.

Bar-Peled, M., Urbanowicz, B. R. and O'Neill, M. (2012). The Synthesis and origin of the pectic polysaccharide rhamnogalacturonan II - insights from nucleotide sugarformation and diversity. *Frontiers in Plant Science*, 3 (May), 92. [Online]. Available at: doi:10.3389/fpls.2012.00092.

Baxter, I. et al. (2007). Purdue ionomics information management system. An integrated functional genomics platform. *Plant Physiology*, 143 (2), pp.600–611. [Online]. Available at: doi:10.1104/pp.106.092528.

Baxter, I. et al. (2009). Root suberin forms an extracellular barrier that affects water relations and mineral nutrition in *Arabidopsis*. *PLoS Genetics*, 5 (5), e1000492. [Online]. Available at: doi:10.1371/journal.pgen.1000492.

Baxter, I. et al. (2010). A coastal cline in sodium accumulation in *Arabidopsis thaliana* is driven by natural variation of the sodium transporter *AtHKT1;1*. *PLoS Genetics*, 6 (11). e1001193 [Online]. Available at: oi:10.1371/journal.pgen. 1001193.

Baxter, I. (2010). Ionomics: The functional genomics of elements. *Briefings in Functional Genomics and Proteomics*, 9 (2), pp.149–156. [Online]. Available at: doi:10.1093/bfgp/elp055.

Baxter, I. et al. (2012). Biodiversity of mineral nutrient and trace element accumulation in *Arabidopsis thaliana*. *PLoS ONE*, 7 (4), e31521. [Online]. Available at: doi:10.1371/journal.pone.0035121.

Baxter, I. (2015). Should we treat the ionome as a combination of individual elements, or should we be deriving novel combined traits? *Journal of Experimental Botany*, 66 (8), pp.2127–2131. [Online]. Available at: doi:10.1093/jxb/erv040.

Baxter, I. R. et al. (2008). The leaf ionome as a multivariable system to detect a plant ’s physiological status. *Proceedings of the National Academy of Sciences of the United States of America*, 105 (33), pp.12081–12086.

Bayer, E. M. et al. (2006). *Arabidopsis* cell wall proteome defined using multidimensional protein identification technology. *Proteomics*, 6 (1), pp.301–311. [Online]. Available at: doi:10.1002/pmic.200500046.

Bayer, P. E. et al. (2017). Assembly and comparison of two closely related *Brassica napus* genomes. *Plant Biotechnology Journal*, 15, pp.1602–1610. [Online]. Available at: doi:10.1111/pbi.12742.

Beato, V. M. et al. (2014). Boron deficiency increases expressions of asparagine synthetase, glutamate dehydrogenase and glutamine synthetase genes in tobacco roots irrespective of the nitrogen source. *Soil Science and Plant Nutrition*, 60 (3), pp.314–324. [Online]. Available at: doi:10.1080/00380768.2014.881706.

Bellaloui, N., Smith, J. R. and Mengistu, A. (2017). Seed nutrition and quality, seed coat boron and lignin are influenced by delayed harvest in exotically-derived soybean breeding lines under high heat. *Frontiers in Plant Science*, 8 (September), 1563 [Online]. Available at: doi:10.3389/fpls.2017.01563.

Ben-Gal, A. and Shani, U. (2002). Yield, transpiration and growth of tomatoes under combined excess boron and salinity stress. *Plant and Soil*, 247 (2), pp.211–221. [Online]. Available at: doi:10.1023/A:1021556808595.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate : a practical and powerful spproach to multiple testing. *Journal of the Royal Statistical Society Series B*, 57 (1), pp.289–300.

Beauchamp, E. G. and Hussain, I. (1974). Brown heart in rutabagas frown on southern Ontario soils. *Canadian Journal of Plant Science*. 54, pp. 171-178.

Bienert, G. P. et al. (2008). A subgroup of plant aquaporins facilitate the bi-directional diffusion of $As(OH)^3$ and $Sb(OH)^3$ across membranes. *BMC Biology*, 6 (26). [Online]. Available at: doi:10.1186/1741-7007-6-26.

Blázquez, M. a et al. (1997). LEAFY expression and flower initiation in *Arabidopsis*. *Development*, 124 (19), pp.3835–3844. [Online]. Available at: doi:10.7150/ijbs.12020.

Blevins, D. G. and Lukaszewski, K. M. (1998). Boron in plant structure and function. *Annual Review of Plant Physiology and Plant Molecular Biology*, 49 (1), pp.481–500. [Online]. Available at: doi:10.1146/annurev.arplant.49.1.481.

Bogaard, A. et al. (2013). Crop manuring and intensive land management by Europe's first farmers. *Proceedings of the National Academy of Sciences of the United States of America*, 110 (31), pp.12589–12594. [Online]. Available at: doi:10.1073/pnas.1305918110.

Bogiani, J. C., Amaro, A. C. E. and Rosolem, C. A. (2013). Carbohydrate production and transport in cotton cultivars grown under boron deficiency. *Scientia Agricola*, 70 (6), pp.442–448. [Online]. Available at: doi:10.1590/S0103-90162013000600010.

Bolaños, L. et al. (2004). Why boron? *Plant Physiology and Biochemistry*, 42 (11), pp.907–912. [Online]. Available at: doi:10.1016/j.plaphy.2004.11.002.

Bonilla, I., Garciagonzalez, M. and Mateo, P. (1990). Boron requirement in cyanobacteria - its possible role in the early evolution of photosynthetic organisms. *Plant Physiology*, 94 (4), pp.1554–1560. [Online]. Available at: doi:10.1104/pp.94.4.1554.

Bortesi, L. and Fischer, R. (2015). The CRISPR / Cas9 system for plant genome editing and beyond. *Biotechnology Advances*, 33 (1), pp.41–52. [Online]. Available at: doi:10.1016/j.biotechadv.2014.12.006.

Braatz, J. et al. (2017). CRISPR-Cas9 targeted mutagenesis leads to simultaneous modification of different homoeologous gene copies in polyploid oilseed rape (*Brassica napus*). *Plant Physiology*, 174, pp.935–942. [Online]. Available at: doi:10.1104/pp.17.00426.

Bradbury, P. J. et al. (2007). TASSEL : software for association mapping of complex traits in diverse samples. *Bioinformatics*, 23 (19), pp.2633–2635. [Online]. Available at: doi:10.1093/bioinformatics/btm308.

Broadley, M. R. et al. (2004). Phylogenetic variation in the shoot mineral concentration of angiosperms. *Journal of Experimental Botany*, 55 (396), pp.321–336. [Online]. Available at: doi:10.1093/jxb/erh002.

Broadley, M. R. et al. (2008). Shoot calcium and magnesium concentrations Differ between subtaxa are highly heritable and associate with potentially pleiotropic loci in *Brassica oleracea*. *Plant Physiology*, 146 (4), pp.1707–1720. [Online]. Available at: doi:10.1104/pp.107.114645.

Brown, P. et al. (1999). Transgenically enhanced sorbitol synthesis facilitates phloem boron transport and increases tolerance of tobacco to boron deficiency. *Plant Physiology*, 119 (1), pp.17–20. [Online]. Available at: doi:10.1104/pp.119.1.17.

Brown, P. H. et al. (2002). Boron in plant biology. *Plant Biology*, 4 (2), pp.205–223. [Online]. Available at: doi:10.1055/s-2002-25740.

Brown, P. H. and Hu, H. (1996). Phloem mobility of boron is species dependent : evidence for phloem mobility in sorbitol-rich species. *Annals of Botany*, 77, pp.497–506. [Online]. Available at: doi:10.1006/anbo.1996.0060.

Brown, P. H. and Shelp, B. J. (1997). Boron mobility in plants. *Plant and Soil*, 193, pp.85–101. [Online]. Available at: doi:10.1023/A:1004211925160.

Burget, E. G. et al. (2003). The biosynthesis of L-arabinose in plants : molecular cloning and characterization of a golgi-localized UDP-D-xylose 4-epimerase encoded by the MUR4 gene of *Arabidopsis*. *The Plant Cell*, 15 (February), pp.523–531. [Online]. Available at: doi:10.1105/tpc.008425.response.

Bus, A. et al. (2011). Patterns of molecular variation in a species-wide germplasm set of *Brassica napus*. *Theoretical and Applied Genetics*, 123, pp.1413–1423. [Online]. Available at: doi:10.1007/s00122-011-1676-7.

Bus, A. et al. (2014). Species- and genome-wide dissection of the shoot ionome in *Brassica napus* and its relationship to seedling development. *Frontiers in Plant Science*, 5 (485), 485. [Online]. Available at: doi:10.3389/fpls.2014.00485.

Cai, D. et al. (2014). Association mapping of six yield-related traits in rapeseed. *Theoretical and Applied Genetics*, 127, pp.85–96. [Online]. Available at: doi:10.1007/s00122-013-2203-9

Callaway, E. (2016). EU law deals blow to CRISPR crops. *Nature*, 560, pp.2016.

Camacho-Cristobal, J. J. et al. (2015). Boron deficiency inhibits root cell elongation via an ethylene/auxin/ROS-dependent pathway in *Arabidopsis* seedlings. *Journal of Experimental Botany*, 66 (13), pp.3831–3840. [Online]. Available at: doi:10.1093/jxb/erv186.

Camacho-Cristóbal, J. J. and González-Fontes, A. (1999). Boron deficiency causes a drastic decrease in nitrate content and nitrate reductase activity, and increases the content of carbohydrates in leaves from tobacco plants. *Planta*, 209 (4), pp.528–536. [Online]. Available at: doi:10.1007/s004250050757.

Camacho-Cristóbal, J. J. and González-Fontes, A. (2007). Boron deficiency decreases plasmalemma H+-ATPase expression and nitrate uptake, and promotes ammonium assimilation into asparagine in tobacco roots. *Planta*, 226 (2), pp.443–451. [Online]. Available at: doi:10.1007/s00425-007-0494-2.

Camacho-Cristóbal, J. J., Rexach, J. and González-Fontes, A. (2008). Boron in plants: deficiency and toxicity. *Journal of Integrative Plant Biology*, 50 (10), pp.1247–1255. [Online]. Available at: doi:10.1111/j.1744-7909.2008.00742.x.

Cao-Pham, A. H. et al. (2018). Nudge-nudge, WNK-WNK (kinases), say no more? *New Phytologist*, 220 (1), pp.35–48. [Online]. Available at: doi:10.1111/nph.15276.

Cao, Z. et al. (2010). Analysis of QTLs for erucic acid and oil content in seeds on A8 chromosome and the linkage drag between the alleles for the two traits in *Brassica napus*. *Journal of Genetics and Genomics*, 37 (4), pp.231–240. [Online]. Available at: doi:10.1016/S1673-8527(09)60041-2.

Cernac, A. et al. (1997). The SAR1 gene of *Arabidopsis* acts downstream of the AXR1 gene in auxin response. *Development*, 124 (8), pp.1583–1591. [Online]. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9108374.

Chalhoub, B. et al. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science*, 345 (6199), pp.950–953. [Online]. Available at: doi:10.1126/science.1253435.

Chapin, R. E. et al. (1997). The effects of dietary boron on bone strength in rats. *Toxicological Sciences*, 35 (2), pp.205–215. [Online]. Available at: doi:10.1093/toxsci/35.2.205.

Chen, H. et al. (2018). Molecular characterization of the genome-wide BOR transporter gene family and genetic analysis of BnaC04.BOR1;1c in *Brassica napus*. *BMC Plant Biology*, 18 (193), pp.1–14. [Online]. Available at: doi.org/10.1186/s12870-018-1407-1

Chen, M. et al. (2014). Proteomic analysis of *Arabidopsis thaliana* leaves in response to acute boron deficiency and toxicity reveals effects on photosynthesis, carbohydrate metabolism, and protein synthesis. *Journal of Plant Physiology*, 171 (3–4), pp.235–242. [Online]. Available at: doi:10.1016/j.jplph.2013.07.008.

Chen, S. et al. (2008). Divergent patterns of allelic diversity from similar origins : the case of oilseed rape (Brassica napus L .) in China and Australia. *Genome*, 51, pp.1–10. [Online]. Available at: doi:10.1139/G07-095.

Chen, Y., Zou, T. and McCormick, S. (2016). *S*-adenosylmethionine aynthetase 3 is important for pollen tube growth. *Plant Physiology*, 172 (1), pp.244–253. [Online]. Available at: doi:10.1104/pp.16.00774.

Cheng, Y. et al. (2011). Characterization of the *Arabidopsis* glycerophosphodiester phosphodiesterase (GDPD) family reveals a role of the plastid-localized AtGDPD1 in maintaining cellular phosphate homeostasis under phosphate starvation. *The Plant Journal*, 66 (5), pp.781–795. [Online]. Available at: doi:10.1111/j.1365-313X.2011.04538.x.

Cheung, F. et al. (2009). Comparative analysis between homoeologous genome segments of *Brassica napus* and its progenitor species reveals extensive sequence-level divergence. *The Plant cell*, 21 (7), pp.1912–1928. [Online]. Available at: doi:10.1105/tpc.108.060376.

Choi, H.-S., Seo, M. and Cho, H.-T. (2018). Two TPL-binding motifs of ARF2 are involved in repression of auxin responses. *Frontiers in Plant Science*, 9 (March), 372. [Online]. Available at: doi:10.3389/fpls.2018.00372.

Choi, W. G. and Roberts, D. M. (2007). *Arabidopsis* NIP2;1, a major intrinsic protein transporter of lactic acid induced by anoxic stress. *Journal of Biological Chemistry*, 282 (33), pp.24209–24218. [Online]. Available at: doi:10.1074/jbc.M700982200.

Christian, M. et al. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*, 186, pp.757–761. [Online]. Available at: doi:10.1534/genetics.110.120717.

Collard, B. C. Y. et al. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica*, 142 (1–2), pp.169–196. [Online]. Available at: doi:10.1007/s10681-005-1681-5.

Dannel, F., Pfeffer, H. and Romheld, V. (2000). Characterization of root boron pools, boron uptake and boron translocation in sunflower using the stable isotopes ^{10}B and ^{11}B. *Australian Journal of Plant Physiology*, 27 (5), pp.397-405.

Delhaize, E. and Randall, P. (1995). Characterization of a phosphate-accumulator mutant of *Arabidopsis thaliana*. *Plant Physiology*, 107, pp.207–213. [Online]. Available at: doi:10.1104/pp.107.1.207

Dell, B. and Huang, L. (1997). Physiological response of plants to low boron. *Plant and Soil*, 193 (1), pp.103–120. [Online]. Available at: doi:10.1023/A:1004264009230.

Dermott, B. Y. W. and Trinder, N. (1947). Brown heart In swedes : a Cumbrian survey. *The Journal of Agricultural Science*, 37 (November), pp.152–155. [Online]. Available at: doi:10.1017/S0021859600083519.

Devlin, B. and Roeder, K. (1999). Genomic Control for Association Studies. *Biometrics*, 55 (December), pp.997–1004. [Online]. Available at: doi.org/10.1111/j.0006-341X.1999.00997.x

Dhakate, P. et al. (2017). Functional characterization of a novel *Brassica* LEAFY homolog from Indian mustard: expression pattern and gain-of-function studies. *Plant Science*, 258, pp.29–44. [Online]. Available at: doi:10.1016/j.plantsci.2017.02.003.

Diehn, T. A. et al. (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 (April), 166. [Online]. Available at: doi:10.3389/fpls.2015.00166.

Diehn, T. A. et al. (2019). Boron demanding tissues of *Brassica napus* express specific sets of functional Nodulin26-like Intrinsic Proteins and BOR1 transporters. *The Plant Jounal*, 99 (1), [Online]. Available at: doi:10.1111/tpj.14428.

Drinking Water Inspectorate. (2018). *Water Supply (Water Quality) Regulations*. pp.1-61.

Dumez, S. et al. (2006). Mutants of *Arabidopsis* lacking starch branching enzyme II substitute plastidial starch synthesis by cytoplasmic maltose accumulation. *The Plant Cell*, 18 (10), pp.2694–2709. [Online]. Available at: doi:10.1105/tpc.105.037671.

Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association*, 56 (293), pp.52–64. [Online]. Available at: doi:10.1080/01621459.1961.10482090.

Ecke, W. et al. (2010). Extent and structure of linkage disequilibrium in canola quality winter rapeseed (*Brassica napus L.*). *Theoretical and Applied Genetics*, 120 (5), pp.921–931. [Online]. Available at: doi:10.1007/s00122-009-1221-0.

Eggert, K. and von Wiren, N. (2013). Dynamics and partitioning of the ionome in seeds and germinating seedlings of winter oilseed rape. *Metallomics*, 5, pp.1316–1325. [Online]. Available at: doi:10.1039/c3mt00109a.

Eggert, K. and von Wirén, N. (2016). The role of boron nutrition in seed vigour of oilseed rape (*Brassica napus L.*). *Plant and Soil*, 402 (1–2), pp.63–76. [Online]. Available at: doi:10.1007/s11104-015-2765-1.

Eide, D. J. et al. (2005). Characterization of the yeast ionome: a genome-wide analysis of nutrient mineral and trace element homeostasis in *Saccharomyces cerevisiae*. *Genome Biology*, 6 (9), p.R77. [Online]. Available at: doi:10.1186/gb-2005-6-9-r77.

Eroglu, S. et al. (2016). The vacuolar manganese rransporter MTP8 determines tolerance to iron deficiency induced chlorosis in *Arabidopsis*. *Plant Physiology*, 170 (February), pp.1030–1045. [Online]. Available at: doi:10.1104/pp.15.01194.

European Commission. (2018). *EU Crops Market Observatory-Oilseeds and protein crops. Agriculture and Rural Development*. [Online]. Available at: https://ec.europa.eu/agriculture/market-observatory/crops/oilseeds-protein-crops/statistics_en

Fadhel, F. et al. (2013). Genotypic resistance to brown heart incidence in swede parent lines and F1 hybrids and the influence of applied boron. *The Journal of Agricultural Science*, 153 (2), pp.195–204. [Online]. Available at: doi:10.1017/S0021859613000889.

Fitzpatrick, K. L. and Reid, R. J. (2009). The involvement of aquaglyceroporins in transport of boron in barley roots. *Plant, Cell and Environment*, 32 (10), pp.1357–1365. [Online]. Available at: doi:10.1111/j.1365-3040.2009.02003.x.

Fleischer, a, O'Neill, M. and Ehwald, R. (1999). The pore size of non-graminaceous plant cell walls is rapidly decreased by borate ester cross-linking of the pectic polysaccharide rhamnogalacturonan II. *Plant Physiology*, 121 (3), pp.829–838. [Online]. Available at: doi:10.1104/pp.121.3.829.

Fort, D. J. et al. (1998). Adverse reproductive and developmental effects in *Xenopus* from insufficient boron. *Biological Trace Element Research*, 66 (1–3), pp.237–259. [Online]. Available at: doi:10.1007/BF02783141.

Francisco-Amorim, M. et al. (2018). The U1 snRNP subunit LUC7 modulates plant development and stress responses via regulation of alternative splicing. *The Plant Cell*, 30 (November), pp.2838-2854. [Online]. Available at: doi:10.1105/tpc.18.00244.

Geldner, N. (2013). The endodermis. *Annual Review of Plant Biology*, 64, pp.531–558. [Online]. Available at: doi:10.1146/annurev-arplant-050312-120050.

Ghandilyan, A. et al. (2009). Genetic analysis identifies quantitative trait loci controlling rosette mineral concentrations in *Arabidopsis thaliana* under drought. *New Phytologist*, 184, pp.180–192. [Online]. Available at: doi: 10.1111/j.1469-8137.2009.02953.x

Gille, S. et al. (2009). Identification of plant cell wall mutants by means of a forward chemical genetic approach using hydrolases. *Proceedings of the National Academy of Sciences of the United States of America*, 106 (34), pp.14699–14704. [Online]. Available at: doi:10.1073/pnas.0905434106.

Goldbach, H. E. and Wimmer, M. a. (2007). Boron in plants and animals: is there a role beyond cell-wall structure? *Journal of Plant Nutrition and Soil Science*, 170 (1), pp.39–48. [Online]. Available at: doi:10.1002/jpln.200625161.

Goldberg, S. (1997). Reactions of boron with soils. *Plant and Soil*, 193, pp.35–48. [Online]. Available at: doi:10.1023/A:1004203723343.

Goodwin, S., McPherson, J. D. and McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, 17 (6), pp.333–351. [Online]. Available at: doi:10.1038/nrg.2016.49.

Gou, M. et al. (2009). An F-box gene , CPR30 , functions as a negative regulator of the defense response in *Arabidopsis*. *The Plant Journal*, 60, pp.757–770. [Online]. Available at: doi:10.1111/j.1365-313X.2009.03995.x.

Guilfoyle, T. J. and Hagen, G. (2007). Auxin response factors. *Current Opinion in Plant Biology*, 10 (5), pp.453–460. [Online]. Available at: doi:10.1016/j.pbi.2007.08.014.

Gupta, U. and Cutcliffe, J. (1978). Effects of methods of boron application on leaf tissue concentration of boron and control of brown-heart in rutabaga. *Candanian Journal of Plant Science*, 58, pp.63–68.

Hadas, A and Hagin, J. (1972). Boron adsorption by soils as influenced by potassium. *Soil Science,* 113 (2), pp. 189-193.

Halkier, B. A. and Gershenzon, J. (2006). Biology and Biochemistry of Glucosinolates. *Annual Review of Plant Biology*, 57 (1), pp.303–333. [Online]. Available at: doi:10.1146/annurev.arplant.57.032905.105228.

Han, S. et al. (2007). Boron deficiency decreases growth and photosynthesis, and increases starch and hexoses in leaves of citrus seedlings. *Journal of Plant Physiology*, 165 (13), pp.1331–1341. [Online]. Available at: doi:10.1016/j.jplph.2007.11.002.

Han, S. et al. (2009). CO2 assimilation, photosystem II photochemistry, carbohydrate metabolism and antioxidant system of citrus leaves in response to boron stress. *Plant Science*, 176 (1), pp.143–153. [Online]. Available at: doi:10.1016/j.plantsci.2008.10.004.

Harholt, J., Suttangkakul, A. and Vibe Scheller, H. (2010). Biosynthesis of pectin. *Plant Physiology*, 153 (2), pp.384–395. [Online]. Available at: doi:10.1104/pp.110.156588.

Harper, A. L. et al. (2012). Associative transcriptomics of traits in the polyploid crop species *Brassica napus*. *Nature Biotechnology*, 30 (8), pp.798–802. [Online]. Available at: doi:10.1038/nbt.2302.

Harper, A. L. et al. (2016). Molecular markers for tolerance of European ash (*Fraxinus excelsior*) to dieback disease identified using associative transcriptomics. *Scientific Reports*, 6 (December), 19335. [Online]. Available at: doi:10.1038/srep19335.

Havlickova, L. et al. (2018). Validation of an updated associative transcriptomics platform for the polyploid crop species *Brassica napus* by dissection of the genetic architecture of erucic acid and tocopherol isoform variation in seeds. *The Plant Journal*, 93, pp.181–192. [Online]. Available at: doi:10.1111/tpj.13767.

He, Z. et al. (2015). Construction of *Brassica* A and C genome-based ordered pan-transcriptomes for use in rapeseed genomic research. *Data in Brief*, 4, pp.357–362. [Online]. Available at: doi:10.1016/j.dib.2015.06.016.

He, Z. et al. (2016). Extensive homoeologous genome exchanges in allopolyploid crops revealed by mRNAseq-based visualization. *Plant Biotechnology Journal*, 15, pp.594–604. [Online]. Available at: doi:10.1111/pbi.12657.

Hejna, O. et al. (2019). Analysing the genetic architecture of clubroot resistance variation in Brassica napus by associative transcriptomics. *Molecular Breeding*, 39 (112). [Online]. Available at: doi:10.1007/s11032-019-1021-4.

Helper, P. (2005). Calcium: a central regulator of plant growth and development. *The Plant Cell*. 17, pp. 2142-2155. [Online]. Available at: doi:10.1105/tpc. 105.032508

Higgins, J. et al. (2012). Use of mRNA-seq to discriminate contributions to the transcriptome from the constituent genomes of the polyploid crop species *Brassica napus*. *BMC Genomics*, 13 (1). [Online]. Available at: doi:10.1186/1471-2164-13-247.

Hirai, M. Y. et al. (2007). Omics-based identification of *Arabidopsis* MYB transcription factors regulating aliphatic glucosinolate biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, 104 (15), pp.6478–6483. [Online]. Available at: doi:10.1073/pnas.0611629104.

Hong-Hermesdorf, A. et al. (2006). A WNK kinase binds and phosphorylates V-ATPase subunit C. *FEBS Letters*, 580 (3), pp.932–939. [Online]. Available at: doi:10.1016/j.febslet.2006.01.018.

Hosmani, P. et al. (2013). Dirigent domain-containing protein is part of the machinery required for formation of the lignin-based Casparian strip in the root. *Journal of Biological Chemistry*, 110 (35), pp.14498–14503. [Online]. Available at: doi:10.1073/pnas.1316463110.

Hruz, T. et al. (2008). Genevestigator V3 : a reference expression database for the meta-analysis of transcriptomes. *Advances in Bioinformatics*, e420747. [Online]. Available at: doi:10.1155/2008/420747.

Hu, H. and Brown, P. H. (1997). Absorption of boron by plant roots. *Plant and Soil*, 193 (1), pp.49–58. [Online]. Available at: doi:10.1023/A:1004255707413.

Hu, H., Brown, P. H. and Labavitch, J. M. (1996). Species variability in boron requirement is correlated with cell wall pectin. *Journal of Experimental Botany*, 47 (2), pp.227–232. [Online]. Available at: doi:10.1093/jxb/47.2.227.

Hua, Y. et al. (2016a). Transcriptomics-assisted quantitative trait locus fine mapping for the rapid identification of a nodulin 26-like intrinsic protein gene regulating boron efficiency in allotetraploid rapeseed. *Plant Cell and Environment*, 39 (7), pp.1601–1618. [Online]. Available at: doi:10.1111/pce.12731.

Hua, Y. et al. (2016b). Physiological, genomic and transcriptional diversity in responses to boron deficiency in rapeseed genotypes. *Journal of Experimental Botany*, 67 (19), pp.5769–5784. [Online]. Available at: doi:10.1093/jxb/erw342.

Hua, Y. et al. (2017). Genome-scale mRNA transcriptomic insights into the responses of oilseed rape (*Brassica napus L.*) to varying boron availabilities. *Plant and Soil*, 416 (1–2), pp.205–225. [Online]. Available at: doi:10.1007/s11104-017-3204-2.

Huang, C. K. et al. (2016). The DEAD-box RNA helicase AtRH7/PRH75 participates in pre-rRNA processing, plant development and cold tolerance in *Arabidopsis*. *Plant and Cell Physiology*, 57 (1), pp.174–191. [Online]. Available at: doi:10.1093/pcp/pcv188.

Huang, L., Bell, R. W. and Dell, B. (2008). Evidence of phloem boron transport in response to interrupted boron supply in white lupin (*Lupinus albus L.* cv. *Kiev* Mutant) at the reproductive stage. *Journal of Experimental Botany*, 59 (3), pp.575–583. [Online]. Available at: doi:10.1093/jxb/erm336.

Huang, L., Ye, Z. and Bell, R. (1996). The importance of sampling immature leaves for the diagnosis of boron deficiency in oilseed rape (*Brassica napus cv. Eureka*). *Plant and soil*, 183, pp.187–198. [Online]. Available at: doi:10.1007/BF00011434.

Huang, X. Y. and Salt, D. E. (2016). Plant ionomics: from elemental profiling to environmental adaptation. *Molecular Plant*, 9 (6), pp.787–797. [Online]. Available at: doi:10.1016/j.molp.2016.05.003.

Hunt, C. D. (2003). Dietary doron: an overview of the evidence for its role in immune function. *Journal of Trace Elements in Experimental Medicine*, 16 (4), pp.291–306. [Online]. Available at: doi:10.1002/jtra.10041.

Hurst, R. and MacLeod, D. (1936). Turnip brown-heart. *Scientia Agricola*, 17, pp.209–214.

Hussain, M. J. et al. (2012). Effects of nitrogen and boron on the yield and hollow stem disorder of broccoli (*Brassica oleracea var . italica*). *The Agriculturists*, 10 (2), pp.36–45.

Hwang, J. et al. (2016). Plant ABC transporters enable many unique aspects of a terrestrial plant's lifestyle. *Molecular Plant*, 9 (3), pp.338–355. [Online]. Available at: doi:10.1016/j.molp.2016.02.003.

Inaba, R. and Nishio, T. (2002). Phylogenetic analysis of *Brassiceae* based on the nucleotide sequences of the S-locus related gene, SLR1. *Theoretical and Applied Genetics*, 105 (8), pp.1159–1165. [Online]. Available at: doi:10.1007/s00122-002-0968-3.

Da Ines, O. et al. (2010). Kinetic analyses of plant water relocation using deuterium as tracer - reduced water flux of *Arabidopsis pip2* aquaporin knockout mutants. *Plant Biology*, 12, pp.129–139. [Online]. Available at: doi:10.1111/j.1438-8677.2010.00385.x.

Irshad, M. et al. (2008). A new picture of cell wall protein dynamics in elongating cells of *Arabidopsis thaliana*: confirmed actors and newcomers. *BMC Plant Biology*, 8, 94. [Online]. Available at: doi:10.1186/1471-2229-8-94.

Isayenkov, S. V. and Maathuis, F. J. M. (2008). The *Arabidopsis thaliana* aquaglyceroporin AtNIP7;1 is a pathway for arsenite uptake. *FEBS Letters*, 582 (11), pp.1625–1628. [Online]. Available at: doi:10.1016/j.febslet.2008.04.022.

Ishii, T. et al. (1999). The plant cell wall polysaccharide RG-II self-assembles into a covalently cross-linked Dimer. *Biochemistry*, 274 (19), pp.13098–13104.

Jankowski, K. J. et al. (2016). Yield and quality of winter oilseed rape (Brassica napus L.) seeds in response to foliar application of boron. *Agricultural and Food Science*, 25 (3), pp.164–176.

Jinek, M. et al. (2012). A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337 (August), pp.816–822.

Johnson, C. S., Kolevski, B. and Smyth, D. R. (2002). TRANSPARENT TESTA GLABRA2 , a trichome and seed coat development gene of *Arabidopsis* Encodes a WRKY transcription factor. *The Plant Cell*, 14 (June), pp.1359–1375. [Online]. Available at: doi:10.1105/tpc.001404.covered.

Kamiya, T. et al. (2015). The MYB36 transcription factor orchestrates Casparian strip formation. *Proceedings of the National Academy of Sciences of the United States of America*, 112 (33), pp.10533–10538. [Online]. Available at: doi:10.1073/pnas.1507691112.

Kasai, K. et al. (2011). High boron-induced ubiquitination regulates vacuolar sorting of the BOR1 borate transporter in *Arabidopsis thaliana*. *Journal of Biological Chemistry*, 286 (8), pp.6175–6183. [Online]. Available at: doi:10.1074/jbc.M110.184929.

Kasajima, I. et al. (2010). WRKY6 is involved in the response to boron deficiency in *Arabidopsis thaliana*. *Physiologia Plantarum*, 139 (1), pp.80–92. [Online]. Available at: doi:10.1111/j.1399-3054.2010.01349.x.

Kim, Y-G., Cha, J. and Chandrasegaran, S. (1996). Hybrid restriction enzymes : zinc finger fusions to *Fok* I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America*, 93, pp.1156–1160. [Online]. Available at: doi:10.1073/pnas.93.3.1156

Kobayashi, M., Matoh, T. and Azuma, J. (1996). Two chains of rhamnogalacturonan-II are cross-linked by borate-diol ester bonds in higher Plant Cell Walls. *Plant Physiology*, 110, pp.1017–1020.

Kobayashi, M., Mutoh, T. and Matoh, T. (2004). Boron nutrition of cultured tobacco BY-2 cells. IV. Genes induced under low boron supply. *Journal of Experimental Botany*, 55 (401), pp.1441–1443. [Online]. Available at: doi:10.1093/jxb/erh142.

Koch, M., Haubold, B. and Mitchell-Olds, T. (2001). Molecular systematics of the *Brassicaceae*: evidence from coding plastidic matK and nuclear Chs sequences. *American Journal of Botany*, 88 (3), pp.534–544. [Online]. Available at: doi:10.2307/2657117.

Koh, J. C. O. et al. (2017). A multiplex PCR for rapid identification of *Brassica* species in the triangle of U. *Plant Methods*, 13, 49. [Online]. Available at: doi:10.1186/s13007-017-0200-8.

Korte, A. and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods*, 9, 29. [Online]. Available at: doi:10.1186/1746-4811-9-29

Krysan, P. J., Young, J. C. and Sussman, M. R. (1999). T-DNA as an insertional mutagen in *Arabidopsis*. *The Plant Cell*, 11, pp.2283–2290. [Online]. Available at: doi.org/10.1105/tpc.11.12.2283

Kühn-institut, J. and Kühn-institut, J. (2005). Relation between total sulphur analysed by ICP-AES and glucosinolates in oilseed rape and Indian Mustard seeds. *Landbauforschung Völkenrode*, 55 (December), pp.205–210.

Kumar, K. et al. (2014). Two rice plasma membrane intrinsic proteins, OsPIP2;4 and OsPIP2;7, are involved in transport and providing tolerance to boron toxicity. *Planta*, 239 (1), pp.187–198. [Online]. Available at: doi:10.1007/s00425-013-1969-y.

Kurowska, M. and Daszkowska-golec, A. (2011). TILLING - a shortcut in functional genomics. *Applied Genetics*, 52, pp.371–390. [Online]. Available at: doi:10.1007/s13353-011-0061-1.

Lahner, B. et al. (2003). Genomic scale profiling of nutrient and trace elements in *Arabidopsis* thaliana. *Nature Biotechnology*, 21 (10), pp.1215–1221. [Online]. Available at: doi:10.1038/nbt865.

Lamesch, P. et al. (2012). The *Arabidopsis* information resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research*, 40, pp.1202–1210. [Online]. Available at: doi:10.1093/nar/gkr1090.

Lanoue, L. et al. (1998). Assessing the effects of low boron diets on embryonic and fetal development in rodents using in vitro and in vivo model systems. *Biological Trace Element Research*, 66 (1–3), pp.271–298. [Online]. Available at: doi:10.1007/BF02783143.

Lee, H. et al. (2000). The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in *Arabidopsis*. *Genes and Development*, 14, pp.2366–2376. [Online]. Available at: doi:10.1101/gad.813600.placed.

Lee, J. et al. (2008). SOC1 translocated to the nucleus by interaction with AGL24 directly regulates LEAFY. *The Plant Journal*, 55 (5), pp.832–843. [Online]. Available at: doi:10.1111/j.1365-313X.2008.03552.x.

Lee, J. E. et al. (2014). SEUSS and SEUSS-LIKE 2 coordinate auxin distribution and KNOXI activity during embryogenesis. *The Plant Journal*, 80 (1), pp.122–135. [Online]. Available at: doi:10.1111/tpj.12625.

Lee, M. et al. (2005). Methionine and threonine synthesis are limited by homoserine availability and not the activity of homoserine kinase in *Arabidopsis thaliana*. *The Plant Journal*, 41 (5), pp.685–696. [Online]. Available at: doi:10.1111/j.1365-313X.2004.02329.x.

Lemarchand, D. et al. (2000). The influence of rivers on marine boron isotopes and implications for reconstructing past ocean pH. *Nature*, 408 (6815), pp.951–954. [Online]. Available at: doi:10.1038/35050058.

Lewis, J. and Chen, C. H. (1976). Effects of boron deficiency on the chemical composition of a marine diatom. *Journal of Experimental Biology*, 27 (5), pp.916–921.

Leyser, H. M. O. et al. (1993). *Arabidopsis* auxin-resistance gene AXR1 encodes a protein related to ubiquitin-activating enzyme E1. *Nature*, 364 (6433), pp.161–164. [Online]. Available at: doi:10.1038/364161a0.

Li, G., Santoni, V. and Maurel, C. (2014). Plant aquaporins: roles in plant physiology. *Biochimica et Biophysica Acta*, 1840 (5), pp.1574–1582. [Online]. Available at: doi:10.1016/j.bbagen.2013.11.004.

Li, H., Ruan, J. and Durbin, R. (2008). Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Research*, 18, pp.1851–1858. [Online]. Available at: doi:10.1101/gr.078212.108.

Li, T. et al. (2011). *Arabidopsis thaliana* NIP7;1: an anther-specific boric acid transporter of the aquaporin superfamily regulated by an unusual tyrosine in helix 2 of the transport pore. *Biochemistry*, 50 (31), pp.6633–6641. [Online]. Available at: doi:10.1021/bi2004476.

Lin, C., Choi, H. S. and Cho, H. T. (2011). Root hair-specific expansin A7 is required for root hair elongation in A*rabidopsis*. *Molecules and Cells*, 31 (4), pp.393–397. [Online]. Available at: doi:10.1007/s10059-011-0046-2.

Lin, W.-D. et al. (2011). Coexpression-based clustering of *Arabidopsis* root genes predicts functional modules in early phosphate deficiency signaling. *Plant Physiology*, 155 (3), pp.1383–1402. [Online]. Available at: doi:10.1104/pp.110.166520.

Lindsay, E. R. and Maathuis, F. J. M. (2016). *Arabidopsis thaliana* NIP7;1 is involved in tissue arsenic distribution and tolerance in response to arsenate. *FEBS Letters*, 590 (6), pp.779–786. [Online]. Available at: doi:10.1002/1873-3468.12103.

Lipka, A. E. et al. (2012). GAPIT: Genome association and prediction integrated tool. *Bioinformatics*, 28 (18) pp. 2397-2399 [Online]. Available at: doi:10.1093/bioinformatics/bts444.

Liu, D. Y. T. et al. (2015). Characterisation of *Arabidopsis* calnexin 1 and calnexin 2 in the endoplasmic reticulum and at plasmodesmata. *Protoplasma*, 254 (1), pp.125–136. [Online]. Available at: doi:10.1007/s00709-015-0921-3.

Liu, J. et al. (2009). Analysis of genetic factors that control shoot mineral concentrations in rapeseed (*Brassica napus*) in different boron environments. *Plant and Soil*, 320 (1–2), pp.255–266. [Online]. Available at: doi:10.1007/s11104-009-9891-6.

Liu, Y., Tabata, D. and Imai, R. (2016). A cold-inducible DEAD-box RNA helicase from *Arabidopsis thaliana* regulates plant growth and development under low temperature. *PLoS ONE*, 11, 4. [Online]. Available at: doi:10.1371/journal.pone.0154040.

Liu, Z. et al. (2017). Comparative analysis of root traits and the associated QTLs for maize seedlings grown in paper roll, hydroponics and cermiculite culture system. *Frontiers in Plant Science*, 8 (March), 436. [Online]. Available at: doi:10.3389/fpls.2017.00436.

Lukens, L. et al. (2004). Genome redundancy and plasticity within ancient and recent *Brassica* crop species. *Biological Journal of the Linnean Society*, 82, pp.665–674. [Online]. Available at: doi:10.1111/j.1095-8312.2004.00352.x.

Macho-Rivero, M. A. et al. (2018). Boron toxicity reduces water transport from root to shoot in *Arabidopsis* plants. Evidence for a reduced transpiration rate and expression of major PIP aquaporin genes. *Plant and Cell Physiology*, 59 (4), pp.836–844. [Online]. Available at: doi:10.1093/pcp/pcy026.

Maillard, A. et al. (2015). Leaf mineral nutrient remobilization during leaf senescence and modulation by nutrient deficiency. *Frontiers in Plant Science*, 6 (May), 317. [Online]. Available at: doi:10.3389/fpls.2015.00317.

Marin, E. et al. (2010). miR390 , *Arabidopsis* TAS3 tasiRNAs , and their AUXIN RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth. *The Plant Cell*, 22, pp.1104–1117. [Online]. Available at: doi:10.1105/tpc.109.072553.

Marschner, H. (1995). Functions of Mineral Nutrients: Micronutrients 9. In: *Mineral Nutrition of Higher Plants (Second Edition)*. Academic Press Inc. pp.313–404. [Online]. Available at: doi:http://dx.doi.org/10.1016/B978-012473542-2/50011-0.

Martín-Rejano, E. M. et al. (2011). Auxin and ethylene are involved in the responses of root system architecture to low boron supply in *Arabidopsis* seedlings. *Physiologia Plantarum*, 142 (2), pp.170–178. [Online]. Available at: doi:10.1111/j.1399-3054.2011.01459.x.

Matoh, T., Kawaguchi, S. and Kobayashi, M. (1996). Ubiquity of a borate-rhamnogalacturonan II complex in the cell walls of higher plants. *Plant and Cell Physiology*, 37 (5), pp.636–640. [Online]. Available at: doi:10.1093/oxfordjournals.pcp.a028992.

Maurel, C. and Chrispeels, M. (2001). Aquaporins. A molecular entry into plant water relations. *Plant Physiology*, 125 (1), pp.135–138. [Online]. Available at: doi:10.1104/pp.125.1.135.

Mccallum, C. M. et al. (2000). Targeting induced local lesions in genomes (TILLING) for plant functional genomics. *Plant Physiology*, 123, pp.439–442. [Online]. Available at:doi.org/10.1104/pp.123.2.439

McDonald, G. K., Eglinton, J. K. and Barr, A. R. (2010). Assessment of the agronomic value of QTL on chromosomes 2H and 4H linked to tolerance to boron toxicity in barley (*Hordeum vulgare* L.). *Plant and Soil*, 326 (1), pp.275–290. [Online]. Available at: doi:10.1007/s11104-009-0006-1.

McLachlan. (1977). Effects of nutrients on growth and development of embryos of *Fucus edentatus*. *Phycologia*, 16, pp. 329-338

Mehmood, A. et al. (2018). Effect of plant density, boron nutrition and growth regulation on seed mass, emergence and offspring growth plasticity in cotton. *Scientific Reports*, 8, 7953 [Online]. Available at: doi:10.1038/s41598-018-26308-5.

Miura, A., Yonebayashi, S. and Watanabe, K. (2001). Mobilization of transposons by a mutation abolishing full DNA methylation in *Arabidopsis*. *Nature*, 411 (May), pp.212–214. [Online]. Available at: doi:10.1038/35075612

Miwa, K. et al. (2007). Plants tolerant of high boron levels. *Science*, 318, p.1417. [Online]. Available at: doi:10.1126/science.1146634.

Miwa, K. et al. (2013). Roles of BOR2, a boron exporter, in cross linking of rhamnogalacturonan II and root elongation under boron limitation in *Arabidopsis*. *Plant Physiology*, 163 (4), pp.1699–1709. [Online]. Available at: doi:10.1104/pp.113.225995.

Miwa, K., Aibara, I. and Fujiwara, T. (2014). *Arabidopsis thaliana* BOR4 is upregulated under high boron conditions and confers tolerance to high boron. *Soil Science and Plant Nutrition*, 60 (3), pp.349–355. [Online]. Available at: doi:10.1080/00380768.2013.866524.

Miwa, K., Takano, J. and Fujiwara. (2006). Improvement of seed yields under boron-limiting conditions through overexpression of BOR1, a boron transporter for xylem loading, in *Arabidopsis thaliana*. *The Plant Jounal*, 46, pp.1084–1091. [Online]. Available at: doi:10.1111/j.1365-313X.2006.02763.x

Morrell, P. L., Buckler, E. S. and Ross-ibarra, J. (2011). Crop genomics: advances and applications. *Nature Reviews Genetics*, 13 (2), pp.85–96. [Online]. Available at: doi:10.1038/nrg3097.

Mosa, K. A. et al. (2016). Enhanced boron tolerance in plants mediated by bidirectional transport through plasma membrane intrinsic proteins. *Scientific Reports*, 6, 21640. [Online]. Available at: doi:10.1038/srep21640.

Müller, B. et al. (2015). Amino acid export in developing *Arabidopsis* seeds depends on UmamiT facilitators. *Current Biology*, 25 (23), pp.3126–3131. [Online]. Available at: doi:10.1016/j.cub.2015.10.038.

Murashige, T. and Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacoo tissue cultures. *Physiologia Plantarum*, 15 (3), pp.473–497. [Online]. Available at: doi:10.1111/j.1399-3054.1962.tb08052.x.

Nable, R. O., Bañuelos, G. S. and Paull, J. G. (1997). Boron toxicity. *Plant and Soil*, 193, pp.181–198.

Nesi, N. et al. (2008). Genetic and molecular approaches to improve nutritional value of *Brassica napus* L. seed. *Comptes Rendus - Biologies*, 331 (10), pp.763–771. [Online]. Available at: doi:10.1016/j.crvi.2008.07.018.

Neugebauer, K. et al. (2016). Variation in the angiosperm ionome. *Physiologia Plantarum*, 163 (3), pp.306–322. [Online]. Available at: doi:10.1111/ppl.12700.

Noh, B. et al. (2001). Multidrug resistance-like genes of *Arabidopsis* required for auxin transport and auxin-mediated development. *The Plant Cell*, 13 (11), pp.2441–2454. [Online]. Available at: doi:https://doi.org/10.1016/j.metabol. 2015.08.014.

Notaguchi, M., Wolf, S. and Lucas, W. (20012). Phloem-mobile Aux/IAA transcripts target to the root tip and modify root architecture. *Journal of Integrative Plant Biology*, 54 (4), pp.760–772. [Online]. Available at: doi:10.1243/1748006XJRR226.

O'Neill, M. A. et al. (1996). Rhamnogalacturonan-II, a pectic polysaccharide in the walls of growing plant cell, forms a dimer that is covalently cross-linked by a borate ester. In vitro conditions for the formation and hydrolysis of the dimer. *Journal of Biological Chemistry*, 271 (37), pp.22923–22930. [Online]. Available at: doi:10.1074/jbc.271.37.22923.

O'Neill, M. A. et al. (2004). Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. *Annual Review of Plant Biology*, 55 (1), pp.109–139. [Online]. Available at: doi:10.1146/annurev.arplant.55.031903.141750.

O'Neill, M. A. et al. (2001). Requirement of borate cross-linking of cell wall rhamnogalacturonan II for *Arabidopsis* growth. *Science*, 294 (5543), pp.846–849. [Online]. Available at: doi:10.1126/science.1062319.

Oertli, J. J. (1993). The mobility of boron in plants. *Plant and Soil*, 155–156 (1), pp.301–304. [Online]. Available at: doi:10.1007/BF00025042.

Okushima, Y. et al. (2005). AUXIN RESPONSE FACTOR 2 (ARF2): A pleiotropic developmental regulator. *The Plant Journal*, 43 (1), pp.29–46. [Online]. Available at: doi:10.1111/j.1365-313X.2005.02426.x.

Osanai, T. et al. (2017). ACR11 is an activator of plastid-type glutamine synthetase GS2 in *Arabidopsis thaliana*. *Plant and Cell Physiology*, 58 (4), pp.650–657. [Online]. Available at: doi:10.1093/pcp/pcx033.

Outten, C. and O'Halloran, T. (2001). Femtomolar sensitivity of metallooregulatory protein controlling zinc homeostasis. *Science*, 292 (June), pp.2488–2492. [Online]. Available at: doi:10.1126/science.1060331.

Pabst, M. et al. (2013). Rhamnogalacturonan II structure shows variation in the side chains monosaccharide composition and methylation status within and across different plant species. *The Plant Journal*, 76 (1), pp.61–72. [Online]. Available at: doi:10.1111/tpj.12271.

Park, M. et al. (2004). NaBC1 is a ubiquitous electrogenic Na^+-coupled borate transporter essential for cellular boron homeostasis and cell growth and proliferation. *Molecular Cell*, 16 (3), pp.331–341. [Online]. Available at: doi:10.1016/j.molcel.2004.09.030.

Parry, M. A. J. et al. (2009). Mutation discovery for crop improvement. *Journal of Experimental Botany*, 60 (10), pp.2817–2825. [Online]. Available at: doi:10.1093/jxb/erp189.

Pastore, J. J. et al. (2011). LATE MERISTEM IDENTITY2 acts together with LEAFY to activate APETALA1. *Development*, 138 (15), pp.3189–3198. [Online]. Available at: doi:10.1242/dev.063073.

Peng, L. et al. (2012). Transcriptional profiling reveals adaptive responses to boron deficiency stress in *Arabidopsis*. *Zeitschrift für Naturforschung*, 67, pp. 510-524. [Online]. Available at: doi:10.5560/ZNC.2012.67c0510.

Phukan, U. J., Jeena, G. S. and Shukla, R. K. (2016). WRKY transcription factors: molecular regulation and stress responses in plants. *Frontiers in Plant Science*, 7 (June), 760. [Online]. Available at: doi:10.3389/fpls.2016.00760.

Pinson, S. R. M. et al. (2015). Worldwide genetic diversity for mineral element concentrations in rice grain. *Crop Science*, 55 (1), pp. 294-311. [Online]. Available at: doi:10.2135/cropsci2013.10.0656.

Pommerrenig, B. et al. (2018). Identification of rapeseed (*Brassica napus*) cultivars with a high tolerance to boron-deficient conditions. *Frontiers in Plant Science*, 9 (August), 1142. [Online]. Available at: doi:10.3389/fpls.2018.01142.

Popescu, A. et al. (2014). A novel and fast approach for population structure inference using kernel-PCA and optimization. *Genetics*, 198 (4), pp.1421–1431. [Online]. Available at: doi:10.1534/genetics.114.171314.

Postaire, O. et al. (2010). A PIP1 aquaporin contributes to hydrostatic pressure-induced water transport in both the root and rosette of *Arabidopsis*. *Plant Physiology*, 152 (3), pp.1418–1430. [Online]. Available at: doi:10.1104/pp.109.145326.

Price, A. . et al. (2010). New approaches to population stratification in genome-wide association studies. *National Review of Genetics*, 11 (7), pp.459–463. [Online]. Available at: doi:10.1038/nrg2813.New.

Price, A. L. et al. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, 38 (8), pp.904–909. [Online]. Available at: doi:10.1038/ng1847.

Princi, M. P. et al. (2015). Boron toxicity and tolerance in plants*:* recent advances and future perspectives. In: *Plant Metal Interaction: Emerging Remediation Techniques.* Elsevier, pp. 115-147. [Online]. Available at: doi:10.1016/B978-0-12-803158-2.00005-9.

Pritchard, J. K., Stephens, M. and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics Society of America*, 155, pp.945–959.

Pröfrock, D. and Prange, A. (2012). Inductively coupled plasma-mass spectrometry (ICP-MS) for quantitative analysis in environmental and life sciences: a review of challenges, solutions, and trends. *Applied Spectroscopy*, 66 (8), pp.843–868. [Online]. Available at: doi:10.1366/12-06681.

Puvanesarajah, V., Darvill, A. G. and Albersheim, P. (1991). Structural characterization of two oligosaccharide fragments formed by the selective cleavage of rhamnogalacturonan II: evidence for the anomeric configuration and attachment sites of apiose and 3-deoxy-2-heptulosaric acid. *Carbohydrate Research*, 218, pp.211–222. [Online]. Available at: doi:10.1016/0008-6215(91)84099-Z.

Qin, Y. et al. (2009). Penetration of the stigma and style elicits a novel transcriptome in pollen tubes, pointing to genes critical for growth in a pistil. *PLoS Genetics*, 5 (8), e1000621. [Online]. Available at: doi:10.1371/journal.pgen. 1000621.

Quirino, B. F., Reiter, W. and Amasino, R. D. (2001). One of two tandem *Arabidopsis* genes homologous to monosaccharide transporters is senescence-associated. *Plant Molecular Biology*, 46, pp.447–457. [Online]. Available at: doi:10.1023/A:1010639015959

Reid, R. (2007). Identification of boron transporter genes likely to be responsible for tolerance to boron toxicity in wheat and barley. *Plant and Cell Physiology*, 48 (12), pp.1673–1678. [Online]. Available at: doi:10.1093/pcp/pcm159.

Reid, R. (2014). Understanding the boron transport network in plants. *Plant and Soil*, 385 (1–2), pp.1–13. [Online]. Available at: doi:10.1007/s11104-014-2149-y.

Reid, R. J. et al. (2004). A critical analysis of the causes of boron toxicity in plants. *Plant, Cell and Environment*, 27 (11), pp.1405–1414. [Online]. Available at: doi:10.1111/j.1365-3040.2004.01243.x.

Reiter, W., Chapple, C. C. S. and Somerville, C. R. (1993). Altered growth and cell walls in a of *Arabidopsis* fucose-deficient mutant. *Science*, 261 (5124), pp.1032–1035. [Online]. Available at: doi:10.1126/science.261.5124.1032

Rerkasem, B. et al. (1993). Boron deficiency in soybean [*Glycine max* (L.) Merr] peanut (*Arachis hypogaea* L.) and black gram [*Vigna mungo* (L.) Hepper]: Symptoms in seeds and differences among soybean cultivars in susceptibility to boron deficiency. *Plant and Soil*, 150, pp.289–294. [Online]. Available at: doi:10.1007/BF00013026.

Rerkasem, B. et al. (1997). Relationship of seed boron concentration to germination and growth of soybean (*Glycine max*). *Nutrient Cycling in Agroecosystems, 48*, pp.217–223.

Reuhs, B. L. et al. (2004). L-galactose replaces L-fucose in the pectic polysaccharide rhamnogalacturonan II synthesized by the L-fucose-deficient *mur1 Arabidopsis mutant*. *Planta*, 219 (1), pp.147–157. [Online]. Available at: doi:10.1007/s00425-004-1205-x.

Rhomberg, S. et al. (2006). Reconstitution in vitro of the GDP-fucose biosynthetic pathways of *Caenorhabditis elegans* and *Drosophila melanogaster*. *FEBS Journal*, 273 (10), pp.2244–2256. [Online]. Available at: doi:10.1111/j.1742-4658.2006.05239.x.

Robert, H. S. et al. (2009). BTB and TAZ domain scaffold proteins perform a crucial function in *Arabidopsis* development. *The Plant Journal*, 58 (1), pp.109–121. [Online]. Available at: doi:10.1111/j.1365-313X.2008.03764.x.

Rogers, E. E. and Guerinot, M. Lou. (2002). FRD3, a member of the multidrug and toxin efflux family, controls iron deficiency responses in *Arabidopsis*. *The Plant Cell*, 14, pp.1787–1799. [Online]. Available at: doi:10.1105/tpc.001495.visiae.

Roppolo, D. et al. (2011). A novel protein family mediates Casparian strip formation in the endodermis. *Nature*, 473 (7347), pp.381–384. [Online]. Available at: doi:10.1038/nature10070.

Roppolo, D. et al. (2014). Functional and evolutionary analysis of the CASPARIAN STRIP MEMBRANE DOMAIN PROTEIN family. *Plant Physiology*, 165, pp.1709–1722. [Online]. Available at: doi:10.1104/pp.114.239137.

Routray, P. et al. (2018). Nodulin Intrinsic Protein 7;1 is a tapetal boric acid channel involved in pollen cell wall formation. *Plant Physiology*, 178, pp. 1269-1283. [Online]. Available at: doi:10.1104/pp.18.00604.

Roycewicz, P. S. and Malamy, J. E. (2014). Cell wall properties play an important role in the emergence of lateral root primordia from the parent root. *Journal of Experimental Botany*, 65 (8), pp.2057–2069. [Online]. Available at: doi:10.1093/jxb/eru056.

Rubio, V. et al. (2001). A conserved MYB transcription factor involved in phosphate starvation signaling both in vascular plants and in unicellular algae. *Genes and Development*, 15, pp.2122–2133. [Online]. Available at: doi:10.1101/gad.204401.availability.

Rus, A. et al. (2006). Natural variants of AtHKT1 enhance Na^+ accumulation in two wild populations of *Arabidopsis*. *PLoS Genetics*, 2 (12), pp.1964–1973. [Online]. Available at: doi:10.1371/journal.pgen.0020210.

Ruuhola, T. et al. (2011). Boron nutrition affects the carbon metabolism of silver birch seedlings. *Tree physiology*, 31 (11), pp.1251–1261. [Online]. Available at: doi:10.1093/treephys/tpr109.

Ryden, P. et al. (2003). Tensile properties of *Arabidopsis* cell walls depend on both a xyloglucan cross-linked microfibrillar network and rhamnogalacturonan II-borate complexes. *Plant Physiology*, 132 (2), pp.1033–1040. [Online]. Available at: doi:10.1104/pp.103.021873.

Salt, D. E. (2004). Update on plant ionomics. *Plant Physiology*, 136 (1), pp.2451–2456. [Online]. Available at: doi:10.1104/pp.104.047753.

Salt, D. E., Baxter, I. and Lahner, B. (2008). Ionomics and the study of the [lant ionome. *Annual Review of Plant Biology*, 59, pp.709–733. [Online]. Available at: doi:10.1146/annurev.arplant.59.032607.092942.

Samuel, T., Noble, R. and Parkway, S. N. (2004). Phosphate transport in *Arabidopsis* : Pht1;1 and Pht1;4 play a major role in phosphate acquisition from both low- and high-phosphate environments. *The Plant Jounal*, 39 (1), pp.629–642. [Online]. Available at: doi:10.1111/j.1365-313X.2004.02161.x.

Sanderson, K. R., Sanderson, J. B. and Gupta, U. C. (2002). Boron for brown-heart control on two rutabaga cultivars. *Canadian Journal of Plant Science*, 82 (3), pp.561–565. [Online]. Available at: doi:10.4141/P01-055.

Schachtman, D. P. and Shin, R. (2007). Nutrient sensing and signaling: NPKS. *Annual Review of Plant Biology*, 58 (1), pp.47–69. [Online]. Available at: doi:10.1146/annurev.arplant.58.032806.103750.

Scheller, H. V. and Ulvskov, P. (2010). Hemicelluloses. *Annual Review of Plant Biology*, 61 (1), pp.263–289. [Online]. Available at: doi:10.1146/annurev-arplant-042809-112315.

Schnurbusch, T. et al. (2010). Boron toxicity tolerance in barley through reduced expression of the multifunctional aquaporin HvNIP2;1. *Plant Physiology*, 153 (4), pp.1706–1715. [Online]. Available at: doi:10.1104/pp.110.158832.

Sessions, A. et al. (2002). A high-throughput *Arabidopsis* reverse genetics system. *The Plant Cell*, 14 (1), pp.2985–2994. [Online]. Available at: doi:10.1105/tpc.004630.netics.

Sheldon, C. C. et al. (2000). The molecular basis of vernalization: the central role of FLOWERING LOCUS C (FLC). *Proceedings of the National Academy of Sciences of the United States of America*, 97 (6), pp.3753–3758.

Shelp, B. J. et al. (1995). Boron mobility in plants. *Physiologia Plantarum*, 94, pp.356–361. [Online]. Available at: doi:10.1111/j.1399-3054.1995.tb05323.x.

Shelp, B. J. and Shattuck, V. I. (1987). Boron nutrition and mobility, and its relation to the elemental composition of greenhouse grown root crops I. Rutabaga. *Communications in Soil Science and Plant Analysis*, 18 (2), pp.187–201. [Online]. Available at: doi:10.1080/00103628709367810.

Shen, B., Li, C. and Tarczynski, M. C. (2002). High free-methionine and decreased lignin content result from a mutation in the *Arabidopsis* S-adenosyl-L-methionine synthetase 3 gene. *The Plant Journal*, 29 (3), pp.371–380. [Online]. Available at: doi:10.1046/j.1365-313X.2002.01221.x.

Shi, L. et al. (2012). Identification of quantitative trait loci associated with low boron stress that regulate root and shoot growth in *Brassica napus* seedlings. *Molecular Breeding*, 30 (1), pp.393–406. [Online]. Available at: doi:10.1007/s11032-011-9629-z.

Shorrocks, V. M. (1997). The occurrence and correction of boron deficiency. *Plant and soil*, 193, pp.121–148. [Online]. Available at: doi:10.1023/A:1004216126069.

Singh, S. K. et al. (2018). ACR11 modulates levels of reactive oxygen species and salicylic acid-associated defense response in *Arabidopsis*. *Scientific Reports*, 8 (1), pp.1–9. [Online]. Available at: doi:10.1038/s41598-018-30304-0.

Slatkin, M. (2008). Linkage disequilibrium - understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, 9 (6), pp.477–485. [Online]. Available at: doi:10.1038/nrg2361.

Smith, F. W. (1986). Interpretation of plant analysis: concepts and principles. In *Plant Analysis: An Interpretation Manual (Second Edition)*. Inkata Press. pp. 1-12

Somerville, C. (2006). Cellulose synthesis in higher plants. *Annual Review of Cell and Developmental Biology*, 22, pp.53–78. [Online]. Available at: doi:10.1146/annurev.cellbio.22.022206.160206.

Spencer, C. C. A. et al. (2009). Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. *PLoS Genetics*, 5 (5). e1000477. [Online]. Available at: doi:10.1371/journal.pgen.1000477.

Stangoulis, J. et al. (2010). The mechanism of boron mobility in wheat and canola phloem. *Plant Physiology*, 153 (2), pp.876–881. [Online]. Available at: doi:10.1104/pp.110.155655.

Stangoulis, J. C. R. et al. (2001). The efficiency of boron utilisation in canola. *Australian Journal of Plant Physiology*, 28, pp.1109–1114. [Online]. Available at: doi:http://dx.doi.org/10.1071/PP00164.

Sun, J. et al. (2012). Cloning and characterization of boron transporters in *Brassica napus*. *Molecular Biology Reports*, 39 (2), pp.1963–1973. [Online]. Available at: doi:10.1007/s11033-011-0930-z.

Sutton, T. et al. (2007). Boron-toxicity tolerance in barley arising from efflux transporter amplification. *Science*, 318, pp.1446-1448. [Online]. Available at: doi:10.1126/science.1146853.

Sweeney, A. (2018). Investigating the genetic basis of seed and leaf micronutrient concentrations in *Brassica napus*. PhD Thesis, University of York, York.

Sze, H. et al. (2002). A simple nomenclature for a complex proton pump: VHA genes encode the vacuolar H^+-ATPase. *Trends in Plant Science*, 7 (4), pp.157–161.

Tadege, M. et al. (2001). Control of flowering time by FLC orthologues in *Brassica napus*. *The Plant Journal*, 28 (5), pp.545–553. [Online]. Available at: doi:10.1046/j.1365-313X.2001.01182.x.

Takada, S. et al. (2014). Improved tolerance to boron deficiency by enhanced expression of the boron transporter BOR2. *Journal of Soil Science and Plant Nutrition*, 60 (August), pp.341–348. [Online]. Available at: doi:10.1080/00380768.2014.881705.

Takano, J. et al. (2002). *Arabidopsis* boron transporter for xylem loading. *Nature*. 420, pp. 337-340. [Online]. Available at: doi:10.1038/nature01139.

Takano, J. et al. (2005). Endocytosis and degradation of BOR1, a boron transporter of *Arabidopsis thaliana*, regulated by boron availability. *Proceedings of the National Academy of Sciences of the United States of America*, 102 (34), pp.12276–12281. [Online]. Available at: doi:10.1073/pnas.0502060102.

Takano, J. et al. (2006). The *Arabidopsis* major intrinsic protein NIP5;1 is essential for efficient boron uptake and plant development under boron limitation. *The Plant Cell*. 18, pp. 1498-1509. [Online]. Available at: doi:10.1105/tpc.106.041640.

Tanaka, M. et al. (2008). NIP6;1 Is a boric acid channel for preferential transport of boron to growing shoot tissues in *Arabidopsis*. *The Plant Cell*, 20 (10), pp.2860–2875. [Online]. Available at: doi:10.1105/tpc.108.058628.

Tanaka, M. et al. (2011). Boron-dependent degradation of NIP5;1 mRNA for acclimation to excess boron conditions in *Arabidopsis*. *The Plant Cell*, 23 (9), pp.3547–3559. [Online]. Available at: doi:10.1105/tpc.111.088351.

Tanaka, N. et al. (2014). Characteristics of a root hair-less line of *Arabidopsis thaliana* under physiological stresses. *Journal of Experimental Botany*, 65 (6), pp.1497–1512. [Online]. Available at: doi:10.1093/jxb/eru014.

Tao, P. et al. (2014). Genome-wide identification and characterization of aquaporin genes (AQPs) in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *Molecular Genetics and Genomics*, 289, pp.1131–1145. [Online]. Available at: doi:10.1007/s00438-014-0874-9.

Thomas, C. L. et al. (2016). Root morphology and seed and leaf ionomic traits in a *Brassica napus* L. diversity panel show wide phenotypic variation and are characteristic of crop habit. *BMC Plant Biology*, 16, 214. [Online]. Available at: doi:10.1186/s12870-016-0902-5.

Tian, H. et al. (2010). *Arabidopsis* NPCC6 / NaKR1 Is a phloem mobile metal binding protein necessary for phloem function and root meristem maintenance. *The Plant Cell*, 22, pp.3963–3979. [Online]. Available at: doi:10.1105/tpc.110.080010.

Trick, M. et al. (2009a). A newly-developed community microarray resource for transcriptome profiling in *Brassica* species enables the confirmation of *Brassica* -specific expressed sequences. *BMC Plant Biology*, 10, 50. [Online]. Available at: doi:10.1186/1471-2229-9-50

Trick, M. et al. (2009b). Single nucleotide polymorphism (SNP) discovery in the polyploid *Brassica napus* using Solexa transcriptome sequencing. *Plant Biotechnology Journal*, 7 (4), pp.334–346. [Online]. Available at: doi:10.1111/j.1467-7652.2008.00396.x.

Tyerman, S. D., Niemietz, C. M. and Bramley, H. (2002). Plant aquaporins: multifunctional water and solute channels with expanding roles. *Plant, Cell and Environment*, 25 (2), pp.173–194. [Online]. Available at: doi:10.1046/j.0016-8025.2001.00791.x.

U. N. (1935). Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Journal of Japanese Botony*, 7, pp.389–452.

Ugalde, J. E. et al. (2011). Identification and characterization of the *Chlamydia trachomatis* L2 S-adenosylmethionine transporter. *Microbiology*, 2 (3), pp.5716–5723. [Online]. Available at: doi:10.1128/mBio.00051-11.Editor.

Uluisik, I., Karakaya, H. C. and Koc, A. (2018). The importance of boron in biological systems. *Journal of Trace Elements in Medicine and Biology*, 45, pp.156–162. [Online]. Available at: doi:10.1016/j.jtemb.2017.10.008.

Uzcategui, N. et al. (2008). Alteration in glycerol and metalloid permeability by a single mutation in the extracellular C-loop of *Leishmania major* aquaglyceroporin LmAQP1. *Molecular Microbiology*, 70 (6), pp.1477–1486. [Online]. Available at: doi:10.1016/j.neuroimage.2013.08.045.The.

Velasquez, S. M. et al. (2011). O-Glycosylated Cell Wall Proteins Are Essential in Root Hair Growth. *Science*, 332 (June), pp.1401–1404.

Veneklaas, E. J. et al. (2012). Opportunities for improving phosphorus-use efficiency in crop plants. *New Phytologist*, 195 (2), pp.306–320. [Online]. Available at: doi:10.1111/j.1469-8137.2012.04190.x.

Vu, K. Van et al. (2016). Systematic deletion of the ER lectin chaperone genes reveals their roles in vegetative growth and male gametophyte development in *Arabidopsis*. *The Plant Journal*, 89 (5), pp.972–983. [Online]. Available at: doi:10.1111/tpj.13435.

Wallace, I. S. (2004). Homology modeling of representative subfamilies of *Arabidopsis* major intrinsic proteins. Classification based on the aromatic/arginine selectivity filter. *Plant Physiology*, 135 (2), pp.1059–1068. [Online]. Available at: doi:10.1104/pp.103.033415.

Wang, G. et al. (2006). Involvement of auxin and CKs in boron deficiency induced changes in apical dominance of pea plants (*Pisum sativum* L.). *Plant Physiology*, 163 (6), pp.591–600. [Online]. Available at: doi:10.1016/j.jplph.2005.09.014.

Wang, N. et al. (2014). Genome-wide investigation of genetic changes during modern breeding of *B. napus*. *Theoretical and Applied Genetics*, 127 (8), pp.1817–1829. [Online]. Available at: doi:10.1007/s00122-014-2343-6

Wang, N. et al. (2016). Association mapping of flowering time QTLs and insight into their contributions to rapeseed growth habits. *Frontiers in Plant Science*, 7 (March), 338. [Online]. Available at: doi:10.3389/fpls.2016.00338.

Wang, X. et al. (2017). Breeding histories and selection criteria for oilseed rape in Europe and China identified by genome wide pedigree dissection. *Scientific Reports*, 7, 1917. [Online]. Available at: doi:10.1021/acs.analchem.7b00732.

Wang, Y. et al. (1976). The first study of shrink of cotton leaves and failure of cotton acquire to flower in Shigushan Region. In: *Cotton*. pp.35–36.

Wang, Z. et al. (2011). Proteomics reveals the adaptability mechanism of *Brassica napus* to short-term boron deprivation. *Plant and Soil*, 347 (1), pp.195–210. [Online]. Available at: doi:10.1007/s11104-011-0838-3.

Wang, Z. Y. et al. (1999). Effect of boron and low temperature on membrane integrity of cucumber leaves. *Journal of Plant Nutrition*, 22 (3), pp.543–550. [Online]. Available at: doi:10.1080/01904169909365650.

Warington, K. (1923). The effect of boric acid and borax on the broad bean and certain other plants. *Annals of Botany*, 37 (4), pp.629–672. [Online]. Available at:

doi:http://aob.oxfordjournals.org/content/os-37/4/629.extract.

Watanabe, T. et al. (2016). The ionomic study of vegetable crops. *PLoS ONE*, 11 (8), e0160273. [Online]. Available at: doi:10.1371/journal.pone.0160273.

Watson, C. A. et al. (2012). Using soil and plant properties and farm management practices to improve the micronutrient composition of food and feed. *Journal of Geochemical Exploration*, 121, pp.15–24. [Online]. Available at: doi:10.1016/j.gexplo.2012.06.015.

White, P. J. (2011). Long-distance Transport in the Xylem and Phloem. In *Marschners Mineral Nutrition of Higher Plants* (Third Edition), Elsevier Ltd. pp. 49-70. [Online]. Available at: doi:10.1016/B978-0-12-384905-2.00003-0.

White, P. J. (2013). Improving potassium acquisition and utilisation by crop plants. *Journal of Plant Nutrition and Soil Science*, 176 (3), pp.305–316. [Online]. Available at: doi:10.1002/jpln.201200121.

White, P. J. et al. (2013). Root traits for infertile soils. *Fronteirs in Plant Science*, 4 (June), 193. [Online]. Available at: doi:10.3389/fpls.2013.00193.

White, P. J. and Brown, P. H. (2010). Plant nutrition for sustainable development and global health. *Annals of Botany*, 105, pp. 1073-1080. [Online]. Available at: doi:10.1093/aob/mcq085.

Will, S. et al. (2012). Boron foliar fertilization of soybean and lychee: effects of side of application and formulation adjuvants. *Journal of Plant Nutrition and Soil Science*, 175 (2), pp.180–188. [Online]. Available at: doi:10.1002/jpln.201100107.

Wilson, R. C. and Doudna, J. A. (2013). Molecular mechanisms of RNA interference. *Annual Review of Biophysics*, 42, pp.217–239. [Online]. Available at: doi:10.1146/annurev-biophys-083012-130404.

Won, S.-K. et al. (2009). cis-element- and transcriptome-based screening of root hair-specific genes and their functional characterization in *Arabidopsis*. *Plant Physiology*, 150 (3), pp.1459–1473. [Online]. Available at: doi:10.1104/pp.109.140905.

Wood, I. P. et al. (2017). Carbohydrate microarrays and their use for the identification of molecular markers for plant cell wall composition. *Proceedings of the National Academy of Sciences of the United States of America*, 114 (26), pp.6860-6865. [Online]. Available at: doi:10.1073/pnas.1619033114.

Xiuwen, W. et al. (2018). Leaf structure and chemical compositions are correlated with cotton boron efficiency. *Journal of Plant Nutrition*, 41 (5), pp.552–562. [Online]. Available at: doi:10.1080/01904167.2017.1392571.

Xu, F., Wang, Y. and Meng, J. (2001). Mapping boron efficiency gene (s) in *Brassica napus* using RFLP and AFLP markers. *Plant Breeding*, 120, pp.319–324. [Online]. Available at: doi:10.1046/j.1439-0523.2001.00583.x/full.

Xu, L. et al. (2016). Genome-wide association study reveals the genetic architecture of fl owering time in rapeseed *(Brassica napus* L.). 23 (1), pp.43–52. [Online]. Available at: doi:10.1093/dnares/dsv035.

Yang, H. et al. (2015). High and low affinity urea root uptake: involvement of NIP5;1. *Plant and Cell Physiology*, 56 (8), pp.1588–1597. [Online]. Available at: doi:10.1093/pcp/pcv067.

Yang, M. et al. (2009). Effect of boron on dynamic change of seed yield and quality formation in developing seed of *Brassica napus*. *Journal of Plant Nutrition*, 32 (5), pp.785–797. [Online]. Available at: doi:10.1080/01904160902787883.

Yang, Y. et al. (1993). Responses of rape genotypes to boron application. *Plant and Soil*, 155/156, pp.321–324. [Online]. Available at: doi:10.1007/BF00025047.

Yao, Y. et al. (2014). Root transcriptome analysis on the grape genotypes with contrast translocation pattern of excess manganese from root to shoot. *Plant and Soil*, 387 (1–2), pp.49–67. [Online]. Available at: doi:10.1007/s11104-014-2279-2.

Yuan, D. et al. (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 (August), 1336. [Online]. Available at: doi:10.3389/fpls.2017.01336.

Zeng, C. et al. (2007). Genetic analysis of the physiological responses to low

boron stress in *Arabidopsis thaliana*. *Plant, Cell & Environment*, 31, pp.112–122. [Online]. Available at: doi:10.1111/j.1365-3040.2007.01745.x.

Zhang, D. et al. (2014). A high-density genetic map identifies a novel major QTL for boron efficiency in oilseed rape (*Brassica napus* L.). *PLoS ONE*, 9 (11), e112089. [Online]. Available at: doi:10.1371/journal.pone.0112089.

Zhang, M. and Malhi, S. S. (2010). Perspectives of oilseed rape as a bioenergy crop. *Biofuels*, 1 (4), pp.621–630. [Online]. Available at: doi:10.4155/bfs.10.28.

Zhang, Q. et al. (2017). The boron transporter BnaC4.BOR1;1c is critical for inflorescence development and fertility under boron limitation in *Brassica napus*. *Plant Cell and Environment*, 40 (9), pp.1819–1833. [Online]. Available at: doi:10.1111/pce.12987.

Zhang, Z. et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nature Genetics*, 42 (4), pp.355–360. [Online]. Available at: doi:10.1038/ng.546.

Zhao, C. et al. (2019). Arabinose biosynthesis is critical for salt stress tolerance in *Arabidopsis*. *New Phytologist*, 222 (2). [Online]. Available at: doi:10.1111/nph.15867.

Zhao, H. et al. (2008). Mapping and validation of chromosome regions conferring a new boron-efficient locus in *Brassica napus*. *Molecular Breeding*, 22 (3), pp.495–506. [Online]. Available at: doi:10.1007/s11032-008-9193-3.

Zhao, S. et al. (2016). Phosphorylation of ARF2 relieves its repression of transcription of the K^+ transporter gene HAK5 in response to low potassium stress. *The Plant Cell*, 28 (12), pp.3005–3019. [Online]. Available at: doi:10.1105/tpc.16.00684.

Zhao, Z. et al. (2012). Dissecting quantitative trait loci for boron ffficiency across multiple environments in *Brassica napus*. *PLoS ONE*, 7 (9), e45215. [Online]. Available at: doi:10.1371/journal.pone.0045215.

Zhou, T. et al. (2016). Physiological and transcriptional analyses reveal differential phytohormone responses to boron deficiency in *Brassica napus* genotypes. *Frontiers in Plant Science*, 7, 221. [Online]. Available at: doi:10.3389/fpls.2016.00221.

Zhu, C. et al. (2008). Status and prospects of association mapping in plants. *The Plant Genome*, 1 (1), pp.5–20. [Online]. Available at: doi:10.3835/plantgenome2008.02.0089.