NATURAL GENETIC VARIATION IN ZINC (ZN) ACCUMULATION IN BRASSICACEAE

Seosamh Ó Lochlainn

B.A., Ollscoil na hÉireann, Má Nuad
B.Sc. (Hons.), University of Essex
M.Sc. (Dist.), University of Nottingham

Thesis submitted to the University of Nottingham for the degree of
Doctor of Philosophy

January 2011

ACKNOWLEDGEMENTS

The research presented in this thesis would not have been possible without the help, support, generosity and understanding from many friends and colleagues over the past four years. Firstly, my sincere gratitude goes to the BBSRC for funding my studentship and the University of Nottingham for agreeing to host my PhD. I would like to express my profound gratitude to my supervisors, Drs. John Hammond, Graham King, Philip White, Rupert Fray and Martin Broadley. Graham and Philip, thank you for your valuable and detailed feedback, input and ideas. John and Rupert, thank you for patiently teaching me the finer details of molecular genetics. Martin, I extend my deepest appreciation for welcoming me to your research group and providing almost endless resources to perform such interesting and engaging research. Moreover, thank you for being a constant source of inspiration, energy, advice and enthusiasm. I look forward to remaining in contact over the coming years.

My heartfelt gratitude goes to Helen Bowen, Mark Meacham, Paul Kasprzak, Rory Hayden, William Spracklen, Khalid Alamer and Deval Patel for their friendship, technical advice, guidance, understanding, and for the long hours selflessly spent growing, harvesting and processing in the name of research. I am indebted to Silin Zhong, Zsuzsa Bodi, James Button, Kati Kovács and Paul Anthony for patiently sharing their expertise with me over the years. Many thanks must go to Bob Kear, Helen Smith, Sonoko Mitsui-Angwin, Diane Jones, Jackie Humphreys and Jennifer Dewick in Plant Science, for calmly attending to all those last minute crises.

I am eternally grateful to my wife and love of my life, Anushka. You kept me sane when I was stressed and made me laugh when times were tough. I thank you for all that you have done, and look forward to our life together.

I wish to finally dedicate this thesis to my parents. Without your endless encouragement and support I could not have made it this far. Go raibh míle maith agaibh as bhur dtacaíocht agus bhur gcabhair leanúnach i rith mo shaoil.

Do mo mháthair agus do m'athair…..

PEER REVIEWED PUBLICATIONS RELATED TO THIS THESIS

Papers published

Ó Lochlainn, S., Fray, R. G., Hammond, J. P., King, G. J., White, P. J., Young, S. D. and Broadley, M. R. (2011) Generation of nonvernal-obligate, faster-cycling Noccaea caerulescens lines through fast neutron mutagenesis. New Phytologist **189:** 409-414

Love, C.G., Graham, N.S., **Ó Lochlainn**, S., Bowen, H.C., May, S.T., White, P.J., Broadley, M.R., Hammond, J.P. and King, G.J. (2010) A Brassica exon array for whole-transcript gene expression profiling. PLoS ONE **5:** e12812. doi:10.1371/journal.pone.0012812

Broadley, M.R., **Ó Lochlainn**, S., Hammond, J.P., Bowen, H.C., Cakmak, I., Eker, S., Erdem, H., King, G.J. and White, P.J. (2010) Shoot zinc (Zn) concentration varies widely within Brassica oleracea L. and is affected by soil Zn and phosphorus (P) levels. Journal of Horticultural Science and Biotechnology **85:** 375-380

Broadley, M.R., Hammond, J.P., King, G.J., Bowen, H.C., Hayden, R., Spracklen, W.P., **Ó Lochlainn**, S. and White, P.J. (2009) Biofortifying Brassica with calcium (Ca) and magnesium (Mg). The Proceedings of the International Plant Nutrition Colloquium XVI **1256** (http://repositories.cdlib.org/ipnc/xvi/1256)

Ó Lochlainn, S., Bowen, H., Fray, R., Hammond, J., King, G., Mills, V., White, P. and Broadley, M. (2007) Natural genetic variation in zinc (Zn) accumulation in Brassicaceae species. Comparative Biochemistry and Physiology: A-Molecular and Integrative Physiology, **146:** S252

Papers submitted

Ó Lochlainn, S., Bowen, H. C., Fray, R. G., Hammond, J. P., King, G. J., White, P. J., Graham, N. S. and Broadley, M. R. (2010) Tandem quadruplication of HMA4 in the zinc (Zn) and cadmium (Cd) hyperaccumulator Noccaea caerulescens. Submitted to PLoS ONE.

ABSTRACT

Zinc (Zn) is an essential plant nutrient. Most plant species have a shoot Zn concentration ($[Zn]_{shoot}$) <0.1 mg Zn g^{-1} dry weight (DW), but extensive natural genetic variation occurs. For example, within the Brassicaceae, some Noccaea (Thlaspi) and Arabidopsis species hyperaccumulate $[Zn]_{shoot}$ >10 mg Zn g^{-1} DW. There is compelling evidence that orthologues of the Arabidopsis thaliana P_{IB}-type Heavy-Metal-Associated domain-containing ATPase 4 (AtHMA4), which transport Zn^{2+} and other cations, have a major involvement in the Zn hyperaccumulation trait. The aim of this thesis was to study aspects of genetic variation in the Brassicaceae using a comparative genomic approach, focussing primarily on orthologues of AtHMA4 in Noccaea and Brassica.

The first major objective was to clone the full genomic sequence of NcHMA4. This locus was successfully sequenced in Noccaea caerulescens Saint Laurent Le Minier. First, a new genomic fosmid library was generated comprising 36,864 clones with 40 kb inserts, giving ~5-fold genomic coverage. Through DNA fingerprinting, Genome Sequencer (GS) FLX 454 sequencing and contig assembly, a single region collinear with AtHMA4 flanking genes was identified. Unlike A. thaliana, four novel tandem HMA4 gene repeats with highly conserved coding regions, but substantially divergent promoter regions, were present. Preliminary evidence indicates cis-regulated high expression, supporting previous expression data for N. caerulescens. Notably, this observation is remarkably consistant with recent findings in A. halleri.

In planta analysis of NcHMA4 remains challenging in N. caerulescens due to a vernal obligate lengthy life cycle (7–9 months) and lack of a robust transformation system. To facilitate future analyses, genetically-stable faster cycling M_4 lines were therefore created using fast neutron (FN) mutagenesis. Two non vernal obligate lines have been characterised bearing fruit as soon as 92 days after sowing (DAS) and showing no perturbed $[Zn]_{shoot}$ or obvious pleiotropic effects. Future efforts should focus on their efficient transformation to improve future in planta biological understanding.

In Brassica, data from previously reported glasshouse and field studies on B. oleracea L. $[Zn]_{shoot}$ were further analysed to test for the presence of HMA4 orthologues in QTL regions. However, large QTL and multiple paralogues have hindered progress. A more efficient Targeting Induced Local Lesions In Genomes (TILLing)-based approach has therefore been pursued in B. rapa during the latter stages of this study. Locus specific allelic variants in a candidate metal transporter gene BraA.CAX1.a have been identified and methods for rapid downstream genotyping (High Resolution Melt (HRM)-based efficient SNP detection technology) and characterisation have been developed successfully. These approaches are now underway for BraA.HMA4 and an additional candidate metal transporter BraA.ESB1. Since A. thaliana knock-outs of ESB1, CAX1 and HMA genes have altered nutritional phenotypes, future studies will focus on their characterisation under contrasting mineral environments.

This thesis has pursued a comparative genomics approach. A previously unreported quadruplication and cis-regulation probably contributes to high HMA4 expression in N. caerulescens. Fast cycling Noccaea lines and a robust Brassica genotyping platform were developed. These will become valuable tools for downstream molecular genetic approaches for in planta functional analysis of HMA4 and other transporters to determine their role in regulating mineral accumulation in Brassicaceae. Ultimately, a greater understanding of genetic variation in $[Zn]_{shoot}$ may have downstream application in genetic biofortification or phytoremediation strategies.

TABLE OF CONTENTS

GLOSSARY

~	approximately
>	greater than
<	less than
∞	symbol for infinity
%	percent
±	plus or minus
®	registered trademark
µ	micro; 12th letter of the Greek alphabet
µl	micro litre (0.000001 L)
3'	end of a DNA strand terminating at the hydroxyl group of the third carbon
5'	end of a DNA strand terminating at the phosphate group of the fifth carbon
Antinutrient	a substance that impairs the absorption of an essential element by the gut
A. thaliana 'Colombia'	Arabidopsis thaliana 'Colombia'
A. tumefaciens	Agrobacterium tumefaciens; gram negative soil Rhizobiaceae bacteria within the Alphaproteobacteria class
APM	bacterial growth culture medium
BAP	6-benzyl amino purine; a cytokinin plant growth regulator
Bioavailability	the amount of an element in a food constituent or a meal that can be absorbed and used by a person eating the meal
Biofortification	the process of increasing the bioavailable concentrations of an element in edible portions of crop plants through agronomic intervention or genetic selection
bp	base pair; two nucleotides on opposite complementary DNA or RNA strands
ca.	circa L.; approximately

CAX	cation/H^+ exchanger gene
Cd	cadmium
Co	cobalt
Cu	copper
cv.	cultivar
d.	days
DFS	diversity foundation set
DH	doubled haploid
DMSO	dimethyl sulfoxides
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide-triphosphate; the units from which DNA molecules are constructed, each carrying a single nitrogenous base (adenine, guanine, cytosine or thymine)
DW	dry weight
E. coli	Escherichia coli; gram-negative, facultative anaerobic and non-sporulating Enterobacteriaceae bacteria within the Gammaproteobacteria class
e.g.	exempli gratia L.; for example
EDTA	Ethylene diamine tetraacetic acid ($C_{10}H_{12}N_2Na_4O_8$)
ESB1	enhanced suberin 1 gene
et al.	et alia L.; and others
etc.	et cetera L.; and the rest
ex vitro	from glass (L.) referring to organisms removed from tissue culture and transplanted; generally plants to soil or a potting mixture
Fe	iron
Fig.	figure
Fortification	the addition of an ingredient to food to increase the concentration of a particular element
GUS	β-glucuronidase reporter gene
G x E	gene x environment interaction
h./hrs.	hour/ hours
HCl	Hydrogen chloride (hydrochloric acid)

Head end	5’ of a DNA strand
HMA	heavy metal associated gene
HRM	high resolution melt curve analysis
i.e.	id est L.; that is
IAA	3- Indoleacetic acid; an auxin plant growth regulator
in planta	within plant species (L.)
in vitro	in glass (L.); plants and tissue grown under axenic conditions within a controlled environment i.e. tissue culture
ITS	internal transcribed spacer (measurement of relatedness between plant accessions)
k	linear term in a Michaelis-Menten equation
Kan	Kanamycin
Kbp	kilobase pair
K_m	the affinity constant in a Michaelis-Menten equation
kPa	kilopascals
L.	latin or Linnaean taxonomy
L	litre
La	Lanthanum, atomic number 57
LB	Luria-Bertani medium for bacterial growth culture
Ltd.	limited; a private company limited by shares
M and M	materials and methods
M	molar; unit of concentration, or molarity, of solutions
m	metre
MBD	metal binding domain
mg	milligram
min.	minimum
min/mins	minute/ minutes
ml	millilitre
mM	millimole
Mn	manganese

Mol	mole; an SI unit used for an amount of a particular object based on Avogadro's number, approximately 6.02252 × 1023
MS	Murashige and Skoog medium; an in vitro basal growth medium
MS0	Murashige and Skoog 0; in vitro basal medium containing 8g l^{-1} agar with full strength MS macro and micro salts and vitamins (4.3g l^{-1}), 30g l^{-1} sucrose, but lacking growth regulators, pH 5.8
N	nitrogen
NAA	α-naphthaleneacetic acid; an auxin plant growth regulator
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NASC	National Arabidopsis Stock Centre
Ni	nickel
N. caerulescens	Noccaea caerulescens
NPTII	neomycin phosphotransferase type II gene
O	oxygen
ºC	degrees Celsius
OD	optical density
Pa	Pascals; an SI derived unit of pressure (1Pa≡1kg $m^{-1}s^{-2}$)
PAR	photosynthetically active radiation defined as the incident quantum flux in the waveband 400 to 700 nm
Pb	lead
PCR	polymerase chain reaction
pers. comm.	personal communication
PGR	plant growth regulator
pH	potential of Hydrogen (the logarithm of the reciprocal of hydrogen-ion concentration in gram atoms per litre)
ppm	parts per million

promoter	a substance that stimulates the absorption of an essential element by the gut
QTL	quantitative trait loci
®	registered trademark
RDA	recommended daily allowance
RNAi	ribonucleic acid interference
S	sulphur
SDW	sterile distilled water
s./secs.	second/ seconds
s.l.	sensu lato L.; in the broadest sense
s.s.	sensu stricto L.; in the strictest sense
SNP	single nucleotide polymorphism
ssp.	subspecies
SERCA	a Ca^{2+} ATPase which resides in sarcoplasmic reticulum
SIM	shoot induction medium
SOD	superoxide dismutase
Supplementation	the addition of an element to the diet to make up for an insufficiency
T. arvense 'Wellesbourne'	Thlaspi arvense ecotype Wellesbourne
T. caerulescens 'Ganges'	Thlaspi caerulescens ecotype Ganges
T_1	the first generation of a transformed plant
Tail end	3' of a DNA strand
Taq polymerase	thermostable DNA polymerase from Thermus aquaticus (Taq)
TDZ	Thiadiazuron, N-phenyl, N-1, 2, 3-thiadiazol-5-ylurea; a cytokinin plant growth regulator
™	trademark
U.K.	United Kingdom
U.S.A.	United States of America
v	versus
V_{max}	the maximum velocity of a reaction in a Michaelis-Menten equation e.g. V_{max} = Zn influx when $[Zn]_{ext} = \infty$

v/v	volume to volume ratio of a solution
W.	watts; an SI unit of power, equal to one joule per second
w/v	weight to volume ratio of a solution
wk./wks.	week/ weeks
WT	wild type
Zn	zinc
$[Zn]_{bss}$	mean zinc concentration in the bulk soil solution
$[Zn]_{ext}$	mean zinc availability at the soil root interface
$[Zn]_{leaf}$	mean zinc concentration in the leaf
$[Zn]_{seed}$	mean zinc concentration in the seed
$[Zn]_{shoot}$	mean zinc concentration in the shoot
$[Zn]_{soil}$	mean zinc concentration in the soil
α	alpha; first letter of the Greek alphabet
β	beta; second letter of the Greek alphabet

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

1.1. ZINC AND BIOFORTIFICATION.

Zinc (Zn) is one of the most important essential trace elements for life on this planet (Aras et al., 2001). Although discovered in pure metallic form in 1746 by Andreas Marggraf (Weeks and Leicester, 1968), Zn ore mining and smelting has existed since 1300 BC, with written references to its medicinal properties since 800 BC (Craddock, 1987; Barceloux, 1999; Biswas, 2006). Today Zn has become an intricate and indispensable element in modern electrical, automotive, hardware, cosmetic, plastic and pharmaceutical industries worldwide (Biswas, 2006; LANL, 2007).

More importantly however, is its contribution to biological evolution, from its fundamental role in the creation of the initial primitive nucleic acids to its crucial activity in cellular metabolism for all living organisms (Pfeiffer and Braverman, 1982; Fischer and Davie, 1998; Nielsen, 2000; Dardenne, 2002; Broadley et al., 2007).

In humans, zinc deficiency exists in 33% (2.2 billion) of the world's population, and is one of the ten leading factors contributing to major global diseases including malaria (16%), diarrhoea (10%) or in severe cases, death (1.4% or 0.8 million) (WHO, 2002; Shrimpton et al., 2005; Palmgren et al., 2008). Several curative (i.e. Zn supplementation and Zn food fortification) as well as preventative (i.e. food diversification) measures have been proposed, however their success is limited by economic, practical or sustainability issues (Welch and Graham, 2004; Shrimpton et al., 2005; White and Broadley, 2005, 2009).

Biofortification of staple crops however, has received more favourable attention following the success of transgenic approaches such as, β-carotene enriched 'golden rice' (Ye et al., 2000) and ferritin-Fe enriched rice grain (Goto et al., 1999).

In the case of zinc biofortification, Brassica spp., have been proposed as model crops for several reasons. First, B. oleracea is the most widely directly consumed

leaf in the world, and B. rapa based varietas represent over 12% of global edible vegetable oil production (FAO, 2005; Choi et al, 2007; Broadley et al., 2008). Second, Brassica are the nearest crop relatives to the fully annotated model plant species Arabidopsis thaliana L. as well as Zn hyperaccumulators, predominately evolving at the base of the Noccaea/Raparia clades (van de Mortel et al., 2006) and, with well characterised genomic relationships, comparative genetic studies can be performed relatively simply (Sebastian et al., 2000; Parkin et al., 2005; Kim et al., 2006). Third, the B. rapa genome is currently being fully sequenced with over 130,000 Bacterial Artificial Chromosomes (BAC) clone end-sequences publically available (http://www.brassica.info/resource/sequencing.php). Fourth, a diversity foundation set (DFS) has been established from over 6000 ex situ publicly available accessions Broadley et al., 2008). Fifth, doubled-haploid (DH) populations and substitution accessions with associated linkage maps of segregating co-dominant markers are publicly available (Sebastian et al., 2000). Sixth, unlike the majority of staple crops, B. oleracea contains an abundance of Zn promoters and few Zn anti-nutrients (substances which promote or inhibit Zn metabolism and absorption respectively), thus increasing the potential Zn nutrient bio-availability i.e. the extent to which humans metabolise this Zn (White and Broadley, 2009). Finally, through its enormous natural genome variation, comparative studies have shown that B. oleracea crop genotypes can contain more available Zn in their edible fractions than almost all major global food crops (Welch and Graham, 2004; White and Broadley, 2005, 2009).

1.2. ZINC: PHYSICAL AND CHEMICAL PROPERTIES.

Globally zinc (Zn) is the 23rd most abundant element and 4th most commonly exploited metal with 21 known isotopes, five of which are stable and naturally occurring; ^{64}Zn (48.63%), ^{66}Zn (27.90%), ^{67}Zn (4.90%), ^{68}Zn (18.75%) and ^{70}Zn (0.62%) with ^{65}Zn ($t_{1/2}$=244.26 d) being the most commonly used and longest lived Zn tracer in plants (Zhang, 1996; Broadley et al., 2007; LANL, 2007). The lustrous blue-white transition metal (atomic number 30, melting point 419.58°C, boiling point of 907°C), exists in the +2 oxidation state and, unlike Fe^{2+} and Cu^{2+}, remains redox stable under physiological conditions due to a completed outer shell of electrons (Barak and Helmke, 1993; Auld, 2001; Broadley et al., 2007; LANL, 2007). Furthermore Zn^{2+} has pronounced Lewis acid

characteristics forming strong covalent bonds with S, N and O donors due to its small radius to charge ratio (0.83 Å, coordination number, CN=6). This electron configuration of aqueous Zn^{2+} complexes favours octahedral coordination geometries (CN=6), although CN=4 and CN=5 geometries also occur (Barak and Helmke, 1993; Zhang, 1996; Broadley et al., 2007). Zinc forms numerous soluble salts, including halides, sulphates, nitrates, formates, acetates, thiocyanates, perchlorates, fluosilicates, cyanides, alkali metal zincates and Zn-ammonia salts; sparingly soluble compounds, including Zn-ammonium phosphate, Zn hydroxide and Zn carbonate; and a range of soluble and insoluble organic complexes (Lindsay, 1979; Barak and Helmke, 1993; Zhang, 1996).

1.3. ZINC: BIOCHEMICAL PROPERTIES.

Zinc is the only metal represented in all six enzyme classes (oxidoreductases, transferases, hydrolases, lyases, isomerases and ligases) and following Fe, it is typically the second most abundant transition metal in organisms, with its three principally recognised Zn^{2+} ligand binding sites being structural, catalytic and co-catalytic (Auld, 2001; Maret, 2005; Broadley et al., 2007). In structural Zn sites, it controls protein folding free from bound water molecules. Zn^{2+} is complexed with water and either S, N or O donors to directly affect the catalytic function of the enzyme in catalytic sites. In co-catalytic sites two or three Zn^{2+} are closely linked by amino acid residues (chiefly aspartic acid [Asp] or glutamic acid [Glu] but also His and potentially a water molecule but not Cys) to be employed in catalytic, regulatory and structural functions. Ligands from the surface of two protein molecules can also bind to a single Zn atom to create a fourth class of catalytic or structural sites. Many proteins, membranes and DNA/RNA molecules can also contain Zn-binding sites including zinc finger domain containing proteins, which can directly regulate transcription through DNA/RNA binding, and also through site-specific modifications, regulation of chromatin structure, RNA metabolism and protein–protein interactions (Klug, 1999; Englbrecht et al., 2004; Broadley et al., 2007).

1.4. ZINC ASSOCIATED PROTEINS.

Despite the minimal Zn cellular content of the prokaryote Escherichia coli being ca. 200000 atoms, it is thought that cytosolic free Zn^{2+} pools do not endure, since femtomolar ($1*10^{-15}$M) cytosolic concentrations, which are 106 fold lower than one Zn^{2+} ion per cell, induce Zn influx and efflux protein activity (Outten and O'Halloran, 2001). Cellular Zn^{2+} probably experience cytosolic exclusion through direct protein transfer, 10% of which are considered tightly bound to six proteins, specifically an RNA polymerase with two Zn atoms bound per copy of 5000 copies per cell and five tRNA synthetases with one Zn atom bound per copy of 2000-3000 copies per cell (Outten and O'Halloran, 2001; Broadley et al., 2007). Indeed cytosolic Zn-binding overcapacity probably exists in E. coli, with >30 high affinity Zn binding proteins of unknown copy number and a significant amount of low Zn affinity proteins, amino acids and nucleotides (Outten and O'Halloran, 2001; Katayama et al., 2002).

In Noccaea caerulescens, high Zn-status leaf epidermal cell vacuoles, cell walls and cytoplasm (excluding the vacuole) have been shown to contain, 74 305, 11,577 and 3,205 µg Zn g^{-1}DW (dry weight) respectively (Fig. 1.1.), while those of lower Zn-status leaf mesophyll cells contain, 327, 9,353 and ≤262 µg Zn g^{-1}DW respectively; and root cells contain, ≤262, 589 and ≤262 µg Zn g^{-1}DW respectively (Frey et al., 2000). Leaf epidermal cell vacuoles, cell walls and cytoplasm can contain 9.6×10^{10}, 2.7×10^{9} and 1.8×10^{9} Zn atoms respectively while those of leaf mesophyll cells contain 4.2×10^{8}, 2.2×10^{9} and $\leq1.5\times10^{8}$, and root cortical cells contain $\leq3.4\times10^{8}$, 1.4×10^{8} and $\leq1.5\times10^{8}$ Zn atoms respectively (Broadley et al., 2007). When compared with E. coli it is probable that cytosolic Zn^{2+} concentrations are maintained particularly low through high affinity Zn-binding and compartmentalisation into cytoplasmic organelles to facilitate metalloregulatory and other signalling proteins.

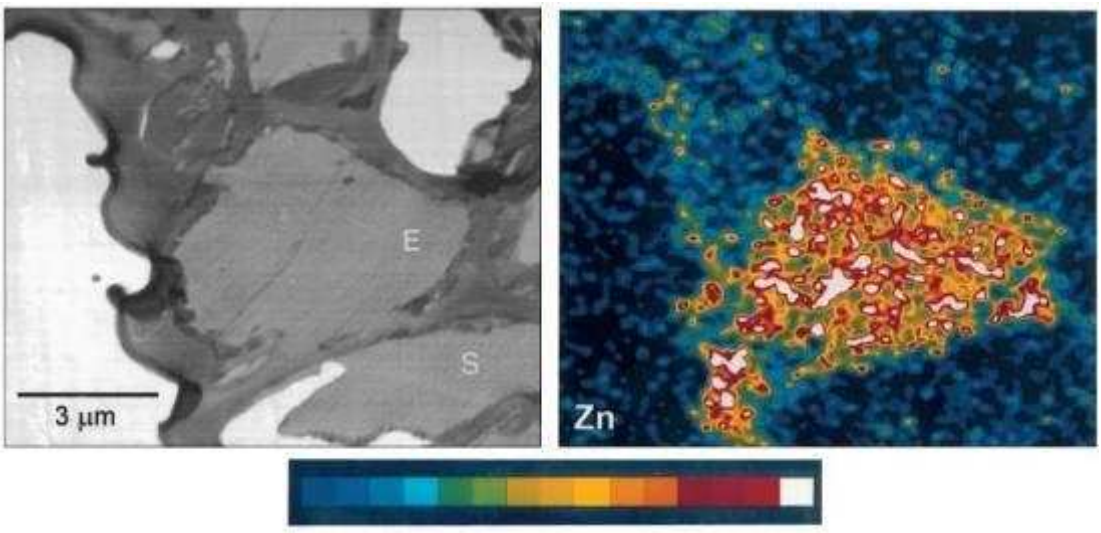


Figure 1.1. Elemental map of a high-pressure-frozen and freeze-dried cryosection of a leaf of the zinc hyperaccumulator Thlaspi caerulescens derived from a metal contaminated soil (Giornico). The STEM micrograph (bright/dark field image) and the corresponding element distribution map for Zn. The colour scale indicates X-ray intensity from low (black, blue) to high (red, white). The following cells were identified: epidermal cell (E), subsidiary cells (S). Zn concentrations (mmol kg^{-1} dry wt [DW]) in epidermal cell wall = 177±36; in vacuole = 1136±130; in organelle = 49±11. Bar = 3 µm.
Source: (Frey et al., 2000).

In Arabidopsis thaliana, Broadley et al. (2007) catalogued proteins implicated in Zn homeostasis, performed an in silico search for proteins with observed and predicted Zn binding domains, and performed an annotation search using the words 'zinc' or 'Zn' to identify 2367 proteins in 181 gene families as predicted Zn related while approximately 120 putative Zn-binding protein domains were identified with 2042 proteins as containing one or more of these domains corresponding to 1096 genes. The largest group of Zn-binding proteins are Zn finger domains assigned to transcription regulator activity (GO:30528) and binding (GO:5488) functional subcategories. The catalytic activity (GO:3824) subcategory comprises proteins with hydrolase activity (GO:16787, e.g. P_{1B}-ATPases) and transferase activity (GO:16740, e.g. mitogen-activated proteinkinases (MAPKs)). The transporter activity (GO:5215) subcategory includes ABC transporters, P_{IB}-ATPases, various divalent cation transporters (for example, the cation diffusion facilitator family (CDFs)), Zn-Fe permeases (ZIPs) and nonspecificcation channels (Broadley et al., 2007).

1.5. ZINC IN SOILS.

1.5.1. Zinc inputs to soils.

Soil zinc content originates primarily from chemical and physical erosion of parent rocks with the lithosphere usually containing 70–80 µg Zn g^{-1}, and sedimentary rocks containing 10–120 µg Zn g^{-1} (Friedland, 1990; Barak and Helmke, 1993; Alloway, 1995; Broadley et al., 2007). Mineral and organic soils characteristically contain 50 and 66 µg total Zn g^{-1} soil ($[Zn]_{soil}$) respectively. However, 'calamine' soils, such as in Plombières in Belgium, can contain Zn sulphide (sphalerite, wurtzite), sulphate (zincosite, goslarite), oxide (zincinte, franklinite, gahnite), carbonate (smithsonite), phosphate (hopeite) or silicate (hemimorphite, willemite) and may exceed 100,000 µg total Zn g^{-1} soil (Alloway, 1995; Barber, 1995; Barceloux, 1999; Cappuyns et al., 2006).

Soil zinc content is also influenced by natural secondary inputs including atmospheric and biotic processes (Broadley et al., 2007). Unnatural secondary inputs, which are 20 fold higher than natural inputs, originate principally from human activities i.e. Zn mining, smelting, fossil fuel combustion, sewage and fertilizer and agrochemical applications which by the early 1990s resulted in mean annual Zn emissions of 2.7 Mt Zn yr^{-1} (Chaney, 1983; Alloway, 1995; Nriagu, 1996).

1.5.2. Zinc behaviour in soils.

Soil Zn occurs in three principal fractions: (i) water-soluble Zn (Zn^{2+} and soluble organic fractions) (ii) adsorbed and exchangeable Zn in the colloidal fraction (associated with clay particles, humic compounds and Al and Fe hydroxides); and (iii) insoluble Zn complexes and minerals (Lindsay, 1979; Broadley et al., 2007). The distribution of Zn between these fractions is chiefly determined by soil-specific precipitation, complexation and adsorption reactions (Broadley et al., 2007). Other factors which affect Zn distribution and phyto-availability in the soil include soil pH (Zn is more readily adsorbed on cation exchange sites at higher pH and adsorbed Zn is more readily displaced by $CaCl_2$ at lower pH), soil moisture, mineral and clay types and contents, diffusion and mass flow rates, weathering rates, soil organic matter, soil biota, plant uptake, available phosphate content, soil temperatures and nitrogen content (Fig. 1.2) (Alloway, 2004).

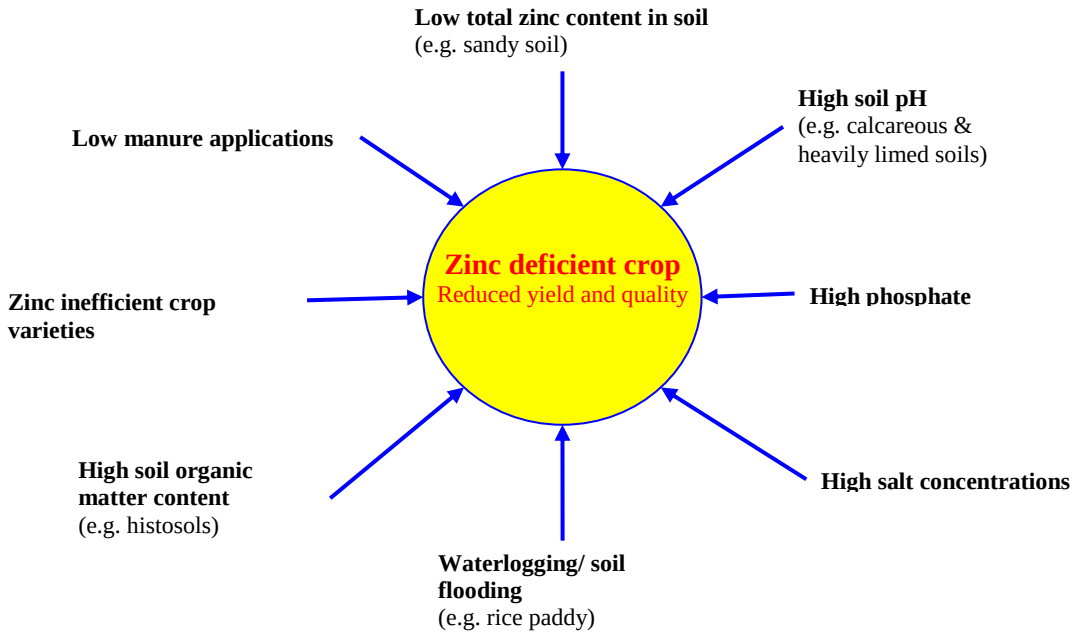


Figure 1.2. Schematic diagram of the causes of Zn deficiency in crops.
Source: adapted from Alloway (2004).

Approximately 90% of soil Zn is insoluble and unavailable for plant uptake, 0.1 to 2 µg Zn g^{-1} exists as exchangeable Zn, and concentrations between 4 * 10^{-10} and 4 * 10^{-6} M of water-soluble Zn exist in the bulk soil solution for both contaminated and uncontaminated soils (Barber, 1995; Knight et al., 1997; Broadley et al., 2007). Although a growth limiting factor at low 10^{-11} to 10^{-9} M in calcareous soils, Zn^{2+} is predominately phytoavailable and characteristically comprises up to 50% of the soluble Zn fraction (Hacisalihoglu and Kochian, 2003; White and Broadley, 2009).

Several modelling approaches to establish soil and plant effects on Zn dynamics have been developed which, adequately determine Zn relationships from soil to root ($[Zn]_{ext}$), despite complexities such as the development of Zn depletion regions around high Zn-uptake roots (Barber and Claassen, 1977; Bar-Yosef et al., 1980; Barber, 1995; Whiting et al., 2003; Qian et al., 2005; Lehto et al., 2006).

The principal regulators of $[Zn]_{ext}$ for plants of similar proportions and with similar root Zn and water-absorbing power are $[Zn]_{bss}$ (concentration of Zn in the bulk soil solution) and b (the Zn-buffering power of the soil), where at high b (>200), $[Zn]_{ext}$ approximates, $[Zn]_{bss}$, i.e., plant-available Zn is determined by the competence of the soil to replenish the soluble Zn fraction as it is removed by the

plant (Whiting et al., 2003; Broadley et al., 2007). Conversely, if b ≤200 or especially <50, $[Zn]_{ext}$ is <<$[Zn]_{bss}$ (Broadley et al., 2007). Despite encompassing some shortcomings, including several necessary but uncorroborated assumptions, modelling $[Zn]_{ext}$ helps to understand plant physiological understanding, optimise crop Zn nutritional fertilisation, and improving risk assessments for metal-contaminated environments (Broadley et al., 2007).

1.5.3. Zinc deficient soils.

Zinc is the most frequent micronutrient deficiency in crops especially in soils of high pH (calcareous), high organic matter, low water, low $[Zn]_{bss,}$ sandy, saline and sodic soils, vertisols and gleysols (Section 1.5.2) (Graham et al., 1992; White and Zasoski, 1999; Cakmak, 2002, 2004; Alloway, 2004). Marschner (1995) demonstrated that a typical leaf Zn concentration ($[Zn]_{leaf}$) of approximately 15 – 20 mg Zn kg^{-1} DW was required for healthy crops. Zn deficiencies in crops are found on many soil types in different bio - climatic zones of the world, (Fig 1.3 and 1.4) (Alloway, 2004).

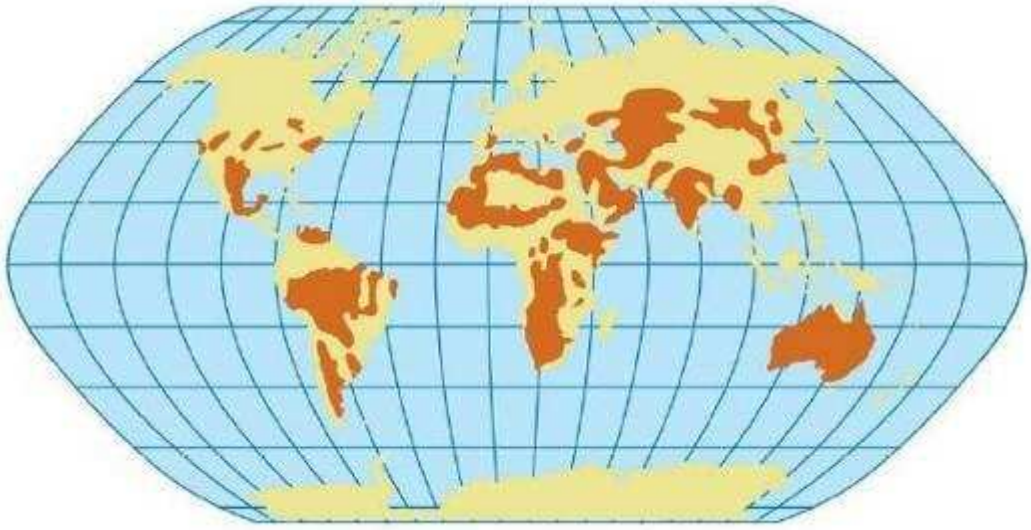


Figure 1.3. Combined global distribution of the main types of soils associated with Zn deficiency. Note, not all the areas of soil shown on this map have conditions suitable for crop production (e.g. desert areas).
Source: (Alloway, 2004).

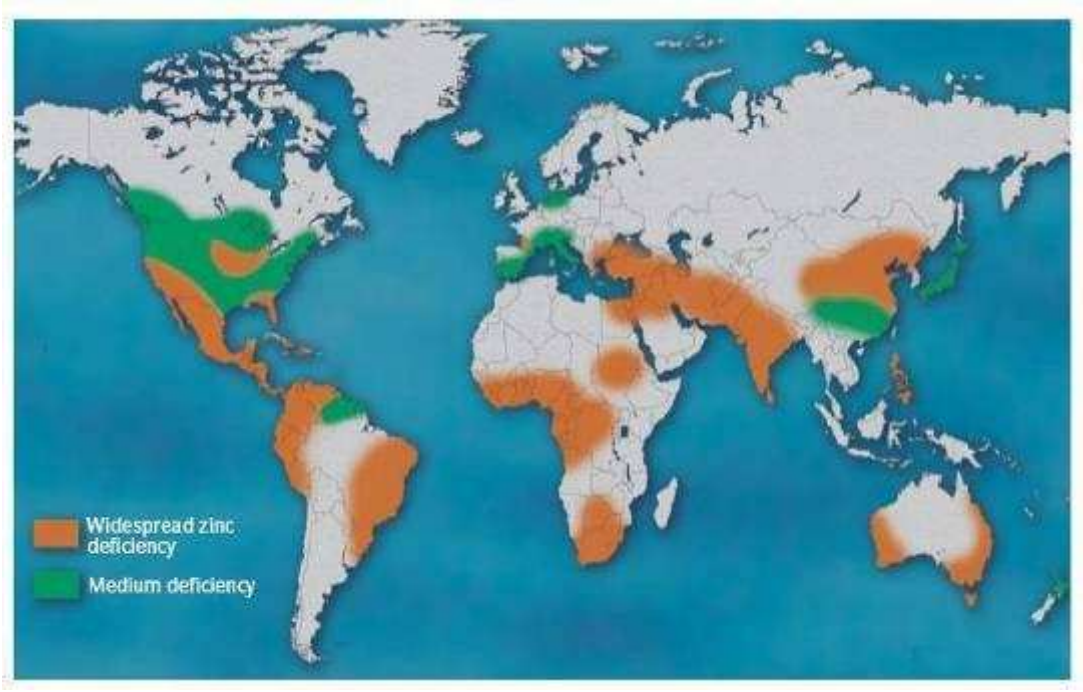


Figure 1.4. Major areas of reported Zn deficiency in world crops.
Source: (Alloway, 2004).

Agronomic and genetic improvement have been undertaken to counteract Zn deficient soil effects, accounting for 50% of the total cultivated soil in Turkey and India, a third of cultivated soil in China, and most soils in Western Australia (Broadley et al., 2007). Zn enriched fertiliser applications successfully resolved little leaf in peach orchards, mottle leaf in citrus, rosetting in pecan and stunting in pineapple (Hoagland, 1948) and more recently a 600% increase in wheat grain yields and a corresponding US$100 million increase in annual economic return in Turkey (Cakmak, 2004). More recently however, interest has been focussed on exploring genotypic differences in Zn uptake and tissue-use-efficiency within crop species as a more sustainable and environmentally safe approach to addressing zinc crop deficiencies (Rengel, 2001; Cakmak, 2002; Hacisalihoglu and Kochian, 2003; Alloway, 2004)

1.5.4. Significance of zinc in plants.

In 1869, Raulin first demonstrated that common bread mould (Aspegillus niger) was found to be unable to grow in the absence of Zn (Alloway, 2004) with similar findings in zinc deficient maize (Mazé, 1915) and barley and dwarf sunflower (Sommer and Lipman, 1926). Zn deficiency in tomato (poor stem

elongation and diminished protein and starch synthesis) were corrected through exogenous Zn applications (Skoog, 1940; Hoagland, 1948; Scaife and Turner, 1983; Marschner, 1995; Sharma, 2006). Zn deficiency phenotypes prevalent in crops include root apex necrosis (‘dieback’), defoliation, small and misshapen fruit or seeds, shoot die-back, interveinal chlorosis (‘mottle leaf’), the formation of bronze hues (‘bronzing’) in addition to auxin deficiency-like responses including internode shortening (‘rosetting’), epinasty, inward curling of leaf lamina (‘goblet’leaves) and leaf size reductions (‘little leaf ’) (Broadley et al., 2007; Alloway, 2004).

1.5.5. Physiological functions of zinc in plants.

1.5.5.1. Carbohydrate metabolism.

Zn deficiency significantly affects carbohydrate metabolism in plants, specifically, sugar transformation, with up to a 70% reduction in photosynthesis (Alloway, 2004). This is assumed to be a result of reduced carbonic anhydrase activity affecting the carbon dioxide assimilation pathway in C_4 crops (Marschner, 1995) or an increase in photosynthetic limiting HCO_3 levels (Brown et al., 1993). Zn is also a major constituent of ribulose 1,5-biphosphate carboxylase (RuBPC) and its reduction would constrain CO_2 fixation in photosynthesis (Brown et al., 1993).

Sucrose and starch formation are also reduced in some zinc deficient plants due primarily to a reduction in sucrose and starch synthetase activities. Conversely, zinc deficient cabbage leaves were shown to increase their concentrations of carbohydrates while those of bean roots showed a decrease, implying interference with the translocation of sucrose from leaves to roots which may be attributed to the function of zinc biomembrane integrity (Brown et al., 1993; Alloway, 2004)

1.5.5.2. Membrane integrity.

Zn performs a critical function in the structure and regulation of biomembranes, including ion-transport, with a greater ^{32}P isotope labelled root exudate leakage demonstrated in Zn-deficient compared to Zn repleat wheat (Welch et al., 1982). Zn may interact with phospholipids and sulphydryl groups of membrane proteins and thus limit photooxidation by controlling the generation and detoxification of

free oxygen radicals (O_2) (Brown et al., 1993). In Zn-deficient root cells of cotton, bean and tomato, superoxide dismutase (SOD) activity is reduced and NADH-dependent free oxygen radicle production is increased (Cakmak and Marschner, 1987, 1988). A subsequent study illustrated that Zn deficient plants leached more phosphorus than those supplied with Zn, as a result of membrane damage (Marschner et al., 1987; Graham et al., 1992). Chandler et al. (1931) reported that the bark of zinc deficient fruit trees were 90% deficient in K while (Cakmak and Marschner, 1987, 1988) reported that Zn deficient cotton, wheat and tomato plant roots had among others, increased K leakage, a major constituent of cell sap, and indicated significant membrane impairment.

1.5.5.3. Protein metabolism.

Commonly, Zn levels in plants are positively correlated to protein levels since Zn is a prerequisite for RNA polymerase enzyme activity, is required to prevent RNA degradation by ribonuclease, and regulates production and formation of ribosomes (Brown et al., 1993). Zn deficiency therefore, initially affects meristematic tissue, even before ribonuclease activity increases as high Zn concentrations are required by this tissue for active cell division and nucleic acid and protein synthesis (Brown et al., 1993). The stability and function of genetic material therefore, is the primary effect of Zn on protein metabolism (Brown et al., 1993).

1.5.5.4. Auxin Metabolism.

Production of the plant growth regulator (PGR) auxin (indole acetic acid, IAA) is positively correlated to Zn content in plants (Brown et al., 1993). ‘Little leaf’ is a common phenotype of Zn deficient plants resulting from inhibited IAA synthesis or increased degradation of IAA. It is thought that Zn is required for Tryptophan synthesis, a precursor for IAA biosynthesis, since rice grains from plants fertilised with Zn contained increased levels of the amino acid (Brown et al., 1993).

1.5.5.5. Reproduction.

Zn is one of the most essential regulators of flowering and seed production in plants. Zn applications on subterranean clover resulted in more significant

increases in the number of inflorescences and seed yield than on dry matter production or seed size (Alloway, 2004). It was proposed by Brown et al. (1993) that Zn affected seed yield by initially restraining abscisic acid synthesis, thus preventing premature leaf and flower bud loss. Evidence from Zn deficient wheat, producing small anthers and aberrant pollen grains, suggested that Zn regulated anther and pollen grain development and physiology.

1.6. ZINC AND ITS NUTRITIONAL IMPORTANCE FOR HUMANS.

Zinc plays a central role in maintaining the human immune system (Shankar and Prasad, 1998). Prasad (1991) first reported cases of delayed maturity among Middle Eastern adolescents and found that zinc supplementation improved their height, weight, bone development and sexual maturation significantly. Global zinc supplementation has demonstrated increased growth among malnourished children and reductions in many childhood infections, such as diarrhoea and pneumonia (Prasad, 1991). Deficiencies in Zn have been associated with some forms of 'head and neck' cancers while modest deficiencies significantly reduced human defence against viral diseases such as the common cold (Shankar and Prasad, 1998; Black et al., 2008; Dijkhuizen et al., 2008). Zn also plays a vital role in human cellular metabolism, with over 100 different enzymes depending on zinc for their ability to catalyse vital chemical reactions. It plays an important structural role with copper (Cu) forming copper-zinc superoxide dismutase (CuZnSOD) which protects cells from free radical damage. Zinc finger proteins regulate gene expression by acting as transcription factors while also having a role in cell signaling, influencing hormone release and nerve impulse transmission (Gizard et al., 2001; Mason, 2006). Organisations such as the Harvest Plus International Food Policy Research Institute have acknowledged the role Zn plays in human health, and thus have concentrated their resources to adopt techniques to alleviate world micro-nutrient deficiencies with a special emphasis on Fe and Zn (Nestel et al., 2006; Stein et al., 2007; Broadley et al., 2009)

1.6.1. Methods to increase Zn in human diets.

Many world regions are plagued with Zn-deficient or phyto-unavailable Zn soils, resulting in inevitable low Zn-density crops (Welch and Graham, 2002; Alloway,

2004; White and Broadley, 2005, 2009), while some produce low Zn accumulating staple crops resulting in 10% of disease and high mortality in developing countries (WHO, 2002; Shrimpton et al., 2005). Currently Zn daily intake is inadequate for a third of the world's population (2.2 billion) with populations in South East Asia and sub-Saharan Africa at the greatest risk of zinc deficiency, affecting 40% of preschool children (Fig. 1.5) (Hotz and Brown, 2004). Complementary food mixes which are fed to infants during weaning are typically Zn deficient (Dewey and Brown, 2003). The long-term sustainable approach is to introduce diet diversification, however this is currently economically unsustainable (Shrimpton et al., 2005; White and Broadley, 2005, 2009). Thus, Zn interventions such as Zn supplementation, food Zn-fortification and Zn-biofortification have been proposed to help reduce child deaths globally by 63% (Jones et al., 2003).

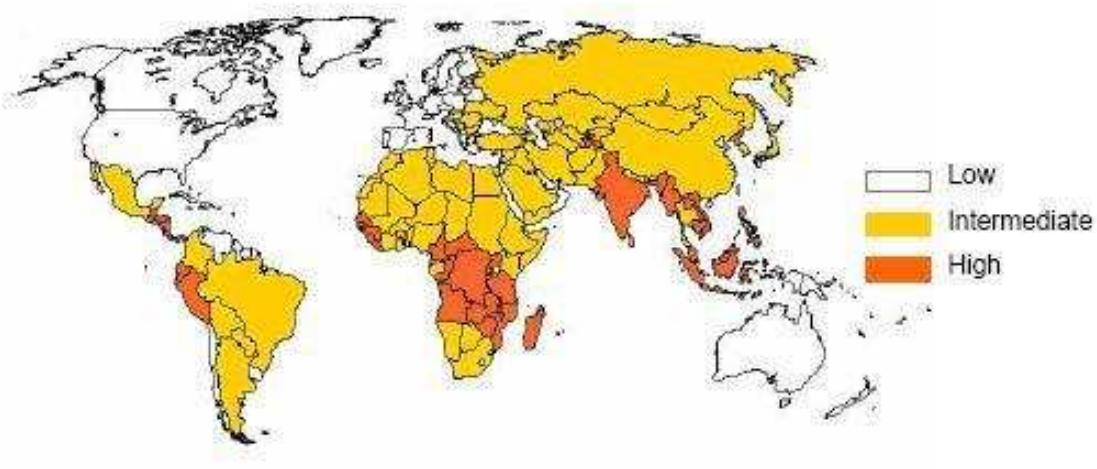


Figure 1.5 Countries with zinc deficiency in the population based on combined information regarding the percent of individuals at risk of inadequate zinc intake and the prevalence of childhood growth-stunting.
Source: (IZiNCG, 2006).

1.6.2. Food fortification.

Success in food fortification campaigns, such as Fe fortified wheat flour in Latin America, has not yet been substantiated (Shrimpton et al., 2005). Research relating to this intervention strategy is in its infancy (Gibson and Ferguson, 1998; Salgueiro et al., 2002), and for Zn, potential problems already surfacing are outlined below (Martinez-Navarrete et al., 2002; White and Broadley, 2005).

First, a bioavailable Zn compound must be selected that does not negatively affect the condition of the food.

Second, staples including cereals, corn meal, corn grits, pasta and breakfast cereals are the favoured food vehicles for such a strategy however, these products contain high levels of phytic acid (Zn absorption inhibitor) and are extremely sensitive to fat oxidation (Martinez-Navarrete et al., 2002). Similarly milk products contain Ca, a known Zn inhibitor (Sandström, 2001). Finally the costs of production could prove prohibitive for many developing countries or its produce too expensive for the poor, mineral deficient population (White and Broadley, 2005, 2009).

1.6.3. Zinc supplementation.

Zinc supplementation has been shown to have many significant health benefits to recipients globally. In Brazil, Zn supplementation significantly reduced diarrhoeal diseases and pneumonia in infants and preschool children respectively, while correcting prepubertal stunting and child birth weights when administered to pregnant mothers (Shrimpton et al., 2005). However, some major concerns regarding this technique centre on how to administer the supplementation, since the most practical i.e. oral rehydration appears least effective as a treatment for diarrhoea. Many countries promote using less effective 'home-made' fluids while the cost of producing and distributing these supplements is dependent on political will, and therefore relatively unstable (White and Broadley, 2009). Zn deficiencies commonly exist alongside Fe deficiencies, which are further increased upon Zn supplementation. Multi-micronutrient supplementation would thus be required, however knowledge surrounding mineral interactions, individual absorption potentials remain insufficient (Shrimpton et al., 2005). In planta micronutrients however, appear to utilise specific absorption mechanisms which protect against any damaging micronutrient interactions (Sandström, 2001), implying that high nutrient containing plant material could replete Zn levels sufficiently without affecting the status of other nutrients in the human diet.

1.6.4. Zinc fertilisation of crops.

Applications of high Zn fertilisers to crops (Hacisalihoglu and Kochian, 2003) was investigated as a biofortification technique using direct foliar applications (Frossard et al., 2000) and soil applications (Cakmak, 2002) and was highly successful in terms of crop yield, vigour and leaf-Zn content, and therefore suited to most Brassica oleracea crop production systems. However, Zn in fruit, seed or grain often increased insignificantly, thus limiting its use in staple cereal or legume crops (White and Broadley, 2005, 2009; Palmgren et al., 2008). Additionally economic factors, including chemical and application frequency costs have proven to be unfeasible for developing world farmers (White and Broadley, 2005). Environmentally, this strategy can be very harmful leading to an eventual accumulation of chemicals in the soil and inhibition of future crop yields. Often the levels of Zn are frustratingly sufficient to support Zn-dense crops, but their phyto-availability is minimal (White and Broadley, 2005). For these reasons, developing nations require crops which are Zn efficient.

1.6.5. Zinc biofortification.

Biofortification using Zn efficient crops is one avenue which can be explored to optimise Zn intake in diet. White and Broadley (2005) defined biofortification as the process of increasing "the bioavailable concentrations of an element in edible portions of crop plants through agronomic intervention or genetic selection". Increasingly investigations are advocating 'genetic selection' since it is more practical and more successful than previously promoted solutions (Welch and Graham, 2004; Palmgren et al., 2008; Barabasz et al., 2010), and has demonstrated the potential to deliver desired results i.e. β-carotene enriched 'golden rice' (Ye et al., 2000) and ferritin-Fe enriched rice grain (Goto et al., 1999). It relies on either natural genetic variation within the species and subsequent selection and plant breeding or genetic modification based approaches to introduce a desired trait into the crop. Bouis et al. (2003) reported on its potential to become sustainable, affordable and access remote rural populations. Staple foods, which are consumed in large quantities by poorer communities, can be targeted for essential mineral increases (White and Broadley, 2009). The economic benefits of this system are immense since the only expense will be the establishment of mineral-dense lines, which thereafter

will cost little to introduce into current breeding programs independent of government or international funding (Welch and Graham, 2002; Bouis et al., 2003; Timmer, 2003; Stein et al., 2007). Zinc biofortification of rice and wheat in India could reduce disability-adjusted life years (DALYs) by 20–51% and save 0.6–1.4 million DALYs each year, costing just $US 0.73–7.31 per DALY. This is considerably lower than that of most other micronutrient interventions (Stein et al., 2007). Finally this system also has agronomic benefits for the producer since studies show that the seed of mineral-dense crops produce more vigorous seedlings, have greater water use efficiency and improved grain yield on infertile soils than unimproved crops (Rengel and Graham, 1995; Yilmaz et al., 1998; Khan et al., 2003; Palmgren et al., 2008).

1.6.6. Improving crops through Zn biofortification.

It has been accepted therefore that there exists two main approaches to control global malnutrition. The first and most sustainable is dietary diversification, however in the current world economic climate this does not appear universally feasible. A second approach involving agronomy, plant breeding and biotechnology combined to biofortify crops could offer a more durable and acceptable solution (Cakmak, 2004; White and Broadley, 2005, 2009; Stein et al., 2007). The model crop species Brassica may be a leading candidate to resolve these global Zn deficiencies.

1.7. BRASSICACEAE

Brassicaceae is a family of globally widespread herbaceous plants or occasionally subshrubs with usually simple and alternate leaves and bisexual flowers (Carr, 2004). The family forms a monophyletic group, sister to Cleomaceae, with almost all members producing six stamens in a tetradynamous pattern (two short and four long), a cruciform actinomorphic (i.e. radial symmetry) corolla, and typical capsular fruit termed siliques (long and thin) or silicles (short and broad) (Carr, 2004; Beilstein et al., 2006). The family comprises approximately 3000 species from 338 genera, of which approximately 260 (77%) can be cautiously assigned to 25 tribes (Fig. 1.6) (Al-Shehbaz et al., 2006). Brassica oleracea and other members of the monphyletic tribe Brassicaceae, all share two unique characteristics: First, their fruits are laterally

divided into two segments (heterocarpic), and second, they all have conduplicate cotyledons (i.e. folded together around the radicle) (Beilstein et al., 2006).

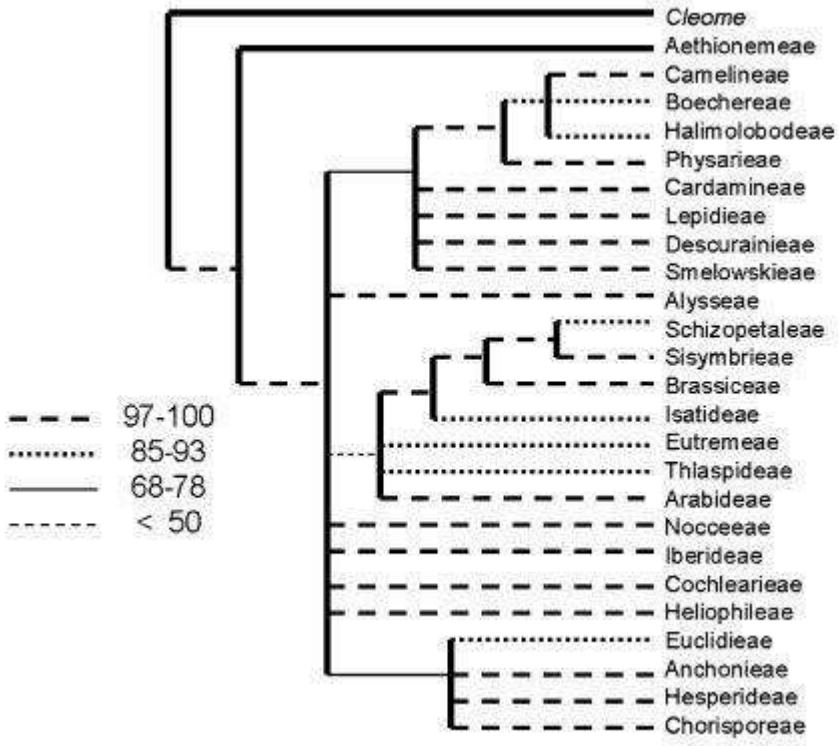


Figure 1.6 Phylogenetic relationships among tribes of the Brassicaceae. Legend represents bootstrap values. **Source:** (Al-Shehbaz et al., 2006).

The genetic proximity of Brassica ssp. L. and their ability to hybridise were originally demonstrated in the "Triangle of U" depicting how important commercial Brassica crop species were derived from a common hexaploid ancestor (Fig. 1.7) (U, 1935). All three ancestral genomes of these crops were classified into three paleaopolypoids, AA, BB, and CC representing B. rapa (2n = 2x = 20), B. nigra (2n = 2x = 16) and B. oleracea (2n = 2x = 18) respectively (U, 1935). Comparative mapping demonstrated that the genomes of these three diploids were rearranged variants of a common hexaploid ancestor (Lagercrantz et al., 1996). The interspecific breeding or fusion events between ancestors of these diploid crops naturally gave rise to three globally important allotetraploid oil crops specifically AABB (2n=4x=36) B. juncea (Indian mustard), AACC (2n=4x=38) B. napus (Rapeseed) and BBCC (2n=4x=34) B. carinata (Ethiopian mustard).

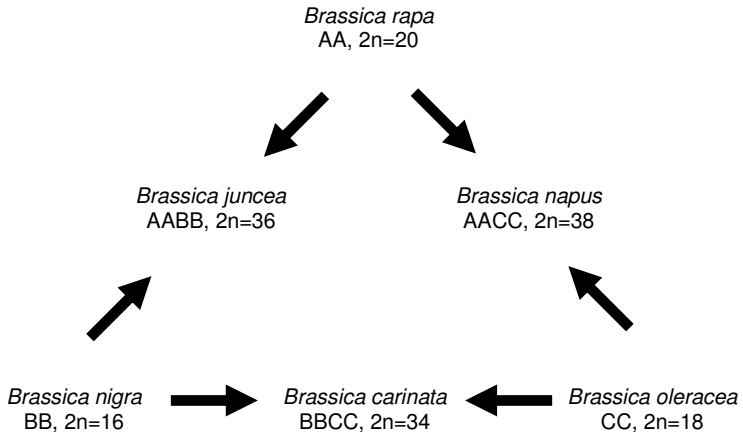


Figure 1.7 The Brassica triangle of U, representing the A, B and C genomes and their respective amphidiploids that arose from spontaneous chromosome doubling via meiotic nondisjunction after interspecific hybridizations in regions of overlapping geographical distribution of the respective diploid progenitors.
Source: Adapted from Snowdon, 2007.

1.7.1. Diverse morphology in Brassica oleracea.

Brassica oleracea is an extremely diverse crop species (GCDT, 2006) with several distinct domesticated varietas including, acephala (kale/collards), alboglabra (Oriental kale), botrytis (cauliflower), capitata (cabbage), gemmifera (Brussels sprout), gongylodes (kohlrabi), italica (broccoli/calabrese), sabauda (Savoy cabbage), and sabellica (borecole/curly kale) (Broadley et al., 2008).

The global and economic importance of this genus is profound with Brassica spp. being characterised into oilseed, vegetable and condiment crops (Paterson et al., 2001). B. napus, B. rapa and B. juncea provided 12% of the worldwide edible vegetable oil supplies (Labana and Gupta, 1993; Choi et al., 2007) and the vegetable crops directly consumed by a significant proportion of the world's population (Paterson et al., 2001) with 68.1 million t yr^{-1} of 'cabbages and other B*rassica'*, and an additional 17.6 Mt yr^{-1} of 'cauliflower and broccoli' consumed in 2005 compared with 61.4 Mt yr^{-1} of 'onions and shallots' (FAO, 2005; Broadley et al., 2008).

1.7.2. Brassica spp. as model crops to study Zn accumulation.

1.7.2.1. Brassica spp.: health and nutritional values.

Brassicaceae food crops have significant anticarcinogenic properties and all contain a rich source of bioavailable minerals and vitamins (Singh et al., 2007). A recently published WHO/FAO report recommended the consumption of a

minimum of 400 g of fruit and vegetables per day (excluding potatoes and other starchy tubers) to prevent chronic diseases such as heart disease, cancer, diabetes and obesity, and prevent micronutrient deficiencies, especially in less developed countries (WHO, 2006). A diet rich in Brassica oleracea has been associated with preventing chemically induced carcinogenesis in humans and laboratory animals and specifically, the hydrolytic products of its endogenous glucosinolates are reputed to have anticarcinogenic properties (Singh et al., 2007). Brassica spp. have a high content of antioxidants including carotenes, DL-*α*-tocopherol and ascorbate, associated with reduced cardiovascular diseases and certain cancers. Mineral content of broccoli remained unaffected by cooking, with cooked florets containing the recommended daily allowance (RDA) per portion size of all 8 minerals examined including Zn (USA RDA levels) (Moreno et al., 2007). Natural synergistic reactions among these phyto- minerals and vitamins have been demonstrated to reduce the likelihood of consumers developing many cancers and diseases (Palozza and Krinsky, 1992). Moreno et al. (2007) concluded, that exploiting natural concentration of phytonutrients through crop breeding, increased production and consumption may better serve human health, than the addition of reduced bio-available elements in supplements or salts.

1.7.2.2. Model zinc crop system: Brassica spp.

The closest crop relatives to Arabidopsis thaliana are represented by Brassica spp., with typically >85% nucleotide identity in coding regions and are therefore, the premium choice for comparative genomic approaches to understand and control biological processes and traits in crops (O'Neill and Bancroft, 2000; Lukens et al., 2003; Parkin et al., 2005). In terms of functional genomics, Brassicaceae is the most studied plant family globally, encompassing two important model systems (Beilstein et al., 2006). First, Arabidopsis thaliana (L.) Heynh. was the first flowering plant to have its entire genome sequenced and is the most extensively investigated plant model species (The Arabidopsis Genome Initiative, 2000). Second, the closely related Brassica complex (B. oleracea L., B. rapa L., B. nigra L., W. D. J. Koch, and their three hybrids) is considered the model system for agricultural crops and has been studied regarding numerous physiological aspects including, the genetics of flowering time, hybridization,

gene silencing (Schranz et al., 2002; Beilstein et al., 2006; Stephenson et al., 2010) and mineral accumulation (Broadley et al., 2008, 2010).

1.7.2.3. Colinearity between B. oleracea and A. thaliana genomes.

It was demonstrated using physical mapping that triplication of the Brassica oleracea genome followed its divergence from Arabidopsis ca. 20 million years ago (Rana et al., 2004), but significant colinearity still exists between these triplicated units (Cavell et al., 1998; O'Neill and Bancroft, 2000). In contrast however using transcriptome mapping, it was found that events such as aneuploidy and chromosomal rearrangement formed the basis of the Brassica genome and not simple polyploidization (Li et al., 2003). Colinearity was again shown between A. thaliana and B. napus (AACC) (which has conserved C and A genome marker order and linkage group structure with B. oleracea and B. rapa respectively) (Parkin et al., 2005). It was demonstrated that the average length of conserved blocks between the two genomes was 14.8 cM in B. napus and 4.8 Mb in A. thaliana and 50% of the mapped lengths of at least seven B. napus linkage groups were equivalent to corresponding conserved regions of the A. thaliana genome (Parkin et al., 2005). In A. thaliana, 21 conserved blocks could be replicated and rearranged to cover 90% of the mapped length of B. napus and since their divergence, only 74 large rearrangements (38 in the A genome and 36 in the C genome) separate B. napus from A. thaliana (Parkin et al., 2005).

1.7.2.4. Natural genetic variation in Zn content.

Brassica oleracea has enormous natural potential to accumulate Zn in its edible fractions. A study into species wide variation of minerals in staple crops including rice, wheat, maize, beans, cassava and B. oleracea was reviewed by White and Broadley (2005), demonstrating that for both Zn and Fe, in a B. oleracea core collection (1089 mg kg^{-1} $[Fe]_{shoot}$ and 229 mg kg^{-1} $[Zn]_{shoot}$), concentrations were higher than that of any other crop excluding spinach (387 mg kg^{-1} $[Zn]_{shoot}$). The species wide Fe and Zn variation was similarly highest in B. oleracea with a 30 $[Fe]_{shoot}$ and 19 $[Zn]_{shoot}$ fold variation compared with a maximum variation of 4.5 fold $[Fe]_{shoot}$ Pea, 6.6 $[Zn]_{shoot}$ Pea and 4 $[Zn]_{shoot}$ among all cereals. It is clear that the natural genetic diversity of B. oleracea offers the greatest immediate alternative for crop breeders to enhance Zn

accumulation for Zn deficient human populations. Analyses of doubled haploid (DH) mapping populations in both B. oleracea (Broadley et al., 2010) and B. rapa (Wu et al., 2001, 2008) have identified significant quantitative trait loci (QTLs) involved in Zn accumulation in leaf tissues. With the advent of comparative QTL mapping experiments, well characterised physical maps of related Brassica spp. and the fully sequenced A. thaliana genome, should hasten the elucidation of loci/genes and pathways involved in Zn uptake in food crops.

1.7.2.5. Micronutrient inhibitor and promoter substances.

Micronutrient inhibitors referred to as antinutrients are substances which hinder Zn metabolism and absorption in the gut (White and Broadley, 2009). Some, such as phytate and tannins are present in many plant tissues but are concentrated in seeds or grains, including staple cereals (Welch and Graham, 2004; Broadley et al., 2006). Many cereals accumulate Zn in their bran which is entirely removed during milling and despite transgenic approaches to increase the Zn and Fe content in the endosperm of the grains i.e. high expression of the soybean ferritin gene, increased levels of Zn bioavailability have not yet been confirmed (Vasconcelos et al., 2003). In a study addressing the nutrient content of all major global staple crops, it was shown that cereals (maize, rice and wheat) and legumes (peas and beans) contain more endogenous anti-nutrients than B. oleracea crop species (Graham et al., 2001; Welch and Graham, 2004). Additionally it was demonstrated that plant based products which improved Zn and Fe uptake were 'fresh fruits and vegetables' and 'green and orange vegetables' producing organic acids e.g. ascorbic acid β-carotene – abundantly available in B. oleracea (Graham et al., 2001).

1.8. ZINC IN PLANTS

1.8.1. Zinc fluxes into plants

Roots transport nutrients to the shoots through the xylem after obtaining typically Zn^{2+} and some Zn organic ligands from the soil solution. Most plant species demonstrate a Michaelis-Menten function, with an affinity constant (K_m) of 1.5 – 50 µm or higher and a V_{max} (the rate at $[Zn]_{ext} = \infty$) of 5.74 µmol g^{-1} FW h^{-1} (Broadley et al., 2007). Two studies however, have described wheat plants with K_m values between 0.6 – 2.3 nM for Zn^{2+} uptake (Wheal and Rengel, 1997;

Hacisalihoglu et al., 2001) with little difference between K_m and V_{max} for both Zn-efficient and Zn-inefficient wheat varieties, implying Zn influx kinetics have an insignificant function in Zn efficiency in wheat (Hart et al., 1998, 2002; Hacisalihoglu et al., 2001; Hacisalihoglu and Kochian, 2003) contrary to variations observed between Zn-efficient and Zn-inefficient sugarcane, rice and tomato genotypes (Bowen, 1973, 1986, 1987). Generally, a lower V_{max} and comparable K_m value exists for Zn-satiated plants compared to Zn-deficient plants (Rengel and Wheal, 1997; Hacisalihoglu et al., 2001).

Zinc may enter plants either through a symplastic pathway involving movement from epidermal and cortical cells to the root xylem via cellular cytoplasm channelled by plasma membrane and tonoplast transport activity (Lasat and Kochian, 2000), or extracellularly, to the stelar apoplast through underdeveloped Casparian bands (White, 2001; White et al., 2002), involving a pathway controlled by cell wall cation exchange capacity, water flow and Casparian band development (Sattelmacher, 2001). Uptake of Zn is also affected by interactions with other macro and micronutrients such as Zn-copper (Zn-Cu) interactions, which are mutually inhibititory, implying that both are absorbed by the same uptake mechanisms (Alloway, 2004). Cadmium uptake in Zn-deficient rice (Neue et al., 1998) and durum wheat (Koleli et al., 2004) significantly increases, with little increase in translocation to the shoots, suggesting the presence of a high affinity Zn transporter. (Neue et al., 1998) also demonstrated that increases in Zn had a negative correlation with Fe and P levels but manganese (Mn) levels increased. In rice seedlings, translocation of Zn from roots increases with Mn application. In Brassica juncea it was demonstrated that increases in N and to a much lesser extent P, positively affected $[Zn]_{shoot}$ content (Hamlin et al., 2003).

The Zn hyperaccumulator Noccaea caerulescens (J. & C. Presl.) F.K.Mey. accumulates up to 30,000 µg Zn g^{-1} DW shoot, suggesting enormous kinetic fluxes occurring at high $[Zn]_{ext}$ to deliver Zn via the symplast to the root xylem (White et al., 2002). It has also been reported that in order to catalyse the export of solely Ca^{2+}, a typical dicot would require more Ca^{2+}-transporting ATPase activity than its total membrane protein availability (White, 1998; White et al., 2002). Despite this, N. caerulescens accumulates considerably more shoot Ca, Mg and Zn > 60,000 µg g^{-1} DW than most dicots (Meerts et al., 2003; Dechamps

et al., 2005). A major factor could be Casparian band development since Cd^{2+} fluxes in wheat, N. caerulescens and T. arvense L. are superior at the root tips, where Casparian bands remain underdeveloped (Piñeros et al., 1998).

1.8.2. Zinc hyperaccumulation.

Zn hyperaccumulation in hypertolerant metallophyte aerial fractions was first reported in 1855 in Viola calaminaria (Ging.) Lej. which contained >1% Zn in its ash followed in 1865 by Thlaspi alpestre L. (N. caerulescens) containing > 1% $[Zn]_{shoot}$ (Reeves and Baker, 2000). The phrase hyperaccumulator was primarily applied to plants containing more than 1000 µg g^{-1} (0.1%) Ni DW (Brooks et al., 1977; Wither and Brooks, 1977) and later Thlaspi L. sensu lato species hyperaccumulating Zn above 10,000 µg g^{-1} (1%) DW (Reeves and Brooks, 1983). Zn hyperaccumulation occurs in 10 – 20 species, the majority of which are within the genus Thlaspi sensu lato (s.l.) (Brassicaceae), with few additionally hyperaccumulating Cd (Broadley et al., 2007). Seed coat, anatomical and sequence data suggests that Thlaspi L. s.l is probably polyphyletic being divided into distinct genera including Thlaspi sensu stricto (s.s.), Vania and a clade containing Thlaspiceras, Noccaea, Raparia, Microthlaspi and Neurotropis (Fig. 1.8) (Mummenhoff et al., 1997; Koch and Mummenhoff, 2001; Koch and Al-Shehbaz, 2004).

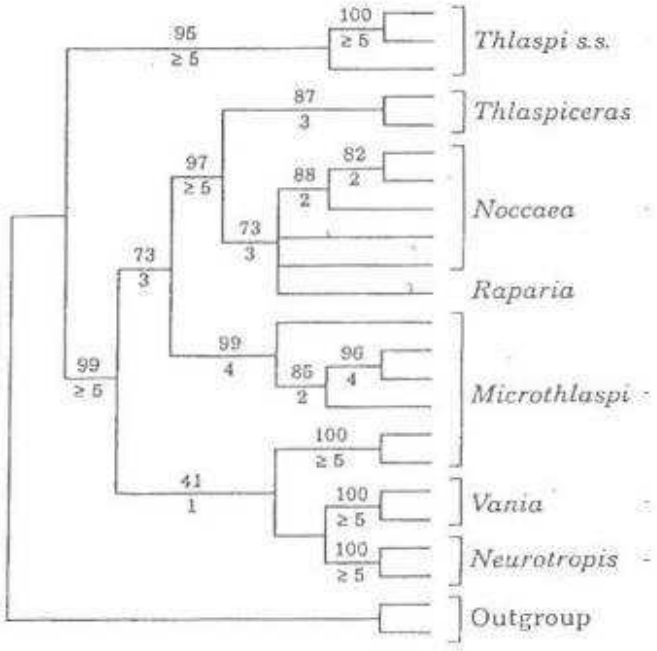


Figure 1.8 Noccaea phylogeny based on ITS sequences. Numbers above branches indicate bootstrap values in percent. Zinc accumulation is hypothesised to have evolved in the section Noccaea and Raparia (Reeves and Brooks, 1983). Noccaea, Raparia and Thlaspiceras form a monophyletic clade.
Source: (Mummenhoff et al., 1997; Tayler, 2004).

The high $[Zn]_{shoot}$ characteristic is assumed to have evolved at the base of the Noccaea/Raparia clades, since species from both express this characteristic (Macnair, 2003; Tayler, 2004) unlike their non-hyperaccumulator related clade Thlaspiceras (Zn hypertolerant) or more distantly related Microthlaspi and Neurotropis (Zn non-hypertolerant) (Broadley et al., 2007; Verbruggen et al., 2009; Krämer et al., 2010). Other plausible proposals include a Zn hyperacccumulation evolution at the base of the Noccaea/Raparia/Thlaspiceras clade followed by a reversion to low $[Zn]_{shoot}$ characteristics in the Thlaspiceras clade, or Ni hyperaccumulation at this base followed by modification to Zn hyperaccumulation in the Noccaea/Raparia clades (Macnair, 2003; Tayler, 2004). More than 400 plant taxa globally are known to hyperaccumulate one or more heavy metals with >300 species in ca. 34 families accumulating nickel and >80% of temperate Ni hyperaccumulaters from Brassicaceae (Reeves and Baker, 2000; Davis et al., 2001; Verbruggen et al., 2009). The presence of the relatively distantly related Arabidopsis halleri as the sole recorded non-Noccaea Zn

hyperaccumulator, implies the occurrence of two relatively recent evolutionary events with possibly other similar events occurring elsewhere in Brassicaceae (Broadley et al., 2007; Verbruggen et al., 2009; Krämer et al., 2010). Interest in this evolutionary feature has heightened recently with the feasibility of transferring these traits to crops with low Zn in their edible fractions (biofortification) as well as opportunities for phytoremediation and phytomining (Chaney, 1983; Baker and Brooks, 1989; Guerinot and Salt, 2001; Macnair, 2003; Barabasz et al., 2010).

1.8.3. Background to N. caerulescens and Zn hyperaccumulation.

The vast majority of known Zn hyperaccumulator genera exist in Brassicaceae, the majority of which form part of the Nocceae/Raparia clades (Section 1.5.4.). Of these, Noccaea caerulescens (2n=2x=14) has received the most physiological attention (Broadley et al., 2007). The short lived, self compatible, non-mycorrhizal biennial (Regvar et al., 2003) is a close relative of the crop model species Arabidopsis thaliana L. (2n=2x=10), with approximately 88.5% DNA identity in coding regions (Rigola et al., 2006) and 87% DNA identity in the intergenic transcribed spacer regions (Peer et al., 2003).

Distributed throughout northern, western and central Europe, it is constitutively Zn hypertolerant and hyperaccumulates foliar Zn up to 3% dry weight (d.w.) (30000 µg g^{-1} d.w.) according to inter- and intra- population variations and often either Cd or Ni at >0.1% d.w. (1000 µg g^{-1} d.w.) (Baker and Brooks, 1989) equating to 100-300 fold more Zn and 100 fold more Cd than most species under normal conditions (Baker et al., 2000; Reeves and Baker, 2000; Bert et al., 2002; Hammond et al., 2006).

Under identical growing conditions however, high-$[Zn]_{soil}$ populations typically accumulate lower $[Zn]_{shoot}$ than low-$[Zn]_{soil}$ populations among U.K (Pollard and Baker, 1996), and European (Roosens et al., 2003) populations. Similarly low-$[Zn]_{soil}$ populations demonstrated lower Zn tolerance but higher $[Zn]_{shoot}$ than high-$[Zn]_{soil}$ populations under soil-based experiments (Escarré et al., 2000). Less significant variation existed in $[Zn]_{shoot}$ between and within half-sib

families, than between populations, substantiating the self fertilizing character of N. caerulescens (Broadley et al., 2007).

Assunção et al. (2006) demonstrated substantial G x E interaction (phenotypic effect of interactions between genes and the environment) affecting $[Zn]_{shoot}$ with populations from a nonmetalliferous (Lellingen, Luxembourg [LE]) site, a serpentine (Monte Prinzera, Italy (MP)) site, and high-$[Zn]_{soil}$ populations from two calamine sites (La Calamine, Belgium [LC] and Ganges, France [GA]). $[Zn]_{shoot}$ decreased GA ≥ LC > LE > MP in field samples and MP > GA > LE > LC in hydroponic conditions with Zn tolerance decreasing GA = LC> MP > LE. Zn accumulation was shown to be a polygenic trait, however genetic independence of Zn tolerance and accumulation in N. caerulescens could not be confirmed following F_3 segregation of an intra specific cross between LE and LC due to the negative correlation between the two traits (Broadley et al., 2007; Verbruggen et al., 2009). Following genetic mapping of these crosses it was shown that QTLs were mapped for $[Zn]_{root}$ (on chromosomes 3 and 5) (Assunção et al., 2006) and later three QTLs explained 44.5% of $[Zn]_{shoot}$ variance in an F_2 population from an LC X GA cross (Deniau et al., 2006). Similar intra specific crossing experiments have shown that F_1 progeny from a high-$[Zn]_{soil}$ x low-$[Zn]_{soil}$ population, were more sensitive to Zn deficiency, however F_1 and F_2 generations demonstrated more tolerance to high $[Zn]_{ext}$ but contained lower $[Zn]_{shoot}$ than progenies from exclusively low-$[Zn]_{soil}$ population crosses (Zha et al., 2004; Frérot et al., 2005). Interestingly a number of studies have shown that $[Cd]_{ext}$ negatively affects $[Zn]_{shoot}$ of Zn/Cd hyperaccumulator (N. caerulescens Ganges) but positively affects $[Zn]_{shoot}$ of Zn hyperaccumulator (N. caerulescens Prayon) whilst high $[Zn]_{ext}$ inhibits Cd accumulation in Prayon, but not in Ganges populations (Lombi et al., 2001; Zhao et al., 2002; Roosens et al., 2003; Zha et al., 2004).

1.8.4. Zn accumulation mechanism in N. caerulescens.

The extraordinary ability for N. caerulescens to hyperaccumulate zinc has received attention from authors wishing to understand the system involved. It has long been considered that three mechanisms may be involved in the high uptake of metals by hyperaccumulators including:

a) root secretions of metal solubilising protons, organic acids and/or chelators;
b) improved absorption of zinc into the roots (Lasat et al., 1996) and elevated rates of metal translocation from roots to the shoots (Shen et al., 1997);
c) and root metal foraging, involving preferential allocation of root biomass into regions of metal enrichment (Whiting et al., 2000).

1.8.4.1. Root exudation.

It was demonstrated that oat plants excreted mucilage to take up Zn from dry soil but its efficacy was 60% lower than uptake from wet soil (Nambiar, 1976). Several investigators since have examined the rhizospheres of hyperaccumulator plants but found no evidence for strong acidification (Bernal et al., 1994; McGrath et al., 1997; Salt et al., 2000). Both N. caerulescens Ganges and Prayon did not demonstrate significantly greater quantities of root exudates than Brassica napus (canola) and wheat and exudates released by all of the Brassicaceae examined contained no significant amounts of chelating compounds with a high affinity for Zn, Cd, Cu and Fe from a contaminated soil (Zhao et al., 2001). These findings were consistent with other non-hyperaccumulating species, that produced increased root exudation of amino acids, sugars and phenolics under Zn deficiency but did not enhance Zn mobilization (Zhang et al., 1991).

1.8.4.2. Root zinc absorption and elevated root – shoot Zn translocation.

Though investigations indicate little likelihood that Zn hyperaccumulation in N. caerulescens is related to root exudation, evidence suggests that it has a higher concentration of Zn transporters in the plasma membranes of root cells. Compared to the non-hyperaccumulator Thlaspi arvense, N. caerulescens exhibited a 4.5-fold higher V_{max} for Zn influx (Lasat et al., 1996) and higher expression of ZNT1 (a Zn transporter gene) in its roots (Pence et al., 2000). It has two extraordinary root features which appear to facilitate Zn-hyperaccumulation and hypertolerance, including an ability to actively search for available Zn in the soil (root Zn foraging) and a unique anatomical root form.

1.8.4.3. Zincophilic roots.

Several studies have shown that N. caerulescens roots have an extraordinary ability to actively forage the soil continuum for heterogeneous Zn containing

sites comparable to macronutrient-directed root growth (Schwartz et al., 1999; Whiting et al., 2000; Haines, 2002). When compared with low $[Zn]_{soil}$ conditions, N. caerulescens roots in elevated $[Zn]_{soil}$ displayed increased length and root hair proliferation (Whiting et al., 2000; Haines, 2002). Approximately 67% of N. caerulescens 'Prayon' roots were located in Zn enriched sections however the total $[Zn]_{shoot}$ accumulation was not significantly different to plants grown on homogeneous Zn treatments (Whiting et al., 2000; Haines, 2002). Conversely, N. caerulescens 'Bradford Dale' displayed no zincophilic activity and had significantly less $[Zn]_{shoot}$ accumulation when grown in heterogeneous $[Zn]_{soil}$ compared with those grown on homogenous $[Zn]_{soil}$ (Haines, 2002). Zincophilic root plasticity appears to be an important though not solitary ecotype dependent trait, with some ecotypes demonstrating possible variations in Zn requirements or genetic selection affecting their response capabilities to $[Zn]_{soil}$ heterogeneity (Haines, 2002).

1.8.4.4. Endodermal development and Peri-endodermal layer.

N. caerulescens roots contain a "peri-endodermal" cell layer, each surrounded by a thick secondary inner tangential cell wall highly concentrated with suberin/lignin deposits. Together these cells form an enclosure around the perimeter of the endodermal cells extending to within 1mm of the root tip (Broadley et al., 2007) (Fig. 1.9 and 1.10). Even more interestingly, this phenomenon is absent in both non-hyperaccumulator Brassicaceae species T. arvense and Arabidopsis thaliana (van de Mortel et al., 2006; Broadley et al., 2007).

Intriguingly it has been shown that endodermal cell development and location differ markedly between T. arvense and N. caerulescens. In N. caerulescens, Casparian bands (initial stage of endodermal development) are found within 1 mm of the root tip and suberin lamellae (the second stage of endodermal development) are formed in all endodermal cells ca. 5 – 6 mm from the apex (Broadley et al., 2007). In contrast T. arvense Casparian bands develop ca. 2 mm from the root tip and suberin lamellae are formed in all endodermal cells ca. 10 mm from the root tip (Broadley et al., 2007) (Fig. 1.9).

The function of these phenomena within N. caerulescens remains speculative however similar peri-endodermal formations were observed within Brassicaceae by early anatomists, describing it as a 'réseau sus-endodermique' (Broadley et al., 2007). Indeed the closely related Brassicaceae halophyte Thellungiella halophila O. E. Schulz, forms a second endodermal layer that may partially explain its inherent salt tolerance (Inan et al., 2004). It would seem logical therefore that additional anatomical studies of metallophytes within and beyond the Thlaspiceras/Noccaea/Raparia/Microthlaspi/Neurotropis clade should clarify the possible impact of these cell developments on apoplastic and symplastic Zn fluxes into and from the stele (Broadley et al., 2007).

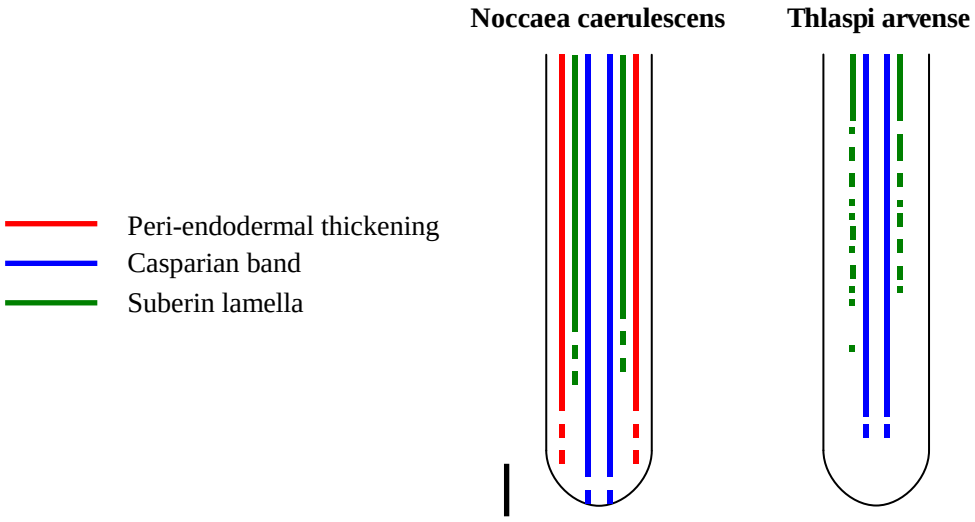


Figure 1.9 Development scheme of apoplastic barriers along the root axis of Noccaea caerulescens and Thlaspi arvense. In the 1st stage of endodermal ontogenesis, Casparian bands, and in the 2nd stage, suberin lamellae deposition, develop closer to the root tip in N. caerulescens than in T. arvense. Note the early formation of a peri-endodermal layer close to the apex in N. caerulescens but not in T. arvense. Bar, 2 mm in the longitudinal axis. Transverse axes are not represented to scale.
Source: (Broadley et al., 2007).

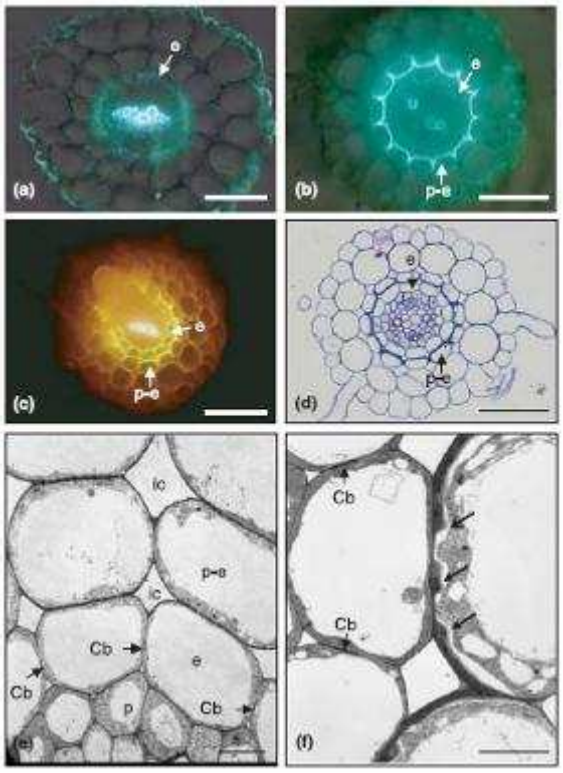


Figure 1.10 Structure of primary roots in cross-sections of Thlaspi arvense (a) and N. caerulescens (b–f). (a) Autofluorescence of T. arvense root section showing typical dicotyledonous root structure with three cortical layers; the innermost is the endodermis exhibiting fluorescence of Casparian bands. (b) Autofluorescence of N. caerulescens root section showing the intensive fluorescence of inner tangential and radial walls of the peri-endodermal layer and light fluorescence of thin endodermal cell walls. (c) N. caerulescens primary root with emerging lateral root primordium influorescence microscope. After Fluorol yellow-toluidine blue staining, endodermal cell walls exhibit intensive bright yellow fluorescence as a result of suberin lamellae deposition (second stage of endodermal development). (d) Semithin section of Spurr embedded N. caerulescens root showing prominent wall ingrowths of the peri-endodermal layer (toluidine blue-basic fuchsin staining). (e) Transmission electronmicroscopy (TEM) of the young root part of N. caerulescens (< 1 mm from the apex) stained by $KMnO_4$ to visualize Casparian bands, which are present in radial walls in endodermis, but absent in the peri-endodermal layer. (f) TEM of the older root part of N. caerulescens (ca. 2.5 mm from the apex) with irregular wall thickening in the peri-endodermal layer. Cb, Casparian band;e, endodermis; ic, intercellular space; p, pericycle; p-e, peri-endodermal layer; arrow, wall ingrowths in peri-endodermal cells. Bars, 50 µm (a–d); 4 µm (e, f).
Source: (Broadley et al., 2007).

1.8.5. Compartmentalisation of Zn in hyperaccumulators.

Typically, in Zn hyperaccumulators, $[Zn]_{shoot}$ is 10 fold greater than $[Zn]_{root}$ depending on $[Zn]_{ext}$ (Shen et al., 1997; Roosens et al., 2003). In plants grown at low $[Zn]_{ext}$, vacuoles of epidermal and mesophyll cells, and leaf cell walls were shown to contain 296, 80 and ca. 110 µg Zn g^{-1}DW, respectively using scanning transmission electron microscopy (STEM) energy-dispersive X-ray microanalysis (EDXMA) (Vazquez et al., 1992). At high $[Zn]_{ext}$, Zn accumulated

in the vacuoles of leaf mesophyll cells and their cell walls at concentrations of 4610 and 333 μg Zn g^{-1} DW, respectively.

The Zn concentration ratio in root epidermal and subepidermal cells at low $[Zn]_{ext}$ was 1 (56 μg Zn g^{-1} DW) : 1.7 : 3 : 3.5 (vacuole centre : vacuole periphery : cell wall : intercellular space) respectively and at high $[Zn]_{ext,}$ background vacuolar Zn concentration was 1 (175 μg Zn g^{-1} DW): 0.5 : 1 (background vacuole : cell wall : intercellular space) respectively (Vazquez et al., 1992). $[Zn]_{ext}$ only therefore appeared to affect vacuolar Zn concentrations while later, Küpper et al. (1999) demonstrated that vacuolation actually promoted Zn accumulation. In Arabidopsis halleri, Zn was concentrated in the trichome bases (up to 1 M) and unlike N. caerulescens mesophyll cells had higher concentrations of Zn than smaller-vacuoled epidermal cells (Küpper et al., 2000; Zhao et al., 2000). With ultrathin (100 nm) cryosections of N. caerulescens and using STEM-EDXMA (scanning transmission electron microscope - energy dispersive X-ray micro-analysis) techniques (Frey et al., 2000), it was demonstrated that Zn concentrations in upper and lower leaf epidermal cells (79% in cell walls) was substantially greater than in leaf mesophyll, guard, or root cortical cells while extremely low concentrations of cytoplasmic Zn was revealed using Newport green diacetate assays (Marquès et al., 2004). Epidermal : mesophyll Zn concentrations were documented at 2.5:1 by Ma et al. (2005) however, mesophyll cell walls contained only 23% of the total cellular Zn content.

1.8.6. Zn storage form in hyperaccumulators.

Unlike many crop species, 80% of Zn is water or weak acid soluble in both N. caerulescens and A. halleri, with very little stored as insoluble Zn : P complexes such as $Zn_3(PO_4)_2$ and Zn-phytates (Zhao et al., 2000; Sarret et al., 2002). In hyperaccumulators, the role of inorganic cation and organic anion millimole equivalents, i.e. the quantity capable of reacting quantitatively with 1 mmole CO, correlates significantly in N. caerulescens shoots with soluble Zn, as well as carboxylic and amino acids (Salt et al., 1999; Küpper et al., 2000; Zhao et al., 2000; Ma et al., 2005). Speculative evidence supports a putative Zn-malate shuttle hypothesis across the tonoplast followed by dissociation and Zn^{2+} binding to stronger chelators such as citrate or oxalate (Broadley et al., 2007), with Zn

stimulating root citrate production in N. caerulescens (Shen et al., 1997). Using X-ray absorption spectroscopy (XAS) and extended X-ray absorption fine structure (EXAFS) analysis it was found that 70% of root Zn could be connected with histidine (His), and 30% with the cell wall (Salt et al., 1999). Zn in xylem saps, was either bound to citrate (21%) or as free Zn^{2+} while citrate, oxalate, His and the cell walls bound 38, 9, 16 and 12% of total Zn respectively in the shoots, with 26% as free Zn^{2+} (Salt et al., 1999). Other investigations found Zn co-ordination with O ligands, potentially representing hydroxyl groups of water, or potentially carboxyl groups of malate, citrate or other organic acids, however none have found associations between Zn-malate, S ligands or other Cys-rich peptides (Broadley et al., 2007).

1.8.7. Molecular aspects of Zn hyperaccumulation

Development in the understanding of Zn homeostasis and hyperaccumulation at the molecular level is being advanced phenomenally through gene and protein expression profiling (Alloway, 2004; Filatov et al., 2006; Hammond et al., 2006; Rigola et al., 2006; Talke et al., 2006; van de Mortel et al., 2006) and through functional analyses of proteins in heterologous plant, yeast and Xenopus laevis oocyte systems (Hussain et al., 2004; Papoyan and Kochian, 2004; Williams and Mills, 2005; Hanikenne et al., 2008; Barabasz et al., 2010; Siemianowski et al., 2010).

Analysis of global shoot transcriptome data published using the full-genome Affymetrix A. thaliana ATH1-121501 (ATH1) GeneChip array for N. caerulescens vs T. arvense (Hammond et al., 2006), A. halleri vs A. lyrata ssp. petraea (Filatov et al., 2006) and A. thaliana (Talke et al., 2006) found that homologues of 60 A. thaliana genes were significantly differentially expressed between hyperaccumulators and non hyperaccumulators and may have conserved roles in brassicaceous Zn hyperaccumulation (Broadley et al., 2007). Six of these genes were found to have functions in Zn transport specifically:

- three cation diffusion facilitator (CDF) family members: (At2g39450 (TAIR6: AtMTP11), At2g46800 (AtZAT1/AtMTP1) and At3g58060 (AtMTPc3));
- a Zn-Fe permease (ZIP) family member (At1g60960 (AtIRT3/TcZNT2))

- and a P_{1B}-type heavy-metal-associated domain-containing ATPase (At4g30120 (AtHMA3)).

CDFs appear to regulate efflux of Zn from the cytoplasm to the vacuole (Persans et al., 2001; Blaudez et al., 2003; Hall and Williams, 2003; Arrivault et al., 2006; Colangelo and Guerinot, 2006).

ZIPs appear to control cellular Zn uptake and are primarily up regulated in response Zn deficiencies (Pence et al., 2000; Assunção et al., 2001; Colangelo and Guerinot, 2006; Krämer et al., 2010).

HMAs control the transport of Zn throughout the cell (Williams et al., 2000; Hussain et al., 2004; Papoyan and Kochian, 2004; Verret et al., 2004; Abdel-Ghany et al., 2005; Colangelo and Guerinot, 2006; Hanikenne et al., 2008; Barabasz et al., 2010; Siemianowski et al., 2010).

Other genes with an unknown function which appear highly expressed in Zn hyperaccumulators included a nutritionally important phosphate-starvation-inducible high-affinity phosphate transporter AtPT2/AtPHT4/AtPHT1;4 (Hammond et al., 2006; Talke et al., 2006).

Plant defensin genes (PDFs) were similarly highly expressed in Zn hyperaccumulators and are known to regulate Zn tolerance and accumulation in heterologous systems, and so may function by impeding Zn-permeable channels (Mirouze et al., 2006; Broadley et al., 2007).

1.9. How plants transport Zn: P_{1B}-ATPases

Many transition metals have vital functions in plant growth and development, however at high levels these metals including Cu, Zn and Mn can be phytotoxic and must be eliminated from sensitive cellular structures through compartmentalisation, chelation and exclusion (Mills et al., 2005). In order to achieve this, transmembrane metal transporting proteins such as P_{1B}-ATPases are employed for transition metal homeostasis and are divided into those that transport either Cu/Ag or Zn/Cd/Co/Pb (Mills et al., 2005). These transporters, pump cations across membranes against their electrochemical gradient through energy from ATP hydrolysis, similar to other P-type ATPases. Unlike other P-type ATPases however, P_{1B}-ATPases have eight transmembrane helices, a CPx/SPC motif in transmembrane domain six, with possible transmembrane binding domains at the N- or C-termini (Mills et al., 2005). Knockout mutants

have been successfully employed to determine their function in planta while lethal effects experienced through deleting more than one gene demonstrates the importance of their role in transition metal transport (Mills et al., 2005). Metals transported by these P_{1B}-ATPases have been exposed through their expression in E.coli and yeast, while methods of this metal transportation have been shown in yeast, bacteria and Arabidopsis, e.g. AtHMA6/PAA1 and AtHMA8/PAA2 are both closely related and involved in transporting Cu to some chloroplast proteins. Mutant studies involving gene knock-out has proved invaluable in determining their function e.g. paa1 paa2 double mutants are seedling lethal (Abdel-Ghany et al., 2005). They also demonstrated that AtHMA6/PAA1 was present in both shoots and roots, indicating a function in both green and non-green plastids. AtHMA8/PAA2 however was thought to have a function in supplying metals to the thylakoid membrane of chloroplasts (Abdel-Ghany et al., 2005). The putative locations of these P_{1B}-ATPases are shown in Figure 1.11.

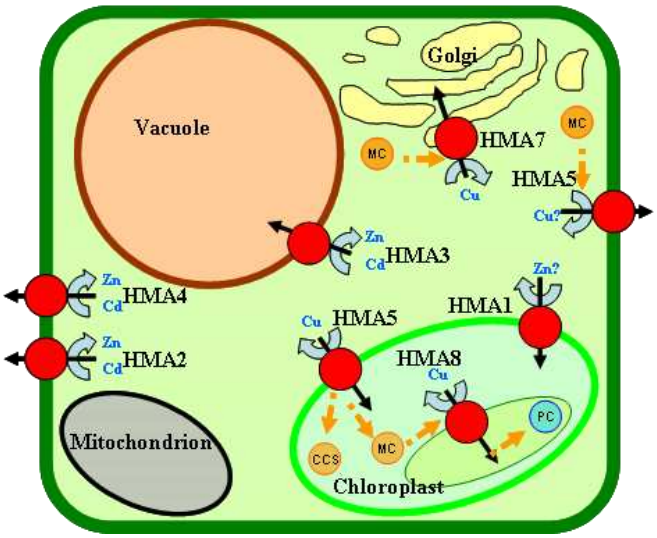


Figure 1.11 Putative sub-cellular locations for Arabidopsis P_{1B}-ATPases in a simplified cell. P_{1B}-ATPases = red circles, together with their physiological substrate(s) (in blue), and possible interactions with metallochaperones (MC, orange circles). Putative locations of AtHMA3 at the tonoplast, AtHMA5 at the plasma membrane and AtHMA1 at the chloroplast positions. AtHMA6 and AtHMA8 transport Cu across the plastid envelope and thylakoid membrane respectively, to Cu-requiring proteins such as plastocyanin (PC, dark-blue circle) and the stromal Cu/Zn superoxide dismutase2 (CSD2, light-blue rectangle). Whether a metallochaperone exists in plant chloroplasts for Cu transfer between AtHMA6 and AtHMA8 is not known, but a Cu chaperone (CCS, orange oval) has been identified that might supply CSD2. AtHMA7 is proposed to transport Cu delivered from a cytoplasmic metallochaperone into a post-Golgi compartment to supply membrane-targeted ethylene receptor apoproteins (ETR1, cream oval) that can then coordinate ethylene. AtHMA4 and AtHMA2 appear to be efflux carriers of Zn and Cd at the plasmalemma.
Source: adapted from Mills et al. (2005).

It has been shown in the cyanobacterium Synechocystis that two P_{1B}-ATPases, which are considered functional homologues to AtHMA6 and AtHMA8, deliver Cu to plastocyanin via a metallochaperone - a protein which binds and delivers metals to their required site of reaction. It is not known whether there is a functional homologue to this metallochaperone in Arabidopsis (Tottey et al., 2002). A number of metallochaperones have been identified in Arabidopsis however such as AtCCH1 and AtCOX17. One chaperone, ahs, has a putative role of delivering Cu to Cu/Zn superoxide dismutase 2 (AtCSD2) (Mills et al., 2005). Some authors suggest that these systems identified in other organisms are conserved in plants (Rosenzweig, 2002). It has not however been confirmed whether chaperones exist for metals such as Zn or if any interaction occurs between the Zn transporting P_{1B}-ATPases.

1.9.1. HMA4 – Nutritional role.

Functional characterisation of AtHMA4 in yeast and bacteria lead to the suggestion that it had a function in Zn efflux at the plasma membrane (Mills et al., 2005). Through the creation of a double mutant hma2,hma4 it was shown that these genes had putative roles with zinc homeostasis in plant cells (Hussain et al., 2004). In order to rescue the mutant exogenous applications of Zn were required, and later promoter-GUS constructs demonstrated that AtHMA4 was expressed in cells surrounding the root vascular tissue. This suggested that AtHMA2 and AtHMA4 were invoved in xylem loading of Zn for transportation to the shoot (Hussain et al., 2004).

AtHMA2 and AtHMA4 have other putative roles in plant nutrition as they are found throughout the vascular tissue of the plant including the phloem, developing anthers and in developing siliques where they are believed to supply Zn e.g. to the reproductive tissues (Hussain et al., 2004). Structural features of the AtHMA4 P_{1B}-ATPase are shown in Figure 1.12

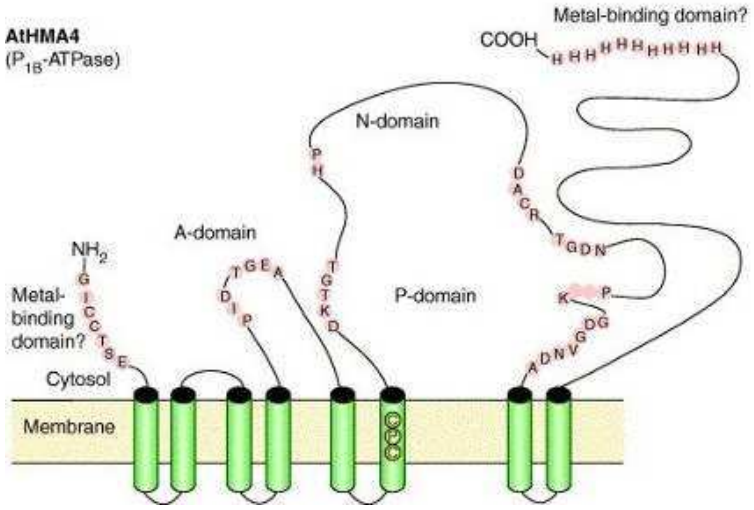


Figure 1.12 Structural features of plant P_{1B}-ATPase HMA4 Main characteristics of P_{1B}-ATPases that were inferred from the crystal structure of P_{2A}-ATPase, SERCA1A .
Source:(Mills et al., 2005)

1.9.2. HMA4 and other P_{1B}-ATPases involved in Zn hyperaccumulation.

Studies in heterologous systems such as yeast have revealed other putative functions for AtHMA4 such as metal detoxification by efflux transport of Cd and Zn (Mills et al., 2005). It has been suggested that AtHMA4 transfer high Zn and Cd levels from the roots to less detrimental areas such as the leaves. The

transcription of the pump involved in delivering high Zn levels to the shoots of Arabidopsis thaliana is regulated by Zn presence, however it is not known if it is post – transcriptionally controlled (Mills et al., 2005).

AtHMA3, another member of the P_{1B}-type ATPase Zn/Cd/Co/Pb subgroup, appears to be involved in metal detoxification through sequestration of Cd and Pb in yeast vacuoles (Gravot et al., 2004).

NcHMA4, an orthologue of AtHMA4 and endogenous to Noccaea caerulescens seems to have a similar role to AtHMA4. It is expressed in the root xylem vasculature and becomes upregulated when Zn or Cd is present (Hammond et al., 2006). Unlike AtHMA4 however, NcHMA4 also tend to show expression when plants are grown on zinc deficient substrate (Mills et al., 2005; Hammond et al., 2006). A similar observation was recorded by Talke et al. (2006), where HMA4 transcript levels were between 4- and 10-fold higher in roots and a minimum of 30-fold higher in shoots of the Zn hyperaccumulator A. halleri than in the non-hyperaccumulator A. thaliana. However HMA2 transcript levels were significantly lower than HMA4 in both species with the hyperaccumulator showing lower levels than A. thaliana, thus suggesting a prevalence of HMA4 expression in A. halleri over HMA2. Their data suggested that A. halleri HMA4 was involved in Zn and Cd hypertolerance and Zn shoot partitioning. Further experimental data by Hanikenne et al. (2008) demonstrated that silencing AhHMA4 expression through AhHMA4RNAi transformation significantly reduced $[Zn]_{shoot}$ concentrations to between 12 – 35% of concentrations found in the wild type Zn hyperaccumulator. Interestingly these levels were similar to $[Zn]_{shoot}$ found in the related non-hyperaccumulator A. thaliana. Roots of A. halleri RNAi lines contained between 49 – 134- fold more $[Zn]_{root}$ than the wild type A. halleri, which were again comparable to levels found in A. thaliana, demonstrating that for Zn hyperaccumulation, A. halleri requires high AhHMA4 transcription for such efficient root – shoot Zn flux. Conversely A. thaliana demonstrates much lower transcription levels of AtHMA4 to maintain Zn-dependent developments in the shoot (Hussain et al., 2004). It was also demonstrated that higher concentrations of Zn were distributed in the xylem vessels, inward from the pericycle layer of the wildtype (WT) A. halleri root vasculature than in the RNAi line which maintained high concentrations localised around the pericycle layer, corroborating similar results in A. thaliana

HMA4 knockout studies obtained by Hussain et al. (2004). Zinc deficiency response genes, including IRT3 and ZIP4 were decreased in the roots of AhHMA4RNAi lines when compared to wild type A. halleri suggesting their expression is secondary or dependent on AhHMA4 expression (Hanikenne et al., 2008). AhHMA4 was shown to have a role in Zn hypertolerance since root growth from AhHMA4RNAi lines grown in the presence of toxic Zn and Cd concentrations demonstrated only 3-15% and 12-37% of the elongation under normal conditions as opposed to WT plants (68 and 60% of normal elongation) (Hanikenne et al., 2008).

A. halleri HMA4 promoter::GUS fusions showed reporter activity was equivalent to the strong constitutive cauliflower mosaic virus (CaMV) 35S promoter expression when transformed into both A. halleri and A. thaliana. When transformed with an A. halleri minigene i.e. AhHMA4cDNA linked to the AhHMA4-1 promoter, A. thaliana demonstrated increased transcript levels of both HMA4 and zinc deficiency response genes, ZIP4 and IRT3 in its roots, and had increased Zn levels in its xylem tissues, when compared with WT A. thaliana (Hanikenne et al., 2008). Similar results were observed when the AhHMA4cDNA construct was expressed in tobacco, thus supporting its role in Zn accumulation in leaf tissue (Barabasz et al., 2010). Enhanced Zn flux from the root symplasm into the xylem vessels and shoot hyperaccumulation as well as upregulation of Zn deficiency response genes in roots, have therefore been demonstrated to be a result of an AhHMA4 tandem triplication in the Zn hyperaccumulator Arabidopsis halleri. It is not known however, whether this genetic mechanism to Zn accumulation is conserved among other Zn hyperaccumulators in the Brassicaceae.

1.10. SUMMARY

The primary functions effecting Zn accumulation are likely to be similar in all hyperaccumulators of the Brassicaceae since all contain a similar collection of genes while physiological differences will probably depend on differences in protein activity and gene expression patterns (Cobbett, 2003). It is most likely that different species and populations have evolved independent means for metal hyperaccumulation, with certain gene duplications often producing variant functions, and leading to higher expression of a particular pathway or conversely

silencing of this pathway (Blanc and Wolfe, 2005). Precisely mapping genes which manipulate such phenotypes in large segregating Brassica populations should permit candidate genes to be inferred from A. thaliana sequences; however, within large crop genomes it has often proved challenging to ascertain if a candidate A. thaliana gene has been similarly preserved (Parkin et al., 2005; Wu et al., 2008; Broadley et al., 2010).

The success of biofortification within Brassica spp. will therefore depend on comparative genomic approaches to extrapolate the molecular mechanisms of Zn accumulation from pre-existing genetic datasets and incorporate this data into crop improvement trials to successfully exploit this natural variation in Zn accumulation.

1.11. AIMS OF THE STUDY

Genetic biofortification, or the exploitation of genes to enhance essential nutrients in plants, through biotechnology and conventional plant breeding, presents an optimistic avenue for sustainable crop improvement with enormous potential to economically and effectively address malnutrition in humans (Bouis et al., 2003; White and Broadley, 2005, 2009; Johns and Eyzaguirre, 2007; Palmgren et al., 2008). It has received substantial scientific consideration (Bouis et al., 2003; Palmgren et al., 2008) and has been promoted by the Consultative Group on International Agriculture Research (CGIAR) and the Harvest Plus program as the primary solution to global micronutrient deficiency (Stein et al., 2007; White and Broadley, 2009). The extent of natural genetic diversity found in Brassica spp. has not yet been observed in other angiosperms and includes an enormous variation in Zn accumulation in its aerial tissues (Wu et al., 2008; White and Broadley, 2009; Broadley et al., 2010). The molecular aspects of $[Zn]_{leaf}$ transport are being more fully elucidated, and through comparative analysis of Zn transporters in related members of the Brassicaceae, such as the model hyperaccumulator N. caerulescens and the model crop Brassica spp., the role of conserved genes in this pathway should become more fully defined (Palmgren et al, 2008; White and Broadley, 2009). Understanding this pathway in the commercially important crop, Brassica spp., should thus provide a basis for the genetic enhancement of numerous Zn deficient food crops in the Brassicaceae and many other important crop families globally.

The primary aim of this study therefore, was to identify through comparative genomic approaches, the level of conservation of genes associating with Zn accumulation in the Brassicaceae, with a primary focus on P_{1B}-type ATPase 4 (HMA4) orthologues in the model Zn hyperaccumulator Noccaea caerulescens and the model crop Brassica spp.

1.11.1. Objectives of research

The investigation was divided into the following main categories.

- To identify and sequence the entire HMA4 locus in the unsequenced Noccaea caerulescens genome and locate syntenic regions with related members of the Brassicaceae (Chapter 3).
- To perform a comparative expression analysis of cis regulatory sequences of HMA4 orthologues from Zn hyperaccumulators N. caerulescens and Arabidopsis halleri and the Zn non-hyperaccumulator A. thaliana (Chapter 4).
- To establish a robust in vitro growth environment for future genetic and physiological trials in N. caerulescens and Arabidopsis thaliana in response to exogenous supplies of Zn (Chapter 5).
- To establish a faster cycling population of Noccaea caerulescens to increase the efficiency and efficacy of future molecular genetic and in planta assays (Chapter 5).
- To identify HMA4 and other putative Zn transporters as candidate genes associating with Zn accumulation in Brassica spp. through comparative genomic analysis of pre-existing Brassica datasets (Chapter 6).
- To identify and efficiently isolate locus specific sequences of Zn transporters from the rapid cycling Brassica rapa R-O-18 genome through a combination of in silico and in vivo approaches (Chapter 6).
- To develop an efficient platform to genotype allelic variation in Zn transporters from a mutant population of B. rapa R-O-18 and therefore complement current sequencing efforts and future reverse genetic approaches (Chapter 6).

CHAPTER 2 GENERAL MATERIALS AND METHODS

2.1. CHEMICALS

Unless otherwise stated, all chemicals were purchased from Sigma-Aldrich (Poole, Dorset, UK). Radiochemicals were obtained from Amersham Pharmacia Biotech Amersham (Little Chalfont, Buckinghamshire, UK).

2.2. PLANT MATERIALS

Plant materials used throughout this thesis included Noccaea caerulescens (J. & C. Presl) F.K.Mey. ecotype Saint Laurent Le Minier seeds, (donated by Guy Delmot from Saint Laurent le Minier in Les Malines region**,** near Ganges, southern France), and Arabidopsis thaliana (L.) Heynh. ecotype Colombia (Col-0) seeds (NASC, University of Nottingham, U.K.). Rapid cycling Brassica rapa mutant TILLing lines (Chapter 6) (purchased from RevGen UK) and wild type lines (donated by Prof. Graham King, Rothamsted Research, U.K.) were derived from Brassica rapa L. ssp. oleifera cv. R-0-18. All plants were generated by either ex vitro or in vitro seed germination. All seeds were stored under darkness at 18 - 20ºC at 30 – 40% relative humidity (RH).

2.3. EX VITRO PLANT CULTURE

2.3.1. Compost mix

Plants were grown in one or all of three compost formulations for all experiments.

(1) Levington M3 high nutrient peat-based compost (pH 5.3 – 5.7), containing nutrients N:280, P:160 and K:350 g m^{-3} + micronutrients (Monro group, Wisbech, Cambridgeshire, U.K.).

(2) Levington M3 mix; containing Levington M3, sand (<1mm) and grit (1-3mm) mix at a ratio of 2:1:1 (v:v:v) respectively (Monro group, Wisbech, Cambridgeshire, U.K.).

(3) Levington M3 mix; containing Levington M3, perlite (2-5 mm) and vermiculite (2-5 mm) mix at a ratio of 2:1:1 (v:v:v) respectively (Monro group, Wisbech, Cambridgeshire, U.K.).

2.3.2. Glasshouse

Plants cultured under glasshouse conditions were subjected to a 16 h photoperiod at 22.3 ±4°C and 13.3 ±2°C mean day and night temperatures respectively. Photoperiods were maintained by light supplementation from 600 watt (W) luminaries using 600 W pressure sodium lamp (Philips® Sun-t Pia green power, Philips Electricals UK Ltd., Guildford, UK), giving a mean photon flux density of 170 µmol quanta m^{-2} s^{-1} within the photosynthetically active radiation (PAR) wavelength. Unless otherwise stated, all plant containers were placed in impermeable growth trays (length 97 cm; width 38 cm; height 5 cm) (Giant Plant Grobag Tray, Sankey, UK) and maintained by basal irrigation, thus preventing aerial humidity and mineral leaching.

Plant pest infestations were prevented through fortnightly applications of biological control agents comprising Swirskiline® (Amblyseius swirskii), Exhibitline® (Steinernema feltiae) and Phytoline® (Phytoseiulus persimilis) to control Bemisia tabaci (Whitefly) and Frankliniella occidentalis (Thrips), Bradysia paupera (Sciarid Fly) and Tetranychus urticae (Red Spider Mite) respectively (Syngenta Bioline, Little Clacton, Essex, UK). Unless otherwise stated, Vitafeed® 2-1-4 nutrient solutions (N-P-K: 16-8-32 + micronutrients) (Vitax Ltd., Coalville, Leicestershire, UK) were applied weekly to plants at a rate of 3 g l^{-1}, eight weeks following germination.

2.4. IN VITRO PLANT CULTURE

Plants grown under in vitro conditions were Noccaea caerulescens and Arabidopsis thaliana.

2.4.1. Seed surface sterilisation

For all plants, seeds were surface sterilised in an axenic flow hood (laminar cabinet, Bassaire, Southampton, UK). For Arabidopsis thaliana, seeds were placed into a sterile 1.5 ml. microcentrifuge tube containing a 1 ml solution of (0.1% v/v) Triton X-100 and milli-Q water (18.2 MΩ cm) (Fisher Scientific, Loughborough, UK) and mixed by vortexing for 30 sec. Solutions were then centrifuged for 5 s in a desk top microcentrifuge, and supernatants aspirated using a sterile 1 ml pipette tip. Seeds were mixed in a 1 ml solution of 50% (v/v)

ethanol (Fisher Scientific, Loughborough, UK) for 30 sec before being rinsed by centrifugation for 5 s in a desk top microcentrifuge, aspiration of supernatants and repeatedly mixing seeds with 1 ml milli-Q water until all traces of ethanol were removed. Seeds were subsequently mixed in a 1 ml solution of 10% (v/v) sodium hypochlorite (NaOCl) (Domestos bleach, Diversey Lever, Northampton, UK) for 5 min before rinsing three times as described previously. Seeds were resuspended in 1 ml milli-Q water (18.2 MΩ cm) and maintained at 5°C until required.

For Noccaea caerulescens, seeds were placed into a sterile 1.5 ml. microcentrifuge tube containing a 1 ml solution of 0.1% (v/v) Triton X-100 and milli-Q water (18.2 MΩ cm) (Fisher Scientific, Loughborough, UK) and mixed by vortexing for 30 sec. Supernatants were aspirated and seeds resuspended and mixed in 1 ml solution of 70% (v/v) ethanol for 2 min before rinsing as for A. thaliana. Subsequently seeds were mixed in a 1 ml solution of 50% (v/v) NaOCl for 15 min before rinsing three times as for A. thaliana. Seeds were resuspended in 1 ml milli-Q water (18.2 MΩ cm) and maintained at 5°C until required.

2.4.2. In vitro growth media and conditions

Unless described otherwise, in vitro growth media contained 8 g l^{-1} agar (A1296, Sigma-Aldrich, Poole, UK) adjusted to pH 5.6 with 0.1 M NaOH, 2.15 g l^{-1} Murashige and Skoog (MS) basal medium (M5524, Sigma-Aldrich, Poole, UK) equivalent to half the strength of the formulations described in Murashige and Skoog (1962) i.e. (half MS), 10 g l^{-1} sucrose (Fisher Scientific, Loughborough, UK), no added growth regulators, and was autoclaved at 121 ºC under 104 kPa for 20 mins. Under axenic flow hood conditions, 75 ml of molten agar was dispensed into translucent polycarbonate boxes (10 x 10 x 10 cm) (Bibby Sterilin, Stone, UK). Surface sterilised seeds were transferred to these using flame sterilised forceps (for N. caerulescens) or pipetted using a 10 µl graduated pipette tip (Star Labs, Milton Keynes, UK) (for Arabidopsis thaliana). Plants were cultured at 20 ±2 °C, under 16h photoperiod, at 50 – 80 µmol photons m^{-2} s^{-1} light intensity from 58 W white halophosphate fluorescent tubes (Cooper Lighting and Security, Doncaster, UK).

2.5. BACTERIAL TECHNIQUES

2.5.1. Escherichia coli strain DH5α

The Escherichia coli strain DH5α was employed for all entry vector cloning during this research. Its full genotype was F^- φ 80dlacZΔM15, Δ(lacZYA-argF)U169 deoR, recA1 endA1 hsdR17(r_k^- m_k^+) phoA, supE44 λ^- thi-1 gyrA96 relA1 (Plant Sciences Division, University of Nottingham, UK). This strain was endA1 thus, did not produce endonuclease I and so permitted cleaner plasmid preps by avoiding non-specific digestion of plasmid DNA. It also lacked the alpha portion of the lacZ gene and so was suitable for blue-white transformation screening.

2.5.2. Agrobacterium tumefaciens strains

2.5.2.1. Agrobacterium tumefaciens 1065

The supervirulent strain 1065 (Curtis et al., 1994; Drake et al., 1997) was derived from strain LBA4404 containing the binary vector pMOG23 (Sijmons et al., 1990) and the hypervirulent pTOK47 which contained the virB, virC and virG operons from pTiBo542 (Jin et al., 1987).

2.5.2.2. Agrobacterium tumefaciens C58

Agrobacterium tumefaciens strain C58 (Wood et al., 2001) containing the pAch5 Ti plasmid to provide in trans and vir functions for the transfer of the integrated T-DNA from a disarmed pBIN19 based binary cloning vector (Hellens et al., 2000).

2.5.2.3. Agrobacterium tumefaciens GV3101

The Agrobacterium tumefaciens strain GV3101 (Koncz and Schell, 1986) was selected to transform Arabidopsis thaliana with promoter::GUS fusion constructs during this research. It carried chromosomal resistance to rifampicin and a Ti-plasmid coding virulence genes and gentamycin resistance. It was sensitive to kanamycin and therefore suitable for transformations with constructs conferring bacterial kanamycin resistance. To maintain selection for this strain, rifampicin and gentamycin were supplemented in growth media at concentrations of 50 µg ml^{-1}.

2.5.3. Transformation vectors

2.5.3.1. Destination vector pMOG23

The destination vector MOG23 contained the chimaeric NOS-NPTII-NOS and CaMV35-GUS-intron cassettes (Vancanneyt et al., 1990; Ohl et al., 2002) inserted between the T-DNA borders conferring kanamycin resistance to plants and was employed to transform both Noccaea caerulescens and Arabidopsis thaliana. This vector was referred to as the 'NPTII::GUS' construct in the experiments described in chapter 5 and donated by Dr. Paul Anthony (University of Nottingham).

2.5.3.2. Destination vector pK7GW1WG2 (II)

The Gateway® destination vector Pk7GW1WG2 (II) contained the attR1 and attR2 recombination sites for Gateway® cloning of fragments from entry vectors harbouring attL1 and attL2 sites mediated by bacteriophage λ (Invitrogen Life Technologies, Paisley, UK.). This vector contained double stranded RNA (hairpin RNA) to trigger post transcriptional gene silencing (Vanhecke and Janitz, 2005; Karimi et al., 2002). During these experiments this formed the basis of the HMA4RNAi construct (Chapter 5) and was donated by Dr. Victoria Mills (University of Nottingham).

2.5.3.3. Destination vector pGWB3

The plant expression vector pGWB3 was employed for the promoter fused with GUS (promoter::GUS) constructs. This vector was based on a modified pABH-Hm1 (Mita et al., 1995), where the HindIII-SacI region in pABH-Hm1, which contained the β-amylase promoter::GUS, was replaced by the Gateway™ cassette and tag (Nakagawa et al., 2007). It contained the NPTII and HPT genes which conferred kanamycin and hygromycin resistance, respectively in transformed plants and bacteria (Nakagawa et al., 2007).

2.5.3.4. Plasmid vector pCR®8/GW/TOPO®

The 2917 bp entry vector contained the attL1 and attL2 sites harbouring bacteriophage λ-derived recombination sequences allowing recombinational cloning of genes of interest in the entry construct with a Gateway® destination vector (Invitrogen Life Technologies, Paisley, UK.). The TOPO® Cloning site

permitted rapid cloning of Taq-amplified PCR products. The pUC origin of replication (ori) permitted high-copy replication and maintenance and the (aadA1) gene conferred spectinomycin resistance to E. coli bacteria (Fig. 2.1).

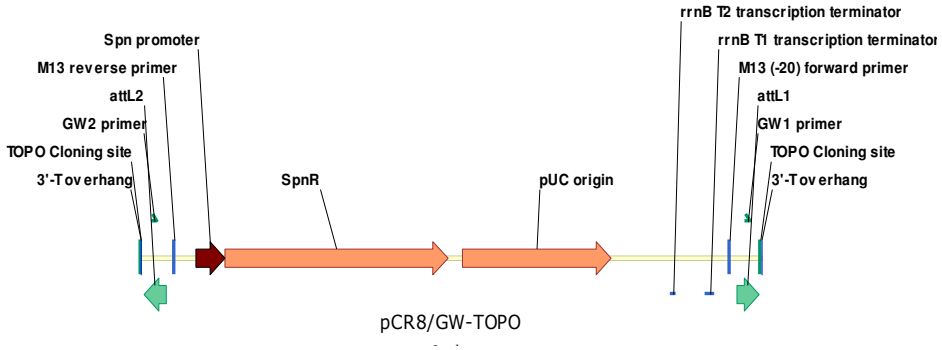


Figure 2.1 Linear illustration of the pCR®8/GW/TOPO® plasmid. Yellow bar represented the entire 2,917 bp plasmid. Green arrows indicated attL1 and attL2 sites for Gateway® recombinational cloning and flanking TOPO® Cloning sites with T-overhangs to facilitate initial cloning of Taq-amplified PCR products. Orange arrows exhibited the pUC origin of replication (ori) and the (aadA1) spectinomycin resistance gene. Maroon arrow indicated the aadA1 promoter. The image was created through Vector NTI 11 (Invitrogen, Paisley, UK).

2.5.4. Growth of bacterial strains

Bacterial strains were grown in Luria-Bertani (LB) broth which was autoclaved prior to appropriate antibiotic supplementation. LB liquid medium contained 10g l^{-1} Bacto-tryptone, 10g l^{-1} Yeast extract and 10g l^{-1} NaCl at pH 7.2 using 0.1 M HCl or 0.1 M NaOH. To make LB agar medium 18g l^{-1} agar was added. E. coli cultures were grown by inoculating antibiotic supplemented LB with a single colony and incubating for 16 h at 37˚C and at 250 rpm in an orbital shaker. Antibiotic supplemented LB broth inoculated with Agrobacterium tumefaciens were grown at 29˚C for 30 hours in darkened conditions.

2.5.5. Storage of bacterial cultures

Bacterial cells were stored for interim periods at 4˚C on 20 ml solidified LB agar in Petri dishes (9 cm diameter). Permanent glycerol stocks of E. coli and A. tumefaciens were established by mixing 1ml of an overnight culture with 500µl of 45% (v/v) glycerol in a sterile cryovial, frozen in liquid nitrogen and stored at -80˚C.

2.5.6. Entry vector transformation into E. coli competent cells

2.5.6.1. Adding adenine (A-Tailing) to blunt-ended PCR products

Amplicons generated using the Phusion® proofreading polymerase were blunt-ended and contained no adenine overhangs which were necessary for the fully efficient pCR8®/GW/TOPO® TA Cloning® system. DNA was adapted by incubation with deoxyadenosine triphosphate (dATP) (A-Tailing). Fragments were gel purified (Section 2.6.3.) and 1 µl was added to a reaction containing 2 µl of GoTaq polymerase buffer (5X), 1 µl of $MgCl_2$ (final concentration of 2.5 mM), dATP (final concentration of 0.2 mM), Taq DNA polymerase (5 units) and SDW to a final volume of 10 µl, incubated at 70°C for 30 min. An aliquot of 1 µl was then employed into the pCR8®/GW/TOPO® TA Cloning® system.

2.5.6.2. Cloning into the pCR8®/GW/TOPO® TA cloning system

A-Tailed (Section 2.5.6.1.) PCR products (1-2 µl) were added to a reaction solution containing 1 µl of salt solution (final concentration of 0.2 M NaCl and 0.01 M $MgCl_2$), 1 µl of pCR8®/GW/TOPO® vector (5 – 10 ng μl^{-1}) containing 3' deoxythymidine (T) residues and SDW to a final volume of 6 µl. By gently mixing the reaction and incubating at room temperature (RT) for 30 min, PCR fragments efficiently cloned into linearised pCR8® vectors by reversing the phosphor-tyrosyl bond introduced by topoisomerase I.

2.5.6.3. Transforming entry vectors into Escherichia coli DH5α

Novel plasmids within the pCR8®/GW/TOPO® cloning reaction solution were added to freshly thawed chemically competent DH5α cells, mixed and incubated on ice for 25 min. Cells were heat shocked for 90 sec at 42˚C before transferring to ice. Room temperature (RT) liquid LB (250 µl) was added prior to incubation at 37˚C, shaking at 250 rpm in an orbital shaker for 1 h to produce an initial growth culture of unselected transformed E. coli DH5α cells.

2.5.6.4. Antibiotic and colony PCR screening of transformed bacteria

Cells were spread (50 µl) on solid LB plates containing appropriate antibiotics (100µg ml^{-1} spectinomycin for pCR8®/GW/TOPO® vectors) and incubated overnight at 37˚C. The largest successful colonies were picked from the plates using sterile pipette tips and dipped into 50 µl of polymerase chain reaction

(PCR) mix i.e. colony PCR (Section 2.7.) to verify the presence of the gene of interest. A PCR reaction was performed using primers specific to the fragment of interest and amplicons separated on agarose gels (Section 2.6.2.). Amplification products of the expected size indicated bacterial colonies containing DNA inserts.

2.5.6.5. Transformation of Agrobacterium tumefaciens GV3101 by electroporation

A single colony of Agrobacterium tumefaciens GV3101 was grown at 29˚C for approximately 30 hours, in 100ml of liquid LB with 50 μg ml^{-1} rifampicin as a selective agent and grown to an optical density $(OD)_{600nm}$ of 0.5 – 0.7. The culture was cooled on ice and centrifuged at 3000 g for 2 mins. The pellet was washed by resuspension in 1 culture volume of cold 10% (v/v) sterile glycerol. Cells were finally resuspended in a 0.01 culture volume of 10% glycerol to give 10^{11} – 10^{12} cells ml^{-1}. Aliquots of 100 μl were pipetted into sterile microcentrifuge tubes, frozen in liquid nitrogen and stored at -80˚C. When ready to transform, cells were thawed on ice and a 40 μl aliquot transferred to a pre-cooled 0.2 cm electroporation cuvette. Plasmid DNA (2 – 10 μg) was mixed with the cell suspension and an electric pulse immediately applied (2.5V, 400Ω, 25 μFD) using a gene pulser (BioRad, Hemel, Hempstead, UK). Cells were then transferred to 1 ml LB and incubated at 29˚C for 3 hours in an orbital shaker at 250 rpm. Aliquots of 100 μl were plated onto solidified LB containing the appropriate antibiotics and incubated for 3 days at 29˚C. Using sterile pipette tips successful colonies were picked from LB plates and colony PCR was performed in a 50 μl PCR mix using fragment specific primers (Section 2.7.). Amplification of products of the expected size indicated bacterial colonies which contained DNA inserts.

2.5.7. Isolation of plasmids from Escherichia coli

A single E. coli colony containing the fragment of interest was grown for 16 h in 10 ml liquid LB supplemented with the appropriate antibiotics, shaking at 250 rpm at 37˚C. According to the manufacturers instructions the plasmid was isolated from these overnight cultures using the Qiaprep Miniprep kit (Qiagen, Crawley, UK). DNA concentrations were established by spectrophotometry

using NanoDrop software at default settings (NanoDrop® ND-1000 UV-Vis Spectrophotometer, Wilmington, DE 19810, USA). A second method of plasmid isolation involved the 'Boil Prep' technique, primarily utilized to quickly verify the presence of a fragment of interest in the E. coli colony, usually through restriction digestion (Section 2.5.7.2).

2.5.7.1. Maxiprep plasmid isolation

Plasmid DNA from bacterial suspension was extracted following the Maxiprep plasmid isolation protocol outlined in Sambrook et al. (1989).

2.5.7.2. Boil prep plasmid isolation

Streaked bacteria were gathered from a solidified LB media and mixed by pipetting in 200 µl STET (10mM Tris-Cl pH8, 0.1M NaCl,1mM EDTA pH8 and 5% v/v Triton X-100) and RNase A (10µg ml^{-1}) in a 1.5 ml microcentrifuge tube. The tube was placed in boiling water for 1 min and then removed and centrifuged at maximum speed in a bench top microcentrifuge for 10 mins. The precipitated pellet was removed using a sterile flattened toothpick and dissolved DNA was precipitated from the liquid by adding 200 µl of Isopropanol (Fisher Scientific, Loughborough, UK). Samples were subjected to -20°C for at least 1 hour and then centrifuged at maximum speed in a microcentrifuge for 10 mins followed by aspiration of the supernatant. Protein and polysaccharide contamination was removed from the pelleted DNA by washing repeatedly with 70% ethanol (Fisher Scientific, Loughborough, UK). The pellet was then dried in a vacuum manifold for 10 min before resuspending the DNA in 10 or 20 µl of sterile distilled H_2O (SDW) and incubating at 5 °C for 1 - 3 hours to fully dissolve.

2.6. NUCLEIC ACID ISOLATION AND MANIPULATION

2.6.1. Plant genomic DNA isolation

Plant genomic DNA was extracted using the GenElute® Plant Genomic DNA Miniprep kit according to the manufacturer's instructions (Sigma-Aldrich, Poole, Dorset, UK). DNA samples were quantified using NanoDrop software (NanoDrop® ND-1000 UV-Vis Spectrophotometer, Wilmington, DE 19810, USA).

2.6.2. Agarose gel electrophoresis of DNA

DNA samples were mixed with 6 X loading buffer and separated by gel electrophoresis at 80 V in 1% (w/v) agarose dissolved in 1 X Tris-acetate-EDTA (TAE) buffer. The gel was supplemented with 5 µl of 10 mg ml^{-1} ethidium bromide. Gels were run in 1 X TAE buffer. TAE buffer was composed of 240 g of Tris base, 57.1 ml of glacial acetic acid, 100 ml of 0.5 M ethylene diamine tetraacetic acid ($C_{10}H_{12}N_2Na_4O_8$) (EDTA) and SDW to a final volume of 1 l.

2.6.3. DNA purification from agarose gels

Gel electrophoresed bands were visualised under a short wavelength UV transilluminator. Bands corresponding to the correct size or molecular weight i.e. containing the correct DNA fragment of interest were excised from the gel using a sterile scalpel, placed into 1.5 ml Eppendorf tubes and flash frozen in liquid nitrogen. DNA extraction was performed using two methods. The first used predominately to obtain highly pure DNA fragments for genomic library probing involved using a MinElute Qiagen Gel Extraction Kit, following the manufacturers protocol (Qiagen, Crawley, UK). The second method, 'Freeze Squeeze', was used primarily to quickly extract DNA for use in cloning.

2.6.3.1. 'Freeze squeeze' gel extraction technique

This protocol was adapted from Islam et al. (2002). An agarose gel was made to a minimum thickness to hold the DNA sample in a single well. Depending on the size of the fragment agarose gel concentrations ranged from 0.7% w/v (fragments > 1 Kbp) to 2% w/v (fragments <100 bp). Immediately following electrophoresis, the gel and casting tray were transferred to a UV light source where the band of interest was excised with a sterile scalpel and placed into a 1.5 ml tube. The slice was then crushed with a spatula and subjected to -80°C for 30 min. The sample was then fully thawed at room temperature and verified by gently flicking the Eppendorf tube to check for the presence of ice particles. Samples were centrifuged at full speed (10,000 x g) on a bench top microcentrifuge for 10 minutes. The supernatant was aspirated using a pipette and transferred to a fresh tube. The gel slice was repeatedly crushed with a pipette tip, re-centrifuged at full speed, and the supernatant transferred to the new tube to recover increased DNA. The pooled supernatants were then centrifuged at

full speed (20,000 x g) for 10 minutes and the supernatant was transferred to a fresh tube. If there was more than 0.4 ml of supernatant, the solution was divided into multiple tubes to a maximum of 0.4 ml per tube. A volume of 5 µl of 10 mg ml^{-1} Dextran (Sigma-Aldrich, Poole, Dorset, UK) and 3 volumes of absolute ethanol were added before mixing and subjecting to -20°C overnight. Samples were then centrifuged at full speed (20,000 x g) for 15 mins. The supernatant was decanted and the tube inverted on a paper towel. 0.5 ml ethanol (70% v/v) was added and the mixture vortexed before centrifuging at full speed (20,000 x g) for five min, decanting the supernatant and inverting the tube on a paper towel. The pellet was dried in a vacuum manifold for 10 min and resuspended in 20 ul of reverse osmosis water.

2.6.4. Endonuclease digestion of DNA

Digestion of DNA using restriction endonuclease enzymes for both DNA PCR fragments of interest and for plasmids was performed with enzymes and corresponding buffers according to manufacturers instructions (Promega, Southampton, UK). For all analytical restriction enzyme digests, reactions were performed in a volume of 20µl on 0.2 - 1.5µg of substrate DNA using a 2- to 10-fold excess of enzyme over DNA, based on unit definition. Unless otherwise stated the following protocol was employed. Under sterile conditions 14 µl SDW was added to a sterile microcentrifuge tube. To this, 2 µl of restriction enzyme 10X buffer, 2ul of 1mg/ml Acetylated BSA, and 0.2 – 1 µg of DNA sample were added. 1 µl of restriction endonuclease (2 – 10 U) was added to a final volume of 20 µl. The reaction was gently mixed by pipetting and centrifuged for 5 sec on a bench top microcentrifuge before incubating for 4 hrs at 37°C. 4µl of Blue/Orange 6X loading dye was added (Promega, Southampton, UK), and the solutions analysed in a 1% (w/v) agarose gel.

2.7. POLYMERASE CHAIN REACTION (PCR) TECHNIQUES

2.7.1. Oligonucleotide primers

Oligonucleotide primers were designed based on gene sequences of interest recorded in the NCBI GenBank database (http://www.ncbi.nlm.nih.gov/) or following contig assemblage. The internet package Primer 3 Version 0.4.0 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) was employed in

designing primers to anneal to specific regions of interest. Unless otherwise stated default settings were employed but with both "Max Self Complementarity" and "Max 3' Self Complementarity" settings were adjusted to 2, to reduce the likelihood of hairpin loops and dimerisation. A BLAST search was performed for each primer against 'Short Nearly Exact Matches' in the NCBI (National Centre for Biotechnology Information) GenBank database (Bethesda, USA) to confirm specificity for the gene of interest. Oligonucleotides were synthesised by MWG Biotech, Ebersberg, Germany (http://www.mwg-biotech.com/html/s_synthesis/s_overview.shtml). Stock solutions were made by resuspension in SDW to a concentration of 100 pmol μl^{-1}. Working solutions were made by dilution with SDW to 10 pmol μl^{-1}.

2.7.2. PCR amplification of DNA using Taq DNA polymerase

For all Taq polymerase PCR reactions, the hotstart technique was employed to reduce the likelihood of amplifying spurious DNA. Go-Taq® polymerase (Promega, Southampton, UK) was employed in reaction volumes containing 10.0 µl of GoTaq buffer (5X), $MgCl_2$ (final concentration (FC) of 2.5 mM), dNTPs (FC 0.2 mM), forward and reverse primers (FC 1 pmol μl^{-1}), DNA (1 - 50 ng), 0.2 µl of Taq polymerase (GoTaq) and SDW to a total volume of 50 µl. PCR conditions required denaturation at 94°C (45 sec), annealing two degrees below the T_m temperature assigned to the primer sequence and extension at 72°C for 1 min per 1 Kbp amplicon.

2.7.3. PCR amplification of DNA using Phusion™ high-fidelity DNA polymerase

Phusion™ high fidelity DNA polymerase (New England Biolabs, Hitchin, Herts, UK), was employed for cloning and sequencing reactions since it was more robust and higher yielding than Taq polymerase, with enhanced extension times of 1 Kbp per 15 sec and low error rates at 4.4 X 10^{-7} in Phusion HF buffer and 9.5 X 10^{-7} in GC buffer. Reaction volumes comprised of 10.0 µl of Phusion HF or GC buffer (5X), dNTPs (FC 0.2 mM), forward and reverse primers (FC 1 pmol μl^{-1}), DNA (1 - 50 ng), 1.5 µl of DMSO, 0.5 µl of Phusion DNA polymerase and SDW to a total volume of 50 µl. PCR conditions required 30 cycles of denaturation at 98°C (5 – 10 sec), annealing two degrees below the T_m

temperature assigned to the primer sequence and extension at 72°C for 15 sec per 1 Kbp amplicon.

2.8. SEQUENCING AND COMPUTATIONAL ANALYSIS OF DNA SEQUENCES

DNA fragments from gel extracted PCR reactions and plasmid extractions from E. coli were initially size resolved to a molecular marker run along side the amplified samples including 100 bp (Promega) and 1 Kbp ladders (Bioline) for appropriate fragment sizes. Once fragments appeared to be the desired size, DNA was gel extracted (Section 2.6.3.) and sequenced.

2.8.1. Dideoxy sequencing of DNA

All PCR fragments and gel extracted plasmid fragments digested with restriction endonucleases were sequenced in both directions using the Value Read, Single Read Service based on Sanger et al. (1974) sequencing methods by Eurofins MWG (Ebersberg, Germany) (http://www.mwg-biotech.com/html/i_custom/i_valueread.shtml). Sequencing was performed using the ABI 3730 XL capillary sequencer with BigDye v.3.1 dye-terminator chemistry as per manufacturer's instructions (Applied Biosystems, Warrington, UK) and typically returned sequencing reads up to 1 Kbp. Each of the four dideoxy terminators was tagged with a different fluorescent dye which fluoresced upon illumination at specific wavelengths and produced a chromatogram from which sequences were deduced.

2.8.2. In silico analysis of DNA sequences

Sequence data were processed using appropriate in silico software for all experiments.

2.8.2.1. Database mining for homologous sequences

To identify sequence orthologues, paralogues and homologues, a Basic Local Alignment Search Tool (BLAST) was employed. Sequences were entered into online searches programs and compared to data deposited in both nucleotide and amino acid sequence databases. High-scoring segment pairs (HSP) contained in both query and database sequences were aligned using a heuristic approach which approximated to the Smith-Waterman algorithm. Depending on requirements either of two algorithms of BLAST were employed: The BLASTn

version 2.2.18 algorithm from GenBank/NCBI (http://www.ncbi.nlm.nih.gov/blast/producttable.shtml#tab31), or the WuBLAST version 2.0 algorithm from the Brassica Genome Sequencing Project (BrGSP) BLAST Server (http://brassica.bbsrc.ac.uk/). The WuBLAST algorithm at default settings was employed to compare interspecific sequences such as A. thaliana with Brassica, since it operated a more flexible sum-of-scores method to identify divergent but similar sequence data. The BLASTn algorithm at default settings operated a less flexible but more rapid Poisson method and so was employed to compare more conserved or closely related sequences

2.8.2.2. Comparative sequence analyses

Two software types were employed to align both nucleotide and amino acid sequences. ClustalW version 1.82 (http://www.ebi.ac.uk/Tools/clustalw2/index.html) was employed to align nucleotide and amino acid sequences at default settings with 'gap open penalty' set to 15.0 and 'gap extension penalty' set to 6.66 for nucleotides and 0.2 for amino acids. AlignX software, VectorNTI version 11 (Invitrogen, Paisley, UK) was similarly employed using identical parameters. Sequence identities were calculated using the Dot Matrix algorithm, VectorNTI version 11 at default settings with the window size set to 5 and stringeny set to 30% (the minimal number of matches in the window to cause a dot to be set in the matrix).

Visual representations of sequence identities, cladograms and phylograms were constructed using programs from the the PHYLogeny Inference Package (Phylip) version 3.66 (http://evolution.genetics.washington.edu/phylip.html). The protein sequence parsimony method (ProtPars) program inferred an unrooted phylogeny from protein sequences. The DNA parsimony program (DNAPars) inferred an unrooted phylogeny from DNA sequences (analogous to Wagner trees).

2.8.2.3. Sequence profiling

All genomic, PCR and plasmid sequence data were characterised and annotated using vector map software from the Vector NTI 11 package (Invitrogen, Paisley, UK). Gene prediction software was employed to detect putative novel genes from de novo sequence data (http://genes.mit.edu/GENSCAN.html). Secondary

structure and thermodynamic profiling of all sequences was performed using the relevant functions within the Vector NTI 11 package. Raw or fragmented sequencing reads from identical loci were assembled into larger overlapping extended contiguous sequences (contigs) using ContigExpress software within Vector NTI 11 package at default settings.

2.9. TRANSFORMATION OF ARABIDOPSIS THALIANA USING FLORAL DIP

Arabidopsis thaliana were grown under glasshouse conditions (Section 2.3.2.) in 0.32 l pots containing Levington M_3 compost for 3 weeks until flowering was initiated. Plants were considered optimal for transformation when there were many immature flower clusters and few fertilized siliques. Agrobacterium tumefaciens strains 1065, C58 or GV3101 containing the construct of interest were grown in liquid LB supplemented with the appropriate selective antibiotic at 29 °C until reaching mid-log phase. A. tumefaciens cultures were then centrifuged at 3000 g for 10 min and the supernatant was removed. The pellet was then resuspended in a sucrose solution (5% w/v) to an OD_{600} = 0.8. Approximately 100 – 200 ml was required for each 0.32 l pot containing A. thaliana plants. Prior to plant exposure to transgenic bacteria, 0.05% (v/v) Silwet L-77 (Lehle Seeds, Round Rock, Texas, USA) was added to the bacterial solutions. All above ground tissue of A. thaliana were submerged and gently agitated in bacterial solutions for 2 – 3 sec (floral dip). Upon removal from solutions, plants appeared to have a film of residual bacterial solution on all tissue. Plants were maintained under high humidity conditions in darkened conditions by placing entire pots into closed plastic sleeves (Zwapak, Alsmeer, Netherlands). This was assumed to optimise bacterial transformation conditions and prevent desiccation as plants began cuticle regeneration, removed by dipping culture surfactants (Clough and Bent, 1998: Dr. Sean May, pers. comm.). Small openings were introduced after 24 hrs, before opening completely after 48 hrs. T_1 seeds were harvested within 3-4 wks.

2.9.1. In vitro selection of T_1 transgenic Arabidopsis thaliana

In vitro growth media for T_1 Arabidopsis thaliana was agar-based containing half MS using the protocol described in (Section 2.4.2). Before decanting 75 ml molten agar to translucent polycarbonate boxes, a solution of ticarcillin

disodium:potassium clavunalate (15:1) solution (Melford Laboratories, UK) was added to a final concentration of 125 mg l^{-1} to eradicate any residual Agrobacterium tumefaciens remaining on seeds from the floral transformations of parent plants. Media was then supplemented with kanamycin sulphate at a final concentration of 50 µg ml^{-1} to select for transgenic plants. All T_1 seeds were surface sterilised (Section 2.4.1), however sterile deionised water (SDW) seed suspensions were decanted onto sterile filter paper in the flow hood and left to dry for 2 – 3 hrs. Once adequately dried, seeds were evenly sprinkled onto agar-based media in each growth box. Box lids were secured using Nescofilm (Bando Chemical Co., Kombe, Japan), before placing in growth rooms at 20 ±2 °C under a 16h photoperiod, and at a light intensity of 50 – 80 µmol photons m^{-2} s^{-1} (Fig. 2.2).

2.9.2. Ex vitro transplantation of T_1 transgenic Arabidopsis thaliana

Following 7 – 10 days of in vitro selection on antibiotic supplemented agar-based media, plantlets which were healthy and actively growing were transplanted to ex vitro conditions into plug trays (2 cm^2 plugs) containing Levington M3 high nutrient peat based compost (Monro Group, Wisbech, Cambridgeshire, UK) and placed in high humidity plant incubators (Multi Tray Pack, Sankey, UK) in a growth room (20 ±2 °C, 16hr photoperiod, 50 – 80 µmol photons m^{-2} s^{-1} light intensity from 58 W white halophosphate fluorescent tubes (Fig. 2.2). Acclimatisation was achieved by gradually removing incubator lids over a 5 day period. Seven day old transplants were transplanted into 0.32 l pots (height 7.9 cm; diameter 9 cm) containing Levington M3. (Monro Group) and grown under glasshouse conditions (Section 2.3.2) to fully mature and self. T_2 seeds were then collected and screened on agar-based media to determine zygosity and number of T-DNA integrations that had taken place within plants.

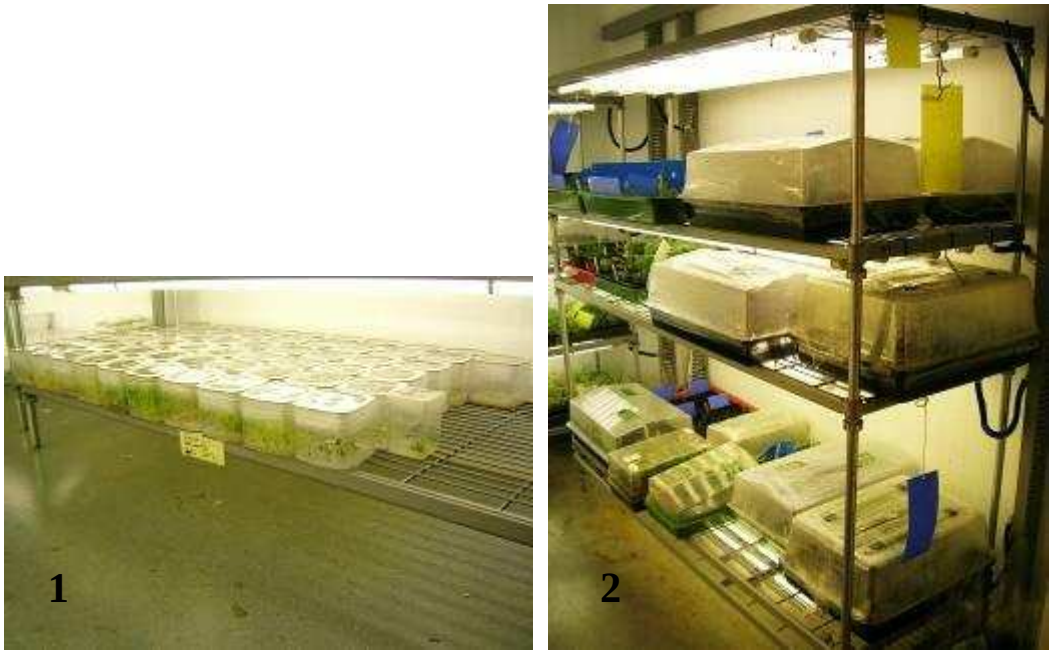


Figure 2.2 (1): Plants growing in vitro under growth room conditions (20 ±2 °C for a 16h photoperiod, at a light intensity of 50 – 80 µmol m^{-2} s^{-1}); (**2):** Ex vitro acclimatisation of plants under similar environmental conditions.

2.10. PRIMERS EMPLOYED

Sequences of primers employed for sequencing and confirmation of both plant and bacterial transformants throughout the experiments were from 5' to 3':

M13_Forward	GTAAAACGACGGCCAG
M13_Reverse	CAGGAAACAGCTATGAC
NPTII_Forward	AATATCACGGGTAGCCAACG
NPTII_Reverse	CGAGGCATGATTGAACAAGA
GUS_Forward	AGTGTACGTATCACCGTTTGTGTGAAC
GUS_Reverse	ATCGCCGCTTTGGACATACCATCCGTA

2.11. DATA ANALYSIS

Unless stated otherwise, data collected from experiments were processed in Microsoft Excel 2003 for statistical analysis. Regression, descriptive statistics and analysis of variance were conducted correspondingly with appropriate parameters in Microsoft Excel 2003. In all analyses, a probability of P = 0.05 or less was taken to be significant.

CHAPTER 3. SEQUENCING NOCCAEA CAERULESCENS HMA4

3.1. INTRODUCTION

Noccaea caerulescens is a model diploid Zn hyperaccumulator plant (Broadley et al., 2007). Studies of mapping populations between different ecotypes of N. caerulescens have made some progress towards identifying candidate quantitative trait loci (QTL) involved in Zn accumulation (Assunção et al., 2006; Deniau et al., 2006; Richau and Schat, 2008). However, one of the main difficulties in resolving QTL to specific candidate genes in these populations lies in the fact that Zn hyperaccumulation is essentially constitutive at the species level in N. caerulescens (Verbruggen et al., 2008). Expression of a N. caerulescens P_{1B}-type ATPase Heavy Metal Associated 4 (NcHMA4) homolog in yeast (Saccharomyces cerevisae) associated with enhanced Zn tolerance and increased Zn transport out of cells which supported a role for Zn efflux across plasma membranes in planta (Papoyan and Kochian, 2004). Studies involving cross species microarrays investigating the expression of the shoot transcriptome have shown that P_{1B}-type ATPases were more highly expressed in the shoots of N. caerulescens than the Zn non-hyperaccumulating relative Thlaspi arvense (Hammond et al., 2006). Further studies characterising HMA4 transcripts in N. caerulescens found that expression increased as exogenous Zn was applied in vitro at levels which were toxic to non-hyperaccumulating species (Hammond et al., 2006). Similarly, higher HMA4 expression levels were also observed in the roots and shoots of N. caerulescens when compared with other closely related non-hyperaccumulator species (Bernard et al., 2004; van de Mortel et al., 2006). In Arabidopsis thaliana the HMA4 ortholog was characterised as a Zn and cadmium (Cd) transporter and localised to the plasma membrane (Mills et al., 2003). Subsequently numerous authors demonstrated a role for AtHMA4 in Zn homeostasis and Cd detoxification as well as the translocation of both metals from the root to the shoot and the localisation of HMA4 expression to the stele (Mills et al., 2003, 2005; Hussain et al., 2004; Verret et al., 2005). Further, overexpression of AtHMA4 in A. thaliana increased both Zn and Cd shoot concentrations (Verret et al., 2004). In Arabidopsis halleri, a close relative to A. thaliana which hyperaccumulates Zn and Cd, QTL involved with Zn and Cd

tolerance were found to co-localize with HMA4 (Courbot et al., 2007). High expression of HMA4 in the BC_1 between the Zn and Cd hyperaccumulator A. halleri, and non-hyperaccumulator, A. lyrata petraea co-segregated with the A. halleri HMA4 allele and with Cd tolerance (Courbot et al., 2007). This evidence was later supported by Hanikenne et al. (2008) who showed through RNA interference (RNAi) that Zn hyperaccumulation and Cd hypertolerance were dependent upon HMA4 expression in A. halleri. These plants were sensitive to increased exogenous Zn and Cd treatments, translocated reduced Zn from the root to the shoot, and were phenotypically more similar to the non-hyperaccumulating relative, A. thaliana (Hanikenne et al., 2008). Subsequent sequencing and functional analyses of AhHMA4 revealed that enhanced HMA4 expression was the result of both gene triplication and altered cis regulation. Despite such overwhelming evidence for its role in Zn hyperaccumulation full sequence data has not been published for HMA4 in N. caerulescens.

The hypothesis that a similar HMA4 gene multiplication event occurred in the Zn and Cd hyperaccumulator Noccaea caerulescens Saint Laurent Le Minier was tested in this chapter. To test this, it was necessary to isolate locus specific genetic sequences offering both coding and non coding data. Thus, a genomic library was first required.

3.1.1. Creating a Noccaea Library

CopyControl™ fosmid pCC1FOS™ vectors were employed to create the Noccaea caerulescens genomic library primarily due to their large insert size and high stability. Unlike high copy number vectors such as cosmids which generally have a high-copy origin of replication and can contain up to 70 copies per cell, fosmids contain the origin and partitioning genes from the F'-episome of E. coli and maintain inserts to single copy (Kim et al., 1992). The use of low copy vectors such as bacterial artificial chromosomes (BACs) and fosmids have been successfully utilised by the scientific community for the preparation of DNA libraries for physical mapping and for large scale DNA sequencing projects (Wang et al., 2009). The primary advantage of these libraries was their stability, and that they contained large, very low or single-copy (SC) clones.

High copy numbers of insert DNA within Escherichia coli hosts which contain repeats or regions with increased AT or GC content have been shown to be susceptible to recombination events (Kim et al., 1992). Moreover, certain genes can express proteins which are toxic to E. coli but under single copy conditions can be tolerated by the host (Kim et al., 1992). Under these conditions, DNA integrity can be maintained and loss of toxic producing sequences is reduced which thus ensures that the maximum of genomic sequence data is represented in the library. During the library production, genomic DNA was sheared, rather than digested with restriction endonucleases, which thus provided a more random distribution of fragments that were not dependent on specific restriction sites (Wild et al., 2002).

However, the major drawback of such low or single copy (SC) clones was that they returned very low levels of DNA, and consequently, reduced purity of DNA with respect to host DNA (Wild et al., 2002). To overcome this issue, CopyControl™ pCC1FOS™ fosmids have two origins of replication, the single copy F-factor ori and a multi copy oriV which can be initiated by the trans-acting replication initiation (trfA) gene product produced by the EPI300™ E. coli cells when subjected to a specific CopyControl™ inducer solution (Epicentre, Madison, USA). This can therefore be used to increase yield for improved fosmid DNA purification. Sequencing such large >40 kb fosmids would be challenging for all traditional sequencing methods including Sanger (Sanger et al., 1997). Therefore more efficient pyrosequencing techniques were selected (Nyrén, 2007).

3.1.2. Pyrosequencing fosmid inserts using the Genome Sequencer FLX 454

Pyrosequencing next generation technologies such as 454 (Roche), Solexa (Illumina) and SOLiD (ABI) have vastly exceeded the capabilities of the traditional Sanger sequencing methods (Sanger et al., 1977), in terms of efficiency and throughput (Meyer et al., 2008). These approaches have now replaced Sanger sequencing in areas which require high throughput, including genome sequencing and serial analysis of gene expression, while also creating new applications such as ultra deep amplicon sequencing (Meyer et al., 2008).

Among these technologies the Next generation Genome Sequencer (GS) FLX 454 offers the greatest read length (350 -450 bp) of all sequencing technologies and is employed for the bulk of all de novo sequencing (Nyrén, 2007, Meyer et al., 2008, Pettersson et al., 2009). GS FLX 454 was therefore used to sequence all fosmids during this study. It can produce > 1 Gbp per day and compares favourably with current Sanger dideoxy sequencing chemistry using the ABI 3730 XL system which produces read lengths of > 1.1 Kbp or 1Mb per day at full capacity (Meyer et al., 2008; Pettersson et al., 2009). The significant difference in data throughput between both systems lies in the different methodologies employed by both sequencing processes. For Sanger sequencing, each read is obtained from an individual sequencing reaction, however for 454 sequencing, reads are obtained from emulsion PCR by amplifying a pool of templates in a single reaction before sequencing. One sequencing reaction on one 16^{th} of a 454 GS FLX sequencing plate can produce approximately 3 Mbp of data. Since this would provide an excessive level of sequence coverage for small sequences such as fosmids, novel approaches such as parallel tagged sequencing (PTS) have been developed which 'barcode' each sample with a specific sequence tag that permits sequences from different genomic origins to be pooled and sequenced together, thus maximising the efficiency of each sequencing output (Meyer et al., 2008). The 454 sequencing reaction is a clonal amplification within an oil aqueous emulsion (emulsion PCR) of randomly sheared DNA fragments of similar lengths, with attached general adaptors to facilitate single primer reactions. Primer coated beads are then added to the emulsion and driven to form a 1:1 bead to fragment ratio. The emulsion is then broken and beads not carrying amplified DNA are rejected through an enrichment process (Pettersson et al., 2009). The remaining beads are distributed on a PicoTiterPlate where beads (28μm in diameter) are fixed to one of 1.6 million wells (44 μm diameter) before sequencing by synthesis (pyrosequencing) is performed where each nucleotide incorporation event is detected as an emitted photon (Pettersson et al., 2009). A general tag primer anneals to the general tags amplified on each bead and emits photons as nucleotides are incorporated. The chemistry surrounding pyrosequencing is thus, polymerase generated nucleotide incorporation releases a diphosphate group (PPi), which forms adenosine triphosphate (ATP) by the use of adenosine phosphosulphate (APS) when

catalysed by ATP sulphurylase. Light is then emitted by a reaction between luciferase and ATP. The enzyme apyrase is employed to degrade any unincorporated dNTPs and ATP to stop the reaction (Nyrén, 2007). In the 454 system luciferase and ATP sulphurylase are fixed to small beads which surround the sequence amplicons. The remainder of reagents is supplied via multiple flows allowing diffusion to templates in the PicoTiterPlate. For each cycle, polymerase and one unique dNTP are added and trigger a light emission in proportion to the number of nucleotide incorporation events. Apyrasae is then used to degrade excess nucleotides before a new cycle begins (Nyrén, 2007; Pettersson et al., 2009). All photons are recorded by a CCD camera. The wash procedure to remove all by-products results in read lengths of 250 bp for the Standard sequencing chemistry, and 350 – 450 bp for the GS FLX Titanium chemistry. An average of 1.2 million wells gives a unique 400 bp sequence thus generating slightly less than 500 Mbp per single run. It has been calculated that a 20 fold coverage of all sequence was sufficient to identify homopolymeric and PCR-introduced errors in sequences (Meyer et al., 2008, Pettersson et al., 2009).

3.2. AIMS

The aim of this chapter was to identify and sequence HMA4 genomic and flanking sequences in Noccaeae caerulescens Saint Laurent Le Minier and test colinearity with orthologous regions in both Arabidopsis thaliana and A. halleri.

3.3. OBJECTIVES

- To create a genomic fosmid library of Noccaea caerulescens Saint Laurent Le Minier.
- To identify fosmid clones containing NcHMA4 from the genomic library.
- To efficiently characterise and sequence fosmid inserts containing NcHMA4 sequences through DNA fingerprinting and sequencing using the Genome Sequencer FLX 454 instrument (GS-FLX)
- To assemble fosmid insert data into larger contiguous sequences and demonstrate colinearity with both A. thaliana and A. halleri orthologous regions.
- Compare genomic, and in silico generated coding and protein sequence data with orthologous sequences from A. thaliana and A. halleri.

3.4. MATERIALS AND METHODS

3.4.1. Southern Blotting.

3.4.1.1. Gel electrophoresis

Approximately 0.2 µg of endonuclease restriction digested plasmid DNA and 10 µg of genomic DNA dissolved in milli-Q water (18.2 MΩ cm) were loaded into a 1 % (w/v) 20 x 20 cm agarose gel supplemented with 5 µl of 10 mg ml^{-1} ethidium bromide and subjected to 20 – 30 volts (V) for an 8 h period. DNA was visualised under UV transillumination to verify the integrity and even loading of extracted samples. The gel was transferred to a glass baking dish and all unused areas were trimmed using a razor blade. DNA was first partially denatured by acid depurination by soaking the gel twice for 15 min in 0.25 M HCL at room temperature (RT), which facilitated the transfer of large DNA fragments. The DNA was then alkali denatured by soaking in several volumes of 1.5 M NaCl and 0.5 M NaOH for 1 h at room temperature with constant stirring and shaking. The gel was then neutralised by soaking in several volumes of a solution containing 1 M Tris Cl (pH 8.0) and 1.5 M NaCl for 0.5 h at RT with constant shaking and stirring. The gel was then ready to be transferred for gel blotting.

3.4.1.2. Gel blotting

The blotting apparatus was constructed using a plastic tub as a reservoir, containing 1L of 10 x SSC (87.65 g NaCl, 44.1 g sodium citrate and H_2O to a volume of 1 L, pH 7). A glass plate was placed over the reservoir and two layers of 3MM Whatman paper (Whatman PLC, Maidstone, UK) were used as a wick, draping over the glass plate, with both ends immersed in the buffer. The gel was inverted onto the wick, with all gel edges masked by Nescofilm to prevent a capillary action short circuit. A precut Hybond-N nylon membrane (GE Healthcare, Buckinghamshire, UK) was carefully laid upon the gel, with all bubbles removed using a glass rod. Four more layers of 3MM Whatman paper were placed on top of the membranes before being covered with a 12 cm high stack of paper towels. A glass plate and 750g weight was placed on top and the apparatus. It was noted that the transfer of large fragments > 15 Kb could take more than 15 hrs therefore the apparatus was left for 20 hrs. Subsequently the construction was dismantled, the well positions marked on the filter in pencil,

and the membrane removed and washed in 2 X SSC. The gel was visualised under UV light to ensure DNA had transferred to the membrane. DNA was fixed to the membrane by placing the membrane between two sheets of 3 MM filter paper and cross-linking with mJ cm^2 of 254 nm UV light in a spectrolinker (Stratagene Stratalinker™ 2400) according to the manufacturer's instructions. The membrane was sealed in a plastic bag and stored at 4˚C for further analysis.

3.4.1.3. Radiolabelling of DNA probes

DNA probes for Sorthern hybridisations were labelled with dCTP α-32P by random priming using a Ready-To-Go DNA Labelling Beads (-dCTP) kit (GE Healthcare, Buckinghamshire, UK) as described by the manufacturer. The probe (50ng) was denatured by boiling for 5 mins followed by quenching on ice to prevent re-annealing. The reaction solution contained 45 µl denatured DNA dissolved in TE buffer (10mM Tris-HCl pH 7.4 and 1mM EDTA pH 8.0), a labelling bead and 5 µl dCTP α-32P (0.4 MBq μl^{-1}) and was incubated at 37˚C for 30 mins. The labelled probe was then separated from unincorporated nucleotides by passing through an Illustra Nick column (GE Healthcare, Buckinghamshire, UK) as described by the manufacturer. The column was rinsed by elution with 400 µl of TE (pH8) and 0.1% SDS. The probe was then denatured for 5 mins at 96 – 100˚C.

3.4.1.4. Hybridisation of radiolabelled probes

Hybond-N nylon membranes were prehybridised for 4 hours at 65˚C in 25 ml prehybridisation solution (25 ml of 20 x SSC, 2 ml 0.5 M EDTA, 5 ml of 100 x Denhardts solution, 5 ml 10% SDS, 1 ml sonicated salmon sperm denatured for 5 mins (100 µg ml^{-1}, Stratagene, UK) and SDW to 100 ml). The radiolabelled probe was denatured, added to the prehybridisation solutions and incubated overnight at 65˚C. Following hybridisation, filters were washed in solutions of 2 x SSC + 0.1% SDS at room temperature for 5 min then at 65˚C for subsequent washes using 2 x SSC + 0.1% SDS for 15 min, 1 x SSC + 0.1% SDS for 10 min and 0.1 x SSC + 0.1% SDS for 10 min until approximately 15 - 30 counts per minute were detected on the membrane using a Geiger-Muller detector. Membranes were prevented from drying by sealing in plastic and then exposed to

autoradigraphy film (Kodak X-Omat AR Film XAR-5, Sigma-Aldrich Gmbh, Steinheim, Germany) in an autoradiography cassette at -80˚C for 4 – 5 days.

3.4.1.5. Film development

Autoradiography film was developed in a dark room using developer and fixer, prepared as directed by the manufacturer (Kodak, Sigma-Aldrich Gmbh, Steinheim, Germany). Film was initially immersed in developer for 5 – 10 mins, rinsed in SDW and transferred to fixer for a further 5 – 10 mins before rinsing thoroughly in SDW. All films were dried at room temperature.

3.4.1.6. Membrane stripping

Membranes which were to be re-used were stripped by immersion in boiling 0.1% SDS for 10 mins. This process was repeated until radioactivity returned to background levels. To verify successful stripping, membranes were exposed to autoradiography film at -80˚C overnight before re-use.

3.4.2. Noccaea caerulescens genomic library

The construction of the N. caerulescens genomic fosmid library was performed by Warwick Plant Genomic Libraries Ltd (WarwickHRI, U.K), from an accession from Saint Laurent Le Minier, Les Malines southern France, and described from sections 3.4.2.1. to 3.4.2.10. A total of 36,864 E-coli EPI300-T1R host cells, each containing one 40kb Noccaea caerulescens genomic insert and 8kb CopyControl™ vector pCC1FOS™ (Epicentre Biotechnologies, Madison, W.I., U.S.A.) were created, resulting in a library giving 5-fold genomic coverage.

3.4.2.1. Noccaea caerulescens genomic DNA shearing

Noccaea caerulescens genomic DNA (0.5 µg μl^{-1}) was sheared into approximately 40-kb fragments in order to create highly random generation of DNA fragments by passing > 2.5 µg through a 200-µl small-bore pipette tip. The DNA was aspirated and expelled from the pipette tip 50-100 times. The extent of DNA shearing was examined by running the samples on a 20-cm long, 1% standard agarose gel at 30-35 V overnight. As a control, 100 ng of the fosmid control DNA was run in an adjacent gel lane. The gel was stained with 5 µl of 10

mg ml^{-1} ethidium bromide after the run to verify that over 10% of the genomic DNA migrated with the fosmid control DNA.

3.4.2.2. End-repair of sheared genomic DNA

This step generated blunt-ended, 5'-phosphorylated DNA. On ice, a reaction was mixed containing 8 µl 10X End-Repair Buffer, final concentrations of 0.25 mM dNTP Mix, 1 mM ATP, sheared DNA (~ 0.5 µg $µl^{-1}$), 4 µl End-Repair Enzyme Mix and SDW to 80 µl and incubated at room temperature for 45 min. Gel loading buffer was then added and the enzyme heat inactivated at 70°C for 10 min.

3.4.2.3. Size selection of the end-repaired DNA

End repaired DNA were analysed on a 20-cm long, 1% (w/v) agarose gel run at 30-35 V overnight. No ethidium bromide was added and the DNA was not exposed to UV light as this was observed to decrease cloning efficiency by > 100 fold. DNA was then fractionated on an agarose gel. A wide comb was used to permit loading of sufficient quantities of DNA into the gel. DNA markers were loaded into the outside lanes of the gel. End repaired DNA was loaded in lanes between both fosmid control DNA lanes. Following overnight gel electrophoresis (30 – 35 V) at room temperature, outer lanes of the gel containing the DNA size markers, the fosmid control DNA, and a small portion of the next lane containing the random sheared end-repaired genomic DNA were excised and subjected to 5 µl of 10 mg ml^{-1} ethidium bromide and visualised under UV light. DNA of the desired size was marked with a divot using a scalpel and used as a reference after reassembly with the portion of the gel not exposed to UV. A gel slice was excised 2 to 4 mm below the position of the fosmid control DNA and the size of the DNA is ≥ 25 Kbp as smaller fragments could give chimeric clones. As an added precaution stained gel portions were rinsed with SDW before reassembly to prevent exposure of DNA to ethidium bromide. The excised gel fragment was transferred to a tared, sterile, screw top 2 ml microcentrifuge tube for extraction.

3.4.2.4. Gel extraction of the size-fractionated DNA

Gel slices were weighed in tared tubes with the assumption that 1 mg of solidified agarose yielded 1 µl of molten agarose. Agarose was melted at 70°C

for 15 min and transferred to 45°C to which prewarmed (45°C) GELase 50X Buffer was added to give a final concentration of 1X. For each 100 µl of melted agarose, 1U (1ul) of GELase Enzyme Preparation was added, the solutions gently mixed and incubated overnight at 45°C. To inactivate the GELase enzyme tubes were transferred to 70°C for 10 min. Aliquots of 500 µl were dispensed into 1.5 ml microfuge tubes and chilled on ice for 5 min. Tubes were centrifuged in a microcentrifuge at maximum speed (>10,000 x g) for 20 minutes to pellet insoluble oligosaccharides. The upper 90 – 95% of the supernatant was removed and aspirated to a sterile 1.5 ml tube leaving behind a gelatinous and translucent to opaque pellet. DNA was precipitated by adding 1/10 volume of 3 M sodium acetate (pH 7.0) and 2.5 volumes of ethanol and mixing by gentle inversion. Precipitation proceeded for 10 min at RT before centrifuging the tube for 20 min in a microcentrifuge, at maximum speed (>10,000 x g). The supernatant was aspirated and the pelleted DNA was washed twice with cold 70% (v/v) ethanol without disruption. The tube was inverted and air dried for 5 – 10 min before resuspension in TE buffer. DNA concentration was estimated by comparing an aliquot with known amounts of the fosmid control standard in an agarose gel. Quantifying DNA by spectrophotometry (A260) was not possible as concentrations were not sufficiently high to be detected.

3.4.2.5. Ligation reaction

It was calculated using the formula N = ln (1-P) / ln (1-f), where P is the desired probability (expressed as a fraction); f is the proportion of the genome (250 Mbp) contained in a single clone; and N is the required number of fosmid clones, that ln (1 – 0.99) / (1 – 0.00016) = 28,780 clones were required at a minimum to ensure that all 40 kb genomic fragments were contained within the library. It was calculated that 0.5 µg CopyControl pCC1FOS Vector = ~ 0.09 pmol vector and 0.25 µg of 40 Kbp genomic DNA = ~ 0.009 pmol; therefore, the ligation reaction contained 1 µl 10X Fast-Link Ligation Buffer, 1 µl 10 mM ATP, 1 µl CopyControl (0.5 µg μl^{-1}), 0.25 µg of ~ 40 kb DNA, 1 µl Fast-Link DNA Ligase and H_2O to 10 µl and was incubated overnight at 16°C. Samples were transferred to 70°C to inactivate the enzyme.

3.4.2.6. Packaging the CopyControl fosmid clones

A solution of 50 ml liquid LB was inoculated with 0.5 ml of a 16 h culture of Escherichia coli EPI300-T1R and grown for ~ 2 hrs to an OD_{600} of 0.8 – 1 before storing at 4°C. One tube of the MaxPlax Lambda Packaging Extracts was thawed, 25 µl was transferred to a fresh microfuge tube on ice and the remaining 25 µl was re-frozen at -70°C. 10 µl of the ligation reaction was added and mixed by pipetting frequently and centrifuged briefly. These packaging reactions were incubated at 30°C for 2 hrs before the remaining 25 µl of MaxPlax Lambda Packaging Extract was added and the reaction incubated for a further 2 hrs at 30°C. Phage Dilution Buffer (PDB) was added to a 1 ml final volume and mixed gently. 25 µl of chloroform was added, mixed and stored at 4°C for infection.

3.4.2.7. Titering the packaged CopyControl fosmid clones

Packaged CopyControl fosmid clones were titered to calculate the number of plates required for the library. Serial dilutions of the 1 ml of packaged phage particles were made using Phage Dilution Buffer (PDB) in sterile microfuge tubes at $1:10^1$, $1:10^2$ and $1:10^3$. From each and from the undiluted phage, 10 µl were individually added to 100 µl of the prepared EPI300-T1R host cells and incubated for 1 hour at 37°C. Infected EPI300-T1R cells were spread on an LB plate + 12.5 µg/ml chloramphenicol and incubated for 16 h at 37°C. Colonies were counted and the titer of the packaged phage particles calculated in colony forming units per ml (cfu ml^{-1}).

3.4.2.8. Plating and selecting the CopyControl fosmid library

10 µl of appropriately diluted packaged phage particles were mixed with 100 µl EPI300-T1R and incubated at 37°C for 1 h. Infected bacteria were spread on solid LB supplemented with 12.5 µg ml^{-1} chloramphenicol and incubated at 37°C for 16 h to select for the CopyControl Fosmid clones. Samples were plated within 24 h as storage for more than 72 hours at 4°C was observed to result in severe losses of phage viability and the plating efficiency. Using a Q-Pix II bench top colony picker by (Genetix, New Milton, Hampshire, UK) colonies were picked into 384 well plates, grown for 16 h at 37°C on horizontal shaker (250 – 300 rpm) and arrayed onto two 22 cm^2 nylon membrane filters using a MicroGrid II, (BioRobotics, Huntingdon, UK). Glycerol stocks were made by

adding sterile glycerol to a final concentration of 20%, mixing and storing at –70°C. A summary of the process is shown in Fig. 3.1.

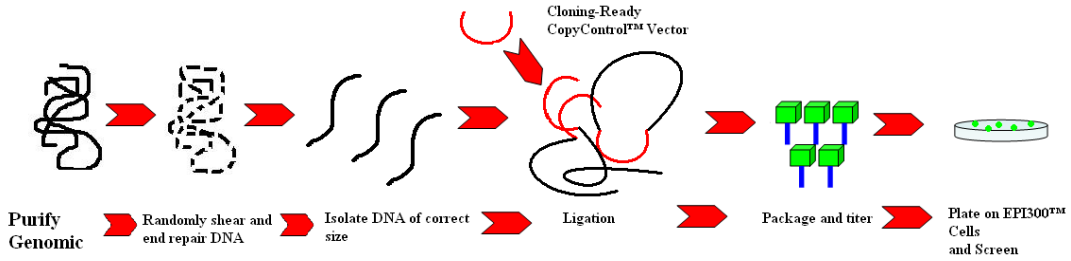


Figure 3.1 Summary of the Noccaea caerulescens genomic library construction.

3.4.2.9. Radiolabelling and hybridisation of library probes

All library probes were labelled as described in (Section 3.4.1.3.). Each pair of library filters were prehybridised for 4 hours at 55˚C in 250 ml prehybridisation solution (50 ml 50 X Denhardts solution, 50 ml 20 X SSC, 25 ml (10%) SDS, 123 ml SDW and 2 ml (10mg ml^{-1}) sonicated salmon sperm (denatured for 5 mins) (Stratagene, UK). The added radiolabelled probe was incubated overnight (55˚C) before washing the filters in solutions of 2 X SSC + 0.1% SDS initially at room temperature, then at 55˚C for subsequent washes until approximately 15 - 30 cpm were detected on the membrane using a Geiger-Muller detector. Filters were sealed in plastic and exposed to autoradigraphy film (Kodak X-Omat AR Film XAR-5, Sigma-Aldrich Gmbh, Steinheim, Germany) at -80˚C for 4 – 5 days.

3.4.2.10. Probing the N. caerulescens genomic library

The N. caerulescens library was arrayed onto two filters divided into 6 compartments, each containing 8 ‘384 well plates’, partitioned into 16 rows (A – P) and 24 (1 – 24) columns, each of which contained a pair of arrayed‘spots’ arranged in a pattern corresponding to their plate of origin (Fig. 3.2). Any positive probe hybridisations therefore, were specifically located to a particular 384 well plate by the arrangement of the ‘spot’ pairs (Fig. 3.2). Filters contained approximately 5% ribosomal and 15% chloroplast contamination.

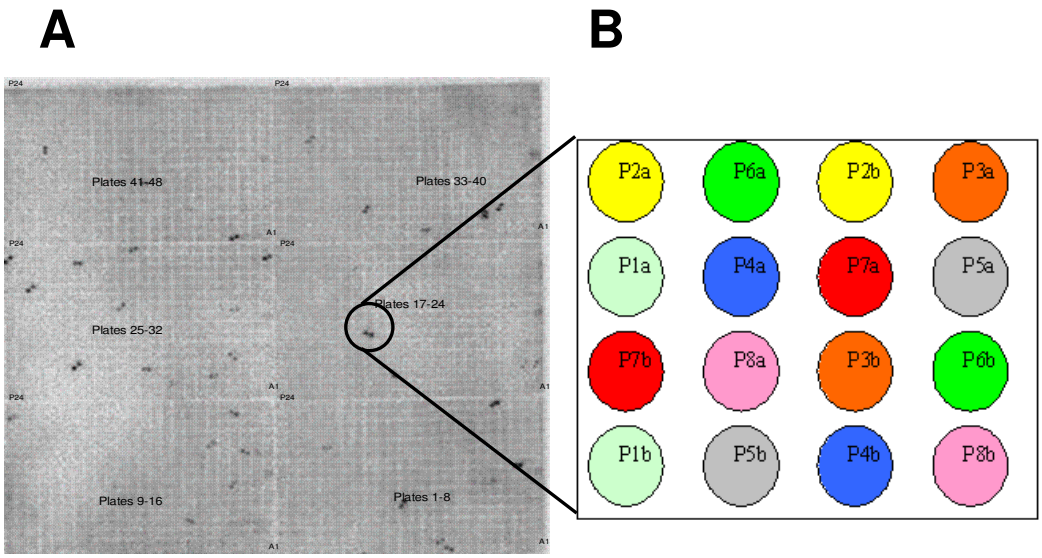


Figure 3.2 Noccaea library filters. **A;** The arrangement of the six compartments each representing a set of 8 plates per filter. Each set of plates was further divided into 384 wells. **B;** A single 'well' containing the arrangement of 'spot' pairs representing their associated plates. 'P' represented '384 well plate'; numbers indicated plate of origin number for that compartment; "a" or "b" indicates 'spot' pairings e.g. "P2a" was the first individual spot on plate 2.

3.4.3. Sequencing N. caerulescens fosmid clones of interest

Pyrosequencing of fosmids was carried out by Cogenics Genome Express (Cambridge, U.K.) using a 454 GS-FLX NextGen sequencing platform (Roche Diagnostics GmbH). Fosmids were extracted using QIAGEN® Large-Construct Kit (Qiagen, Crawley, U.K.) according to manufacturer's recommendations. DNA was individually quantified by optical density (OD) measurement and extracted fosmids were pooled in an equimolar ratio. DNA was transformed into libraries of single-stand template DNA fragments (sstDNA) flanked with amplification and sequencing primer sequences using the GS Library Preparation kit (Roche Diagnostics GmbH), according to the manufacturer's recommendations. 5µg of pooled fosmid DNA was nebulized into fragments ranging from 400 to 800 base pairs. Fragment ends were made blunt before the ligation of the adaptors « A » and « B » allowing for both amplification and sequencing. The ligation mixture was immobilised onto magnetic streptavidin-coated beads, via the biotin moiety of Adaptor B and washed to eliminate any unbound fragments without B adaptors, before melting and purifying the single strand DNA integrating both the A and B adaptor away from the beads. The quality of the sstDNA was verified by a size range analysis with an RNA 6000

Pico-Assay on the 2100 Bioanalyzer (Agilent Technologies, South Queensferry, West Lothian, U.K.) while their quantification was performed by a sensitive fluorescent measurement using the Quant-it™ Ribogreen® RNA assay (Invitrogen Life Technologies, Paisley, U.K.). Subsequently sstDNA was immobilized onto DNA Capture Beads before adding to a mix of amplification solution and oil (GS emPCR Kit I, Roche Diagnostics GmbH), and vigorously shaking on a Tissue lyser (Qiagen, Crawley, U.K.), creating "micro-reactors" containing both amplification mix and a single bead. Emulsion was dispatched in a 96-well plate and the PCR amplification program according to the manufacturer's recommendations was performed. The emulsion was then chemically broken, and Capture Beads were recovered and washed by filtration. Positive beads were purified by the biotinylated primer "A" which bound to streptavidin coated magnetic beads. DNA library beads were separated from magnetic beads by denaturing the double-stranded amplicons leaving a population of bead-bound single-stranded template DNA fragments. The sequencing primer was annealed to the amplified sstDNA prior to evaluating the number of beads carrying amplified sstDNA with a Z2™ Cell Counter (Beckman Coulter, High Wycombe, U.K.). The sample was sequenced in one GS-FLX run using the Small Regions Bead Loading Gasket (Roche Diagnostics GmbH), one 70 X 75mm Pico-Titer plate device (Roche Diagnostics GmbH) and one GS LR-70 sequencing kit (Roche Diagnostics GmbH). According to the manufacturer's instructions, 50,000 DNA beads were loaded per region, followed by the appropriated volume of packing beads and enzyme beads. Following the Pre-Wash run, the sequencing run was launched with the "FullAnalysis" parameter set.

3.4.3.1. Contig alignments of N. caerulescens fosmid sequences

N. caerulescens gene sequence fragments were assembled into contigs by Cogenics Genome Express (Cambridge, U.K.) using a standard whole genome sequencing assembly with the 454/Roche Newbler assembler V 1.1.02.15. Contig sequences corresponding to Noccaea gene regions were then aligned to orthologous 100 Kb regions in the fully sequenced A. thaliana genome via AlignX software using default gap settings. Regions not sequenced by 454 pyrosequencing were sequenced using Sanger sequencing methods. Briefly,

primers were designed to conserved regions in the alignments, located upstream and downstream of unsequenced regions. Following PCR on the fosmid of interest, amplicons were gel purified (Section 2.6.3.) and sequenced (Section 2.8.1.).

3.4.4. Plant genomic DNA extraction

Fresh young leaf tissue was ground to a fine powder in liquid nitrogen using a prechilled mortar and pestle, transferred to a sterile, capped centrifuge tube and mixed with 3 ml of urea extraction buffer (urea 42% (w/v), NaCl (0.31M), Tris-HCl pH 8.0 (50 mM), EDTA pH 8.0 (20mM) and Sodium sarcosine (1% (w/v)) per 1 g ground tissue. Powder and buffer were mixed well by gentle inversion and an equal volume of phenol:chloroform was added. The phases were mixed by inversion, placed at room temperature for 15 minutes, mixed again and spun at 8000 x g for 10 min at 4°C. The aqueous phase was taken, mixed with an equal volume of phenol:chloroform and the centrifugation step repeated. The aqueous phase was then precipitated at -20°C for 30 min following the addition of a sixth volume of 4.4 M ammonium acetate (pH 5.2) and an equal volume of isopropanol. DNA was pelleted by centrifugation in a bench top centrifuge for 3 min at 3000 g. The supernatant was removed and the pellet resuspended in 20 µl SDW and transferred to an microfuge tube. 10µl of a 10 mg ml^{-1} solution of RNaseA was then added and the solution incubated at 37°C for 30 min. An equal volume of phenol:chloroform was then added, the phases gently mixed, and then separated by spinning for 10 min. at 13000 g in a bench top microcentrifuge. The aqueous phase was transferred to a new microfuge tube and an equal volume of chloroform added, after mixing and centrifugation the aqueous phase was taken and DNA precipitated at -20°C for 30 min by the addition of a tenth volume of 3 M sodium acetate and two volumes of ethanol. The DNA was pelleted by centrifugation at 6500 g in a microcentrifuge for 5 min, washed with 70% (v/v) ethanol and dried under vacuum. The pellet was then re-dissolved in 50 µl SDW, 0.5 M NaCl and 10% w/v PEG 8000 and stored at 4°C overnight for DNA precipitation. Following centrifugation at 6500 g for 10 min the supernatant was removed and the pellet washed with 70% ethanol then dried under vacuum. Following resuspension in SDW, genomic DNA was ready for restriction digestion and Southern analysis. Genomic DNA solutions were stored at -20°C.

3.4.5. Primers employed

Sequences of primers employed to identify and characterise fosmid clones were from 5' to 3':

NcHMA4-1_Forward	AAGTCGCCTTGGCCTAGTGG
NcHMA4-1_Reverse	CGGGATGCCCACGGTTAT
NcHMA4-1/4-2_Forward	ATCTCAAGTTGGTCTGCTGTGGA
NcHMA4-1/4-2_Reverse	GGTGATGAAAAGAAAACAGAGGAGAT
NcHMA4-3_Forward	AAGAGCAAATGTGAAAGAGTAATGGG
NcHMA4-3_Reverse	CAAACTCAAACTGGAATAAACACCG
NcHMA4-4_Forward	CATTGTTGCCCTTTTATTTGTTTCA
NcHMA4-4_Reverse	TTCTCCACTTTATTCATCTCCACTTTC
NcHMA4 probe_Forward	GCTAGGGAATGCTTTGGATG
NcHMA4 probe_Reverse	TGTTGCTTCTGCGAGAGAAG
NcHMA4 gene_Forward	GAGTACAACAAAATGTAAGCGAGAAGC
NcHMA4 gene_Reverse	CTTCTCGTTCTTCAACGCCATTAT
NcHMA6_Forward	AACGGGACTCGAGAGATTGA
NcHMA6_Reverse	CTGACGACGGAACAGAGACA
NcHMA7_Forward	GGTGGATGGGATTTTGAATG
NcHMA7_Reverse	TGCAACATAGAATCGCTTGC

3.5. RESULTS

3.5.1. Development of the Noccaea caerulescens 'Ganges' genomic library.

A Noccaea caerulescens 'Ganges' genomic library was constructed by Warwick Plant Genomic Libraries Ltd., University of Warwick, UK, (http://www.wpgl.co.uk/home.htm.) using the low-copy-number fosmid vectors pCC1FOS™. Following extraction and purification, genomic DNA was randomly sheared into 40 kb fragments by passing 2.5 µg (at a concentration of 500 ng/µl) through a 200 µl small bore pipette tip 100 times by aspiration and expulsion. More than 10% of genomic DNA was confirmed as being sheered into 40 kb fragments by running an aliquot (2 µl) of samples with 100 ng of fosmid control DNA on a 20 cm 1% (w/v) agarose gel at 30 volt (V) overnight and visually comparing band intensities.

Blunt-ended, 5'-phosphorylated DNA was then generated for these 40 kb insert fragments using the end repair reaction step (Section 3.4.2.2.). End-repaired fragments were then size resolved on a Low Meting Point (LMP) 20 cm 1% (w/v) agarose gel run at 30 V overnight without staining or exposing to UV light. Fragment lengths > 25 kb were excised and isolated from the gel through the Gelase method. The DNA concentration of the gel extracted N. caerulescens end-repaired 40 kb fragments was estimated by running a DNA aliquot on an agarose gel alongside dilutions of known amounts of the Fosmid Control DNA as a standard. End-repaired N. caerulescens ~40 kb fragments were ligated to CopyControl pCC1FOS vector at a 10:1 (pCC1FOS vector:insert DNA) molar ratio using the Fast-Link system. Using the formula supplied by the manufacturer (Epicentre Biotechnologies, Madison, USA) it was calculated that the N. caerulescens genome (250 Mb) required a minimum of 28,780 clones to reasonably ensure that all 40 kb genomic fragments were contained within the library i.e. N = ln (1-P) / ln (1-f), where P was the desired probability (expressed as a fraction); f was the proportion of the genome contained in a single clone; and N was the required number of fosmid clones; therefore 28,780 = ln (1 – 0.99) / (1 – 0.00016). Based on this information three ligation reactions were performed which would produce an average of $10^{4.5}$ each depending on the insert DNA quality, according to the manufacturer (Epicentre Biotechnologies,

Madison, USA). Ligated products were packaged into MaxPlax Lambda Packaging Extracts following an incubation period of a total of 4 hours at 30°C. Chloroform was then added (25 µl) and the titer of the phage particles determined by making serial dilutions with Phage Dilution Buffer (PDB) and adding 10 µl of each and 10 µl of the undiluted phage, individually, to 100 µl of the prepared Escherichia coli EPI300-T1^{R} strain cells. EPI300-T1^{R} cells were employed due to their Phage T_1-resistance and because of an F factor-based partitioning and single-copy origin of replication system which minimised the accumulation of toxic gene products and selective pressures for mutations. Transformed EPI300™-T1^{R} E. coli cells were resistant to 25 µg ml^{-1} chloramphenicol as a selecting agent. PDB was added to each packaging reaction to obtain a plating density of approximately 200 colonies per plate using the formula supplied by the manufacturer (# of colonies) (dilution factor) (1,000 µl/ml)/ (volume of phage plated [µl]). A ratio of 100 µl EPI300™-T1^{R} cells were incubated with 10 µl of titered phage particles diluted with PDB and incubated for 16 h at 37°C on solid LB medium supplemented with 12.5 µg ml^{-1} chloramphenicol. Infected colonies were picked into 96 x 384-well plates using a Q-Pix II bench top colony picker by (Genetix, New Milton, Hampshire, UK) and grown for 16 h at 37°C on horizontal shaker (250 – 300 rpm). To the solutions, sterile glycerol was added to a final concentration of 20%, mixed and stored at –70°C. All 36,864 individual E. coli clones containing 40 kb genomic Noccaea caerulescens fragments were arrayed onto two 22 cm^2 nylon membrane filters using a MicroGrid II, (BioRobotics, Huntingdon, UK). Quality control provided by Warwick Plant Genomic Libraries Ltd revealed that filters contained approximately 5 % ribosomal and 15 % chloroplast contamination. The specifications of the library therefore were: 2 arrayed filters representing 48 plates of 384 wells each inoculated with ~40Kb of genomic N. caerulescens DNA. This represented a total 1,474,560 Kb and equated to ~5.5 fold coverage of the N. caerulescens (250Mb) genome. In order to induce higher copy numbers of inserts EPI300-T1^{R} cells contained a unique mutant trfA gene which was strictly regulated by an inducible promoter. Once activated through the addition of 500 X CopyControl Fosmid Autoinduction Solution 2 µl ml^{-1} (v/v) (Epicentre Biotechnologies, Madison, USA) to the liquid bacterial growth culture, trfA initiated a high copy oriV origin of replication in the pCC1FOS vector which

increased copy numbers to up to 200 copies per cell which permitted high yields of DNA for all downstream applications.

3.5.2. Screening the Noccaea caerulescens Saint Laurent Le Minier genomic library.

The genomic library was screened for full length genomic Nc-Ganges-HMA4 sequences. A PCR fragment of 421 bp amplified from N. caerulescens genomic DNA was produced, cloned and sequenced using primers designed to the publicly available Noccaea caerulescens Prayon coding sequence (GenBank: AJ567384.1) for use as a library probe.

3.5.2.1. Primer design and amplification of the NcHMA4 probe to hybridise to the genomic library

The internet package Primer 3 Version 0.4.0 (Section 2.7.1.) was employed to design primers specific to flank a 421 bp region of N. caerulescens HMA4 using default settings but with both "Max Self Complementarity" and "Max 3' Self Complementarity" settings adjusted to 2, to reduce the likelihood of hairpin loops and dimerisation. N. caerulescens genomic DNA was extracted from leaf tissue (Section 2.6.1.) and HMA4 specific PCR fragments were amplified using HMA4 probe primers (Section 3.4.5.) and a proofreading polymerase (Phusion, Fisher Scientific, UK) under the conditions described in Section 2.7.3. Fragments were purified by gel electrophoresis and isolated by gel extraction techniques (Section 2.6.3.) for subsequent library hybridisation (Section 3.4.1.4.) (Fig. 3.3). Two positive controls in the PCR reaction used to amplify fragments from N. caerulescens were: "+ve control #1", amplified from using primers specific for NcHMA6 and "+ve control #2", amplified from using primers specific for NcHMA7 (Section 3.4.5.). The negative control "-ve" contained all constituents of the PCR mixture except a DNA template (Fig. 3.3).

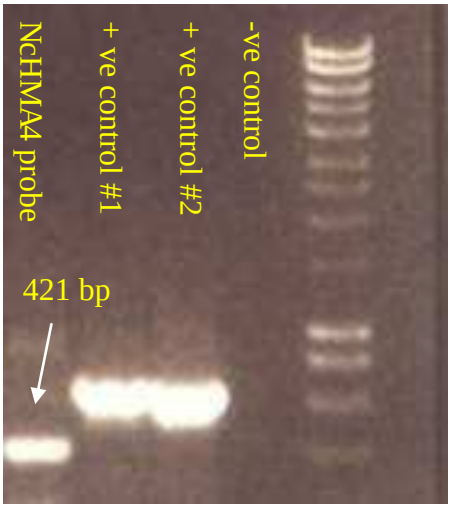


Figure 3.3 PCR amplifying a 421 bp fragment from genomic DNA of Noccaea caerulescens Saint Laurent Le Minier. Lanes labelled '+ve control #' represent amplifications from DNA using primers specific for; 1, HMA6 and 2, HMA7. A no template control was loaded into the lane labelled '-ve control'. Arrow represented bp according to the 1 Kb DNA ladder (Hyperladder I, Bioline). Gel contained 1% (w/v) agarose.

3.5.2.2. Cloning the NcHMA4 probe

Following NcHMA4 probe DNA gel purification and isolation, the 421 bp fragment was incubated for 30 min at 70°C with 1 µl of dATP to add an adenine base to the 3' end of each strand of the PCR product to adapt to cloning into the pCR8®/GW/TOPO® TA Cloning® system (Invitrogen, Paisley, UK) (Section 2.5.6.). The pCR8® compatible NcHMA4 probe fragments (2 µl) were ligated into a pCR8® entry vector as described in Section 2.5.6. Ligated fragments were incubated with freshly thawed competent Escherichia coli DH5α cells (50 µl) and subjected to chemical and heat shock treatments (Section 2.5.6.3.). Solutions were grown overnight at 37°C on solid LB media supplemented with 50 µg ml^{-1} spectinomycin. Spectinomycin resistant colonies were subjected to colony PCR using NcHMA4 primers (Section 2.7.2.) (Fig. 3.4). Control reactions contained all the constituents of the PCR mixture except: the positive "+ve" control template contained genomic N. caerulescens DNA and the negative control "-ve" contained no DNA template (Fig. 3.4). Two colonies producing amplicons of the correct 421 bp size were grown overnight in conical flasks containing 50 ml of Luria-Bertani (LB) broth, supplemented with 50 µg ml^{-1} spectinomycin, at 37°C on a horizontal shaker (250 – 300 rpm). Plasmids were isolated from bacterial cultures using the Qiaprep Miniprep kit (Qiagen, Crawley, UK) and inserts

sequenced using the Value Read, Single Read Service by Sanger sequencing with the ABI 3730 XL capillary sequencer using BigDye v.3.1 dye-terminator chemistry (Eurofins MWG GmbH, Ebersberg, Germany) (Section 2.8.). Insert sequences were aligned using Pairwise Alignment tool at default settings from Clustal W (Vector NTI, Invitrogen, Paisley, UK) to the N. caerulescens Prayon published sequence (GenBank: AJ567384.1) to confirm locus specificity. Glycerol stocks were made of all confirmed transformed E. coli DH5α cells (Section 2.5.).

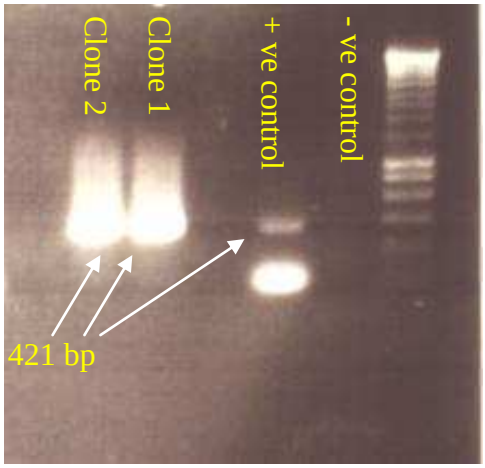


Figure 3.4 Single colony PCR on Escherichia coli DH5α cells (Clone 1 and 2) transformed with the PCR8® entry vector (Invitrogen) containing a 421 bp coding region from NcHMA4. Similar primers were employed to amplify; '+ve control' from genomic N. caerulescens DNA and '-ve control' containing no template. Arrows represented bp according to the 1 Kb DNA ladder (Hyperladder I, Bioline). Gel contained 1% (w/v) agarose.

3.5.2.3. Probing the N. caerulescens genomic library.

Gene specific fragments were PCR amplified from transformed DH5α competent E. coli clones containing the NcHMA4 probe insert (Section 2.5.6.4.), electrophoresis gel purified (Section 2.6.3.) and quantified using nanodrop (Section 2.5.7.). Fragment integrity was confirmed through comparison of gel electrophoresis of 2 µl of purified gel extract of insert DNA and N. caerulescens genomic (positive control) amplicons (Fig. 3.5). The negative control "-ve" contained no DNA template (Fig. 3.5).

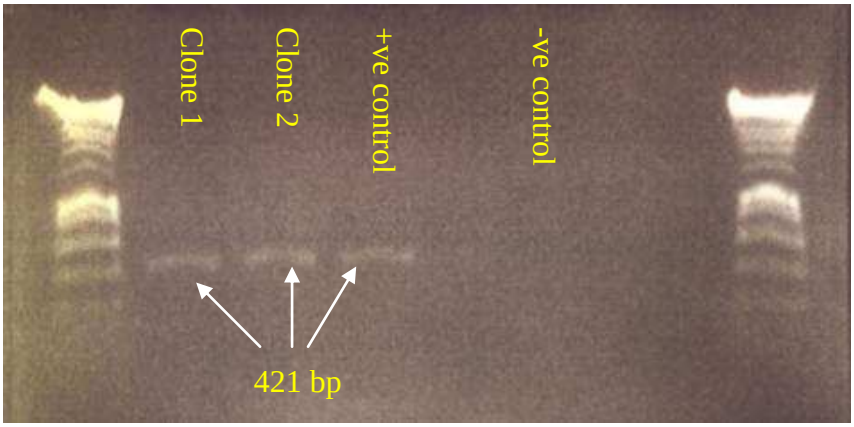


Figure 3.5 Agarose gel electrophoresis of gel extracted PCR amplified products of a 421 bp fragment of NcHMA4 cloned into pCR8® plasmids (Clone1 and 2) or from genomic N. caerulescens DNA (+ve control). A no template control was loaded into the lane labelled '-ve control'. Arrows represented bp according to the 1 Kb DNA ladder (Hyperladder I, Bioline). Gel contained 1% (w/v) agarose.

Both Hybond-N nylon membranes from the N. caerulescens genomic fosmid library were prehybridised for four hours at 65°C in 25 ml of prehybridisation solutions (Section 3.4.1.4.). The NcHMA4 probe fragment was dissolved in 45 µl TE buffer and heat denatured (95°C for 5 min) before adding one labelling bead from a Ready-To-Go DNA Labelling Beads (-dCTP) kit (GE Healthcare, Buckinghamshire, UK) and 5 µl of dCTP α-32P (0.4 MBq μl^{-1}) and incubating at 37˚C for 30 mins. The labelled probe was then purified and added to the prehybridisation solution and incubated overnight at 65°C (Section 3.4.1.4.).

Membranes were washed to reduce radioactivity to between 20 - 30 counts per minute (cpm), then sealed in laminated plastic and exposed to autoradiography film (Kodak X-Omat AR Film XAR-5) in darkened conditions at -80˚C for 4 days (Section 3.4.1.5.). Library bacterial colonies whose insert hybridised to the NcHMA4 probe were identified by darkened 'spots' pairs on the processed autoradiography film (Fig. 3.6). Coordinates for these clones were indicated by their well position highlighted as background marking on the developed films. Within these wells, plates of origin were identified by the arrangement of spot pairs (Section 3.4.2.10.). Bacterial clones were picked from their corresponding plates, streaked onto solid LB supplemented with 12.5 µg ml chloramphenicol and grown overnight at 37°C. A total of 36 clones were identified by probing the library. Seven of these clones generated identical sequences to the NcHMA4 probe sequence using NcHMA4 probe specific primers, confirmed by colony PCR and Sanger sequencing with the ABI 3730 XL capillary sequencer using

BigDye v.3.1 dye-terminator chemistry (Eurofins MWG GmbH, Ebersberg, Germany) (Sections 2.6.1. and 2.8.1.).

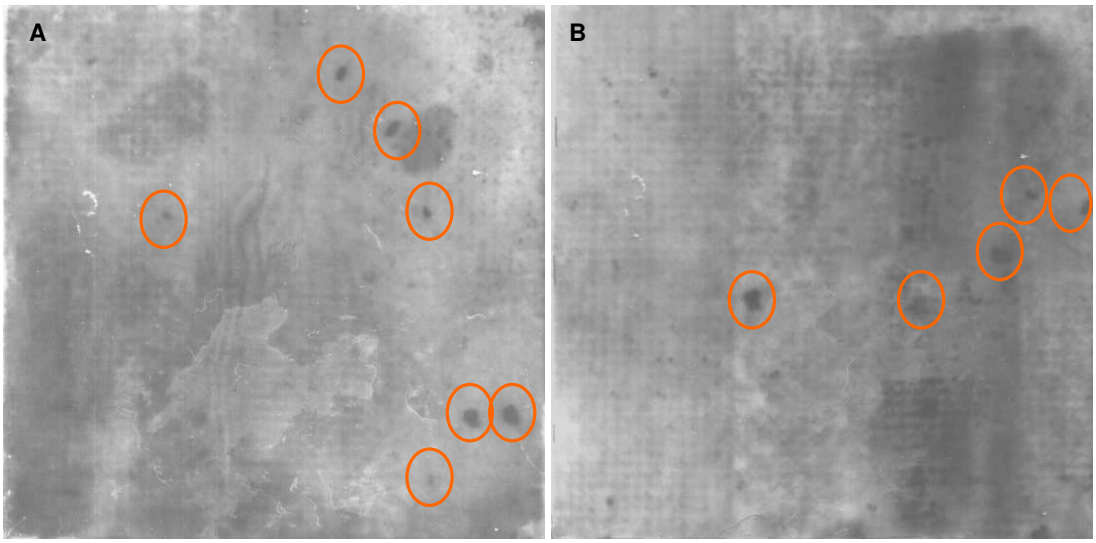


Figure 3.6 Autoradiograph images (Kodak, Sigma-Aldrich Gmbh, Steinheim, Germany) developed after 2 days exposure to a pair (A and B) of Noccaea caerulescens Saint Laurent Le Minier fosmid library filters, containing 36,864 arrayed fragments of approximately 40 kbp, hybridised to the NcHMA4 probe radiolabelled with dCTP α-32P by random priming using a Ready-To-Go DNA Labelling Beads (GE Healthcare, Buckinghamshire, UK). Orange circles surround positive hybridisation events. Each filter measures 22 cm^2.

3.5.3. Sequencing the N. caerulescens HMAs.

Pyrosequencing of a pool of 13 fosmid clones following initial hybridisation to the NcHMA4 probe was carried out by Cogenics Genome Express (Cambridge, U.K.) using the Next Generation Genome Sequencer (GS)-FLX 454 sequencing platform with Standard sequencing chemistry (Roche Diagnostics GmbH) (Section 3.4.3.). Seven clones contained sequences identical to the NcHMA4 probe. All 13 clones were grown overnight in LB media supplemented with 12.5 µg ml chloramphenicol and pCC1FOS™ plasmids were isolated using the Maxi-prep protocol (Section 2.5.7.1.). Extracted plasmids were then examined for integrity through gel electrophoresis by loading 1 µl into a 1% (w/v) agarose gel (Fig. 3.7). Plasmids demonstrated three bands when viewed under UV light which likely represented a possibility of five different plasmid conformations with different electrophoretic mobility including:

1. nicked open-circular DNA where one strand was broken during the lyses processes,

2. relaxed circular DNA where the plasmid was fully intact with both strands uncut, but was enzymatically "relaxed" i.e. supercoils removed,
3. linear DNA which had free ends because both strands were broken,
4. supercoiled or covalently closed-circular DNA which remained fully intact with both strands unbroken, and folding maintained,
5. supercoiled denatured DNA with unpaired regions resulting in a less compact structure caused by excessive alkalinity during plasmid preparation.

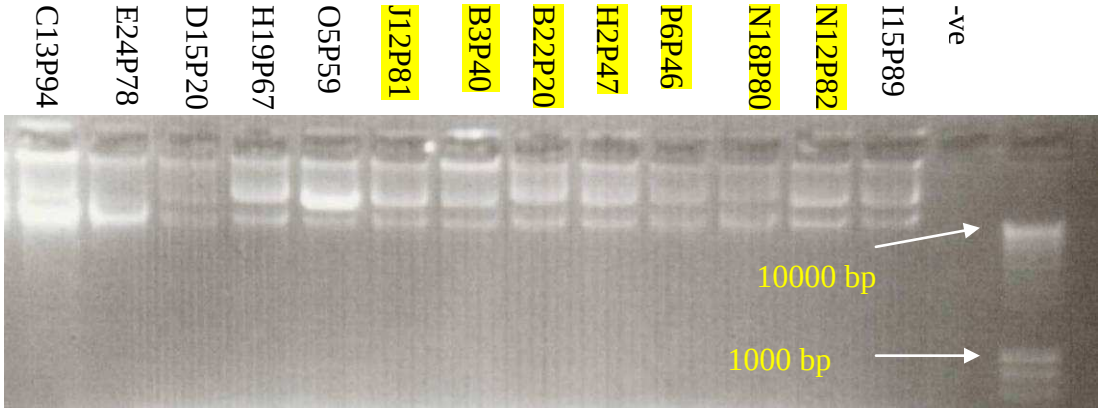


Figure 3.7 Agarose gel electrophoresis of 1 µl of 13 plasmids extracted from fosmid clones from the Noccaea caerulescens library. Lanes are labelled according to fosmid clones. Fosmids highlighted in yellow contain were previously shown to amplify NcHMA4 sequences through PCR. Lane '-ve' contained a no-template control. Arrows represented bp according to the 1 Kb DNA ladder (Hyperladder I, Bioline). Gel contained 1% (w/v) agarose.

Subsequently plasmid DNA was individually quantified by optical density (OD) measurement and extracted fosmids were pooled in an equimolar ratio. Samples were prepared and sequencing performed as described in Section 3.4.3. All 13 fosmids were loaded as one mixture into a 1/16th run in a NextGen GS-FLX 454 sequencing platform with Standard Sequencing chemistry (Roche Diagnostics GmbH) and returned a total of 3 Mbp of sequence, averaging 7-fold coverage per fosmid. These sequences, averaging 100 – 150 bps, were assembled into larger contiguous sequences (contigs) using a standard whole genome sequencing assembly with the 454/Roche Newbler assembler V 1.1.02.15 (Section 3.4.3). All Contigs were aligned to an orthologous HMA4 region in the A. thaliana genome and to the N. caerulescens 'Prayon' published coding sequence (GenBank: AJ567384.1), using ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html)

at default settings (Fig 3.6). Nine contigs containing 528, 438, 628, 2534, 2051, 572, 595, 457 and 1066 bp respectively showed > 80% sequence identity to regions within both genomes. Both 5' and 3' untranslated regions (UTR) in NcHMA4 were sequenced using this method. Four pairs of contigs showed single nucleotide polymorphisms for similar regions within NcHMA4 suggesting allelic variation as was indicated in the alignments of contig00062 and contig00118 which demonstrated 3 SNPs (Fig 3.8).

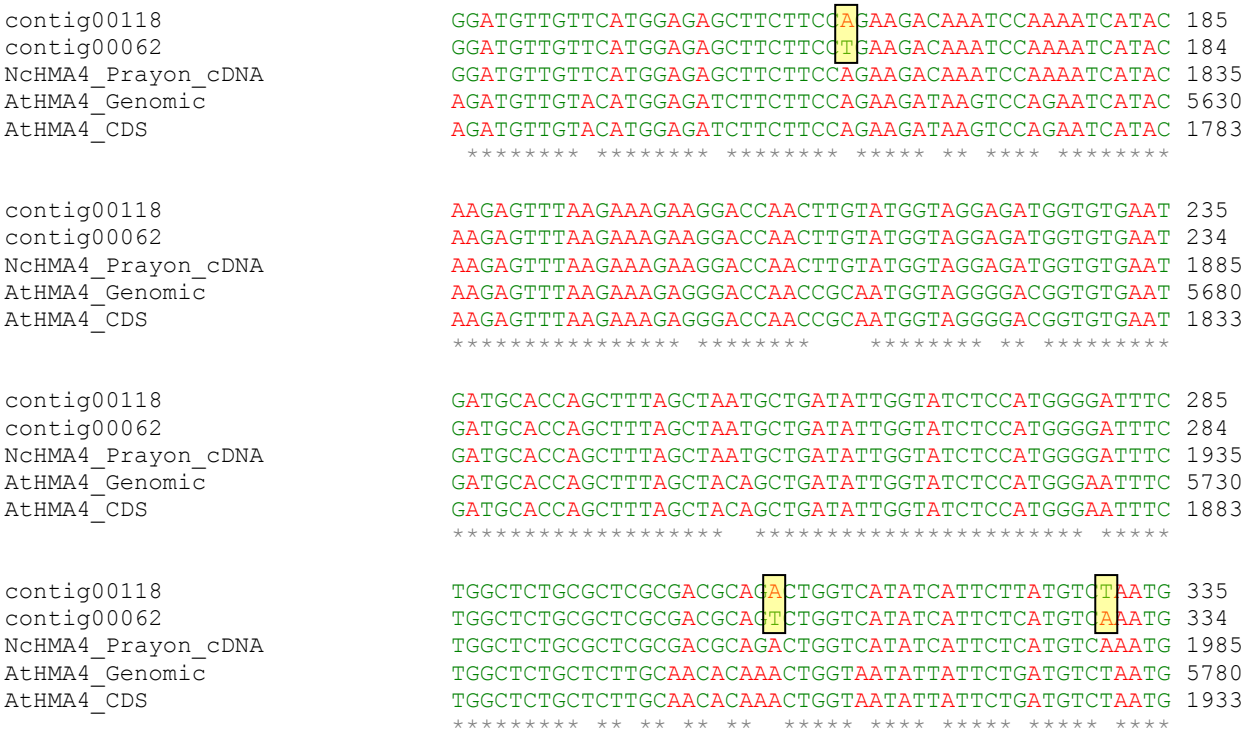


Figure 3.8 Multiple sequence alignment of two contigs representing two Noccaea caerulescens Saint Laurent Le Minier HMA4 paralogues with A. thaliana HMA4 genomic and CDS and one published HMA4 CDS from N. caerulescens 'Prayon' (GenBank: AJ567384.1). Sequences were aligned using ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html). Contig sequences were generated from GS FLX sequences data from a pool of 13 fosmid clones from the Noccaea caerulescens Saint Laurent Le Minier genomic fosmid library. * represent conserved sequences. Number refer to genomic position within each sequence. Boxes highlighted in yellow surround base pair variation between contigs.

Using the alignment data, Noccaea caerulescens primers were designed to sequences demonstrating >95% percentage identity between Arabidopsis thaliana and N. caerulescens contig sequences within the 5' and 3' regions of N. caerulescens HMA4 gene (Section 2.7.1.). All seven fosmid clones containing the NcHMA4 probe sequence were subjected to PCR amplification using these NcHMA4 gene primers and all produced amplicons of between 6.5 and 7.1 kb (Fig 3.9). Control reactions contained all the constituents of the PCR mixture except: the positive "+ve" control template contained genomic N. caerulescens DNA and the negative control "-ve" contained no DNA template. All products were purified by gel electrophoresis and gel extraction (Section 2.6.3.) and approximately 700 bp was sequenced from both ends for each using both forward and reverse NcHMA4 gene primers employed during the PCR amplification and Sanger sequencing with the ABI 3730 XL capillary sequencer using BigDye v.3.1 dye-terminator chemistry (Eurofins MWG GmbH, Ebersberg, Germany). Analysis of sequence chromatograms of these end sequences identified regions with multiple peaks in both P6P46, B22P20 and B3P40 fosmids suggesting the presence of duplicated copies of this region in these clones. Fosmids J12P81, N18P80, H2P47 and N12P82 contained no multiple peaks within the sequence chromatograms. Alignment of sequences from all seven fosmids using Vector NTI 11 at default settings (Invitrogen, Paisley, UK) demonstrated SNPs differences between all fosmids.

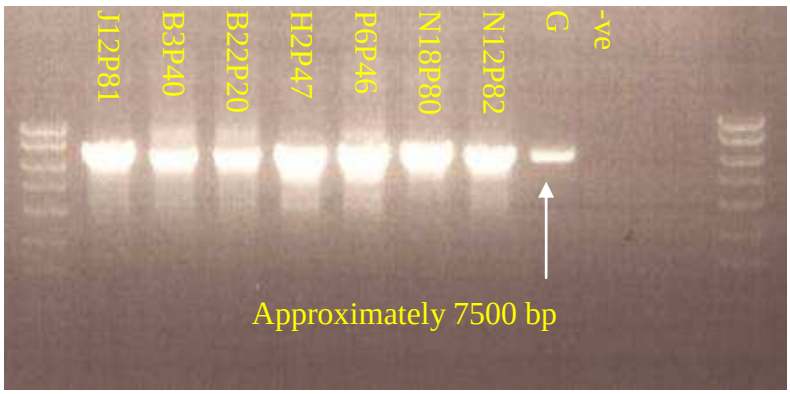


Figure 3.9 Agarose gel electrophoresis of PCR products from fosmid clones which positively hybridised to the radiolabelled NcHMA4 genomic library probe. Primers employed during the PCR were specific for NcHMA4 gene sequences. Lanes were labelled according to fosmid clones containing the N. caerulescens genomic insert DNA. Lane 'G' represents genomic DNA from N. caerulescens Saint Laurent Le Minier. Lane '-ve' contained a no-template control. The molecular ladder was a 1 Kb DNA ladder (Hyperladder I, Bioline). Gel contained 1% (w/v) agarose.

3.5.4. Profiling Noccaea HMA4 locus

In order to efficiently screen and sequence clone inserts, a clone fingerprinting technique was employed to detect clones containing potentially identical insertions. All seven fosmid clones containing NcHMA4 sequences were grown at 37°C for 12 hrs on LB plates supplemented with 12.5 µg ml^{-1} chloramphenicol. Plasmids were isolated through a 'boil prep' technique, where bacterial colonies were resuspended in a STET solution, boiled, precipitated and resuspended in sterile distilled water (SDW) (Section 2.5.7.2.). Genomic DNA extracted from N. caerulescens Saint Laurent Le Minier leaf samples (Section 3.4.4) was used as a positive control to compare all observed hybridisation patterns. Plasmid DNA (2 µl) and N. caerulescens genomic DNA (10 µl) were digested separately with EcoRI, HindIII and BamHI restriction endonucleases for 4 hrs at 37°C (Section 2.6.4.) before running a solution of 10 µl of each digest mixed with 0.1 volume of 10 X loading buffer into a 1 % (w/v) agarose/TAE gel for 12 hrs at 0.5V cm^{-1} (Section 2.6.2.). Gels were blotted onto a precut Hybond-N nylon membrane for a period of 20 hrs. (Section 3.4.1.2.). The membrane was prehybridised for 4 hrs at 65°C in 25 ml of prehybridisation solutions (Section 3.4.1.4.). The NcHMA4 probe fragment was dissolved in 45 µl TE buffer and heat denatured (95°C for 5 min) and radiolabelled with 5 µl of dCTP α-32P (0.4 MBq μl^{-1}) (Section 3.4.1.3.). The labelled probe was then purified and added to the prehybridisation and incubated overnight at 65°C (Section 3.4.1.4.). The membrane was washed to reduce radioactivity to between 20 - 30 counts per minute (cpm), then sealed in laminated plastic and exposed to autoradiography film (Kodak X-Omat AR Film XAR-5) in darkened conditions at -80˚C for 2 days (Section 3.4.1.5.). Autoradiograph images demonstrated two darkened bands where the NcHMA4 probe hybridised to fosmid insert DNA in each lane for B3P40, B22P20 and P6P46 digested with restriction endonucleases (Fig. 3.10). For the remainder of clones J12P81, N18P80 and H2P47 only one hybridisation band was shown and therefore a unique hybridisation event took place. The pattern of bands illustrated a similarity between fosmids B3P40 and B22P20 and to J12P81 (although this contained unique hybridisation bands in each lane). Fosmids N18P80 and H2P47 demonstrated banding patterns that closely resembled those in P6P46 (which contained two bands per lane).

Hybridisation patterns to Noccaea caerulescens genomic DNA digested with identical restriction endonucleases showed similar patterns represented in individual fosmids (Fig. 3.10).

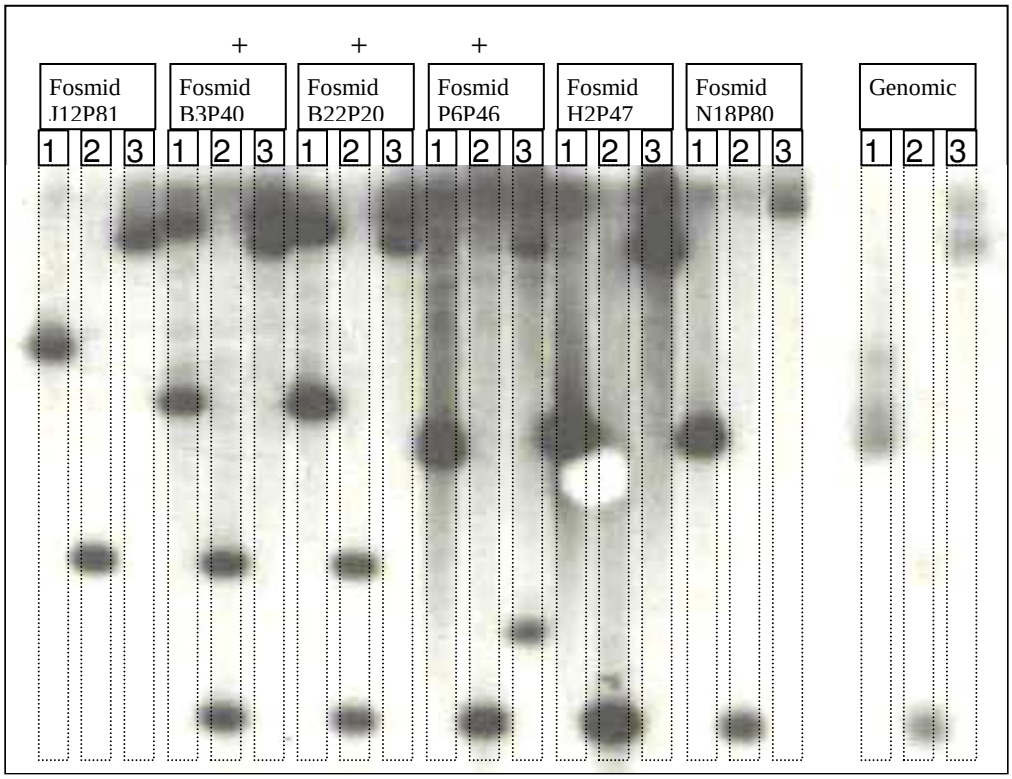


Figure 3.10 Developed autoradiography images of gel blots of N. caerulescens genomic library Fosmid clones and genomic DNA labelled "**Fosmids**" and "**Genomic**" respectively above their corresponding membrane positions. All DNA was digested with EcoRI, HindIII or BamHI corresponding to lanes **1**, **2** or **3** respectively, resolved on a 0.9% (w/v) agarose gel, blotted, and hybridized with the radiolabeled NcHMA4 library probe (represented by darkened regions). Fosmids labelled with '+' contained two hybridisation bands per gel lane.

3.5.5. Sequencing fosmid clone B3P40

Pyrosequencing of selected fosmids was carried out by Eurofins MWG GmbH (Ebersberg, Germany) using the Next Generation Genome Sequencer (GS) FLX 454 sequencing platform using either Standard or Titanium sequencing chemistry (Roche, Basel, Switzerland) (Section 3.4.3). All were contigs were assembled using the 454/Roche Newbler assembler V 1.1.02.15 (Section 3.4.3.). For fosmid clone B3P40 a single colony was isolated and was grown overnight in 500 ml liquid LB supplemented with 12.5 μg ml^{-1} chloramphenicol. The plasmid was subsequently isolated from bacterial cells (Section 2.5.7.1). DNA was quantified and size resolved by gel electrophoresis (Section 2.6.2.).

Purified plasmids were prepared into a shotgun library preparation according to Roche/454 protocols and sequencing was performed using the GS FLX Standard series chemistry (Section 3.4.3.). The total number of reads observed was 4841 representing a total of 1,553,007 bases without GS FLX sequencing keys or tags and with all bad quality bases removed. The average read length without GS FLX sequencing keys, tags and bad quality bases was 320 bp. A total of 4765 sequencing reads were assembled by MWG using the 454/Roche Newbler assembler V 1.1.02.15 (Section 3.4.3.). One large contig was formed containing 36,052 bases with an average sequence coverage of 42.14 fold. The remainder of the 22 smaller contigs ranging from 61 bp to 550 bases in length had an average of between 1 and 6.47 fold sequence coverage and was not employed for further downstream fosmid insert analyses. Fosmid pCC1FOS™ plasmid sequences were identified by the NCBI Basic Local Alignment Search Tool (BLAST) against all nucleotide sequences within the database using the BLASTn algorithm at default parameters (http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastHome). Once plasmid sequences were removed, the 'clipped' large contig was reduced to 27,978 bp (Fig 3.11). Further sequence analysis of this contig revealed the presence of a tandem repetition of N. caerulescens HMA4 genes, orthologous to AtHMA4, labelled from 5' to 3', NcHMA4-1 and NcHMA4-2 respectively. Both tandem repeats shared 94% sequence identity and demonstrated 72% and 73% sequence identity respectively with AtHMA4 at a genomic sequence level using Dot Matrix analysis at default settings (VectorNTI 11, Invitrogen, Paisley, UK). The length of each genomic sequence was 7163 bp and 6833 bp from the deduced translational start and stop sites homologous with the published N. caerulescens 'Prayon' coding sequence (CDS) (GenBank: AJ567384.1). Comparative sequence alignments with N. caerulescens Prayon CDS using AlignX software Vector NTI 11) at default settings and in silico gene prediction software (Genscan) (http://genes.mit.edu/GENSCAN.html) at default settings calculated a total of 10 exons and 9 introns present in both genes copies.

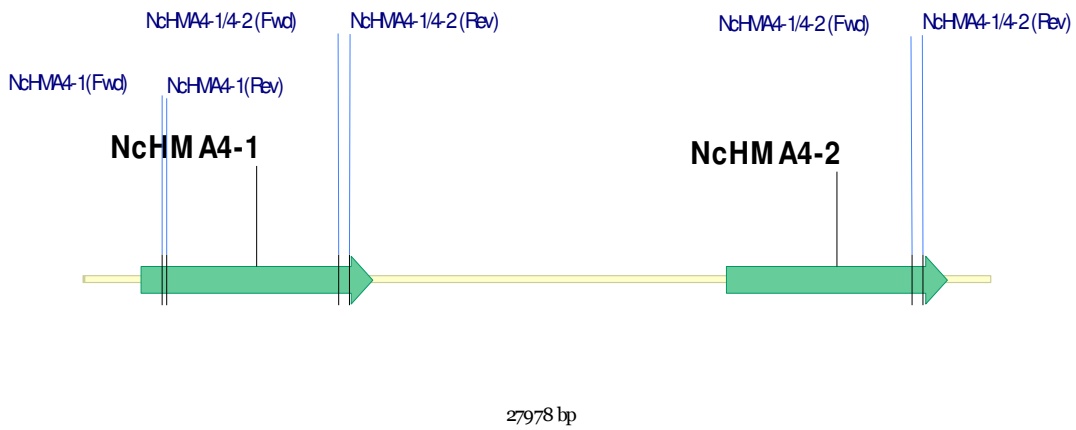


Figure 3.11 Consensus of the genomic illustration of the fosmid B3P40. Yellow bar represents the entire 27978 bp genomic insert. Green arrows illustrate both tandem repeats of NcHMA4-1 and NcHMA4-2 and their transcriptional direction. Blue script and lines highlight sites in the fosmid which were 100% specific for that primer. Image created through Vector NTI 11 (Invitrogen, Paisley, UK).

3.5.6. Sequencing fosmid clone P6P46

Pyrosequencing of fosmids P6P46 was performed similarly by Eurofins MWG GmbH (Ebersberg, Germany) using the Next Generation GS FLX 454 sequencing platform using Titanium sequencing chemistry (Roche, Basel, Switzerland). An isolated P6P46 clone was grown overnight in 500 ml liquid LB supplemented with 12.5 µg ml^{-1} chloramphenicol and plasmid extracted from bacterial cells (Section 2.5.7.1.). DNA was quantified and size resolved by gel electrophoresis (Section 2.6.2.). Shotgun library preparation and GS FLX 454 sequencing and data assembly were performed by Eurofins MWG GmbH, (Ebersberg, Germany) according to Roche/454 protocols (Section 3.4.3.). The total number of reads observed was 3703. The average read length without GS FLX sequencing keys, tags and bad quality bases was between 350 – 450 bp and sequence coverage averaged 20 fold. Closure of unsequenced gaps between contigs was performed by classical Sanger approaches using region specific primers and sequencing of derived PCR products on the ABI 3730xl capillary sequencers using BigDye v.3.1 dye-terminator chemistry. Sanger sequence data assembly was performed using the gap4 module of the Staden package (http://helix.nih.gov/Documentation/staden/manual/gap4_toc.html).

The completely assembled contig contained 39,785 bp and following BLASTn analysis and removal of pCC1FOS™ plasmid sequences, the insert length was reduced to 31,521 bp (Fig 3.12). Further sequence analysis of this contig

revealed the presence of a tandem repetition of N. caerulescens HMA4 genes, orthologous to AtHMA4, labelled from 5' to 3', NcHMA4-3 and NcHMA4-4 respectively. Both tandem repeats shared 99% sequence identity and showed sequence identity of 99% with NcHMA4-1, 93% with NcHMA4-2 and 72% with AtHMA4 using Dot Matrix analysis at default settings (VectorNTI 11, Invitrogen, Paisley, UK). The length of each genomic sequence was 6498 bp and 6488 bp for NcHMA4-3 and NcHMA4-4 from the deduced translational start and stop sites homologous with the published N. caerulescens 'Prayon' coding sequence (CDS) (GenBank: AJ567384.1). Similar to NcHMA4-1 and NcHMA4-2, NcHMA4-3 was predicted to contain a total of 10 exons and 9 introns present. NcHMA4-4 contained a predicted truncation at bp 1980 or amino acid 660 and so contained 9 exons and 8 introns were calculated to be present.

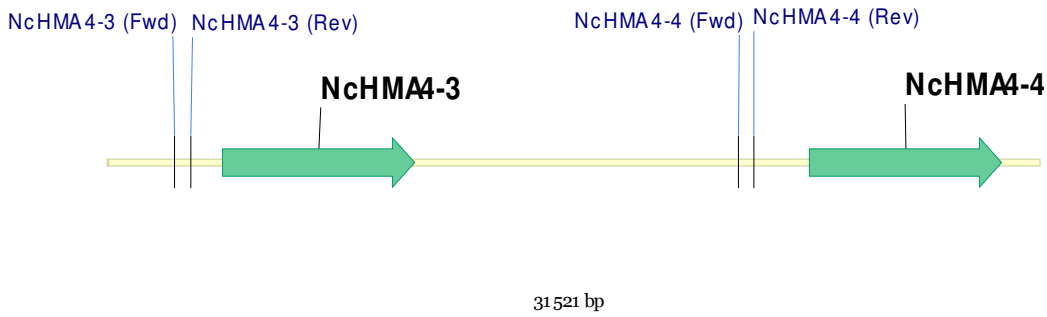


Figure 3.12 Consensus of the genomic illustration of the fosmid P6P46. Yellow bar represents the entire 31521 bp genomic insert. Green arrows illustrate both tandem repeats of NcHMA4-3 and NcHMA4-4 and their transcriptional direction. Blue script and lines highlight sites in the fosmid which were 100% specific for that primer. Image created through Vector NTI 11 (Invitrogen, Paisley, UK).

3.5.7. Sequencing fosmid clone J12P81

Pyrosequencing of fosmid J12P81 using GS FLX Titanium chemistry resulted in 11,030 reads giving an average of 20 fold coverage of one assembled 39,409 bp contig using the 454/Roche Newbler assembler V 1.1.02.15 (Section 3.4.3.) which was reduced to 31,521 bp after BLASTn analysis and removal of pCC1FOS™ plasmid sequences (Fig 3.13). Sequence analysis identified a sequence of 6833 bp which had 72% sequence identity with AtHMA4 and was identical to NcHMA4-2 located in the fosmid B3P40. Sequences flanking the 5' region of this gene showed 100% identity with similar regions upstream of NcHMA4-2 on fosmid B3P40 DotMatrix analysis at default settings, indicating a

common region shared by both J12P81 and B3P40. Two gene sequences were identified downstream of the 3' region of this sequence and demonstrated 75% sequence identity with A. thaliana genes At2g19160 and At2g19170.

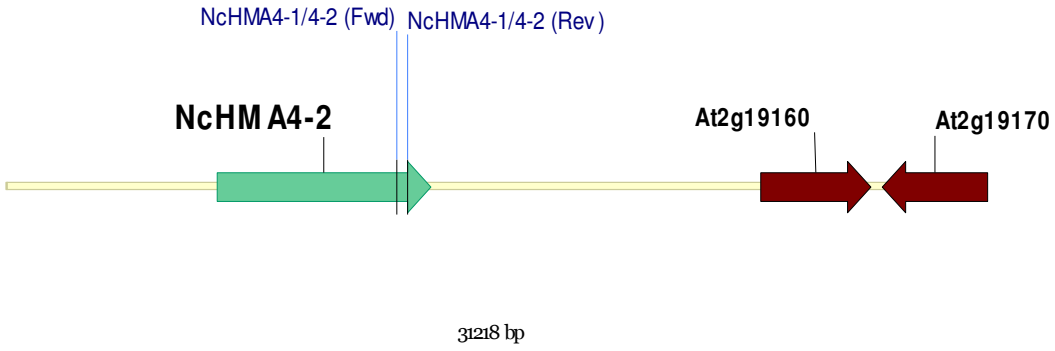


Figure 3.13 Consensus of the genomic illustration of the fosmid J12P81. Yellow bar represents the entire 31218 bp genomic insert. Green arrow illustrates a single copy of NcHMA4-2 its transcriptional direction. Brown arrows illustrate flanking genes At2g19160 and At2g19170 and their transcriptional directions. Flanking genes are labelled according to their A. thaliana orthologues. Blue script and lines highlight sites in the fosmid which were 100% specific for that primer. Image created through Vector NTI 11 (Invitrogen, Paisley, UK).

3.5.8. Sequencing fosmid clone N18P80

GS FLX Titanium sequencing technology and PCR product sequencing using Sanger BigDye v.3.1 dye-terminator chemistry gave 20 fold coverage of the fosmid N18P80 which was assembled into a 20,090 bp genomic insert (Section 3.4.3.). A 940 bp sequence at the 3' end of this contig was identical to the 5' region of NcHMA4-3 (Fig 3.14). A total length of 3853 bp upstream of this sequence showed 99% sequence identity with a similar region upstream of NcHMA4-3 on contig P6P46, indicating a common region shared by both N18P80 and P6P46. Upstream of this NcHMA4-3 sequence, four N. caerulescens orthologues to At2g19060, At2g19070, At2g19080 and At2g19090 showed 100% synteny to the upstream 5' region of the AtHMA4 locus (Fig. 3.14)

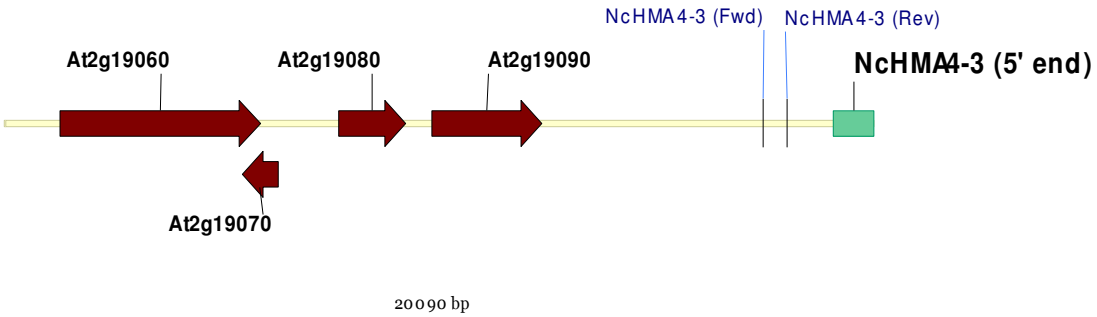


Figure 3.14 Consensus of the genomic illustration of the fosmid N18P80. Yellow bar represents the entire 20090 bp genomic insert. Green box illustrates a single copy of the 5' end of NcHMA4-3. Brown arrows illustrate flanking genes At2g19060, At2g19070, At2g19080 and At2g19090 and their transcriptional directions. Flanking genes are labelled according to their A. thaliana orthologues. Blue script and lines highlight sites in the fosmid which were 100% specific for that primer. Image created through Vector NTI 11 (Invitrogen, Paisley, UK).

3.5.9. Characterising the Noccaea caerulescens HMA4 locus

To test if these NcHMA4 containing fosmid inserts were located in a single HMA4 locus within the Noccaea genome, B3P40, P6P46, H2P47, N12P82 and N18P80 were subjected to PCR using primers specific to NcHMA4-1, NcHMA4-1 and NcHMA4-2, NcHMA4-3 and NcHMA4-4 (Section 3.4.5.). As a positive control genomic DNA from N. caerulescens Saint Laurent Le Minier was similarly amplified using these primers (Fig. 3.15). The N. caerulescens genomic DNA positive control was amplified by all primer sequences (Fig. 3.15). B3P40 was uniquely amplified by primers specific for NcHMA4-1 and NcHMA4-2. P6P46 and N12P82 were very weakly amplified by primers specific for NcHMA4-1 and NcHMA4-2 however amplification was significantly more pronounced for primers specific for NcHMA4-3 and NcHMA4-4. N18P80 was uniquely amplified by primers specific for NcHMA4-3. H2P47 was uniquely amplified by primers specific for NcHMA4-1 and NcHMA4-4, indicating possible shared genomic regions between P6P46 and B3P40, and was therefore selected for further GS FLX 454 sequencing.

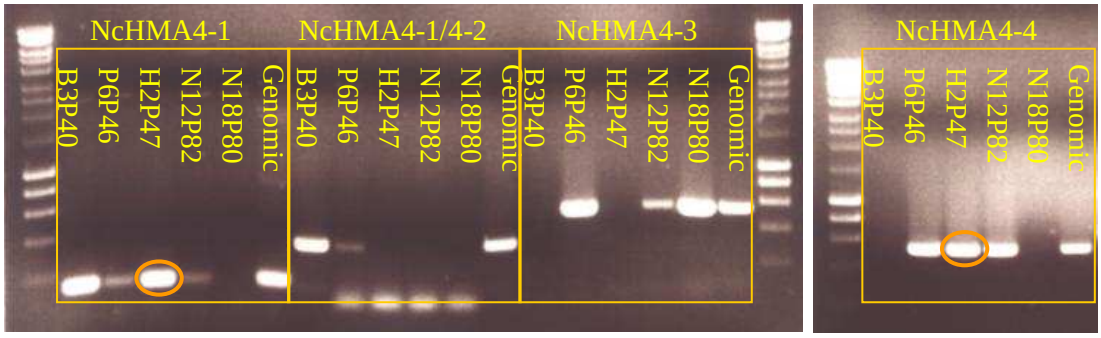


Figure 3.15 Agarose gel electrophoresis of PCR products from fosmid clones containing N. caerulescens HMA4 sequences and Noccaea genomic DNA using primers specific for 1) NcHMA4-1, 2) NcHMA4-1 and 4-2, 3) NcHMA4-3 and 4) NcHMA4-4. Orange circles surround PCR amplicons for fosmid H2P47. Lanes are labelled according to fosmid clones or 'Genomic' Noccaea DNA. The molecular ladder is a 1 Kb DNA ladder (Hyperladder I, Bioline). Gel contained 1% (w/v) agarose.

3.5.10. Sequencing fosmid clone H2P47

Next Generation GS FLX Titanium sequencing technology and Sanger BigDye v.3.1 dye-terminator chemistry for sequencing PCR products generated an average of 20 fold coverage of the fosmid H2P47. It was shown to contain a 20,916 bp genomic insert (Fig. 3.16). Sequence analysis identified a sequence of 6488 bp which was identical to NcHMA4-4 and shared 99% sequence identity with NcHMA4-3 and NcHMA4-1, 93% identity with NcHMA4-2 and 72% with AtHMA4 (Dot Matrix analysis at default settings) (VectorNTI 11, Invitrogen, Paisley, UK). Sequences flanking the 5' region of this gene showed 100% identity with similar regions upstream of NcHMA4-4 on fosmid P6P46 indicating a shared genomic region between P6P46 and H2P47. To the 3' end H2P47 contained a 425 bp sequence which shared 100% sequence identity with the 5' end of fosmid insert B3P40 and indicated a commonly shared genomic region between H2P47 and B3P40.

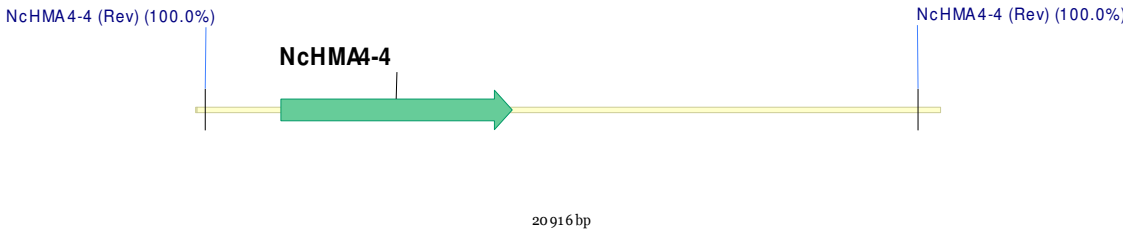


Figure 3.16 Consensus of the genomic illustration of the fosmid H2P47. Yellow bar represents the entire 20916 bp genomic insert. Green arrow illustrates a single copy of NcHMA4-4 its transcriptional direction. Blue script and lines highlight sites in the fosmid which were 100% specific for that primer. Image created through Vector NTI 11 (Invitrogen, Paisley, UK).

3.5.11. Mapping the HMA4 locus in Noccaea caerulescens

Contigs representing fosmid inserts were aligned to a 207,564 bp Arabidopsis thaliana genomic sequence (100 kb flanking either side of AtHMA4) using AlignX software at default settings (Vector NTI 11). Fosmid inserts demonstrating > 99% sequence identity were judged to contain identical segments of the Noccaea caerulescens genomic regions and therefore aligned to each other until a scaffold of 5 fosmid inserts from 5' to 3', N18P80, P6P46, H2P47, B3P40 and J12P81, was formed (Fig. 3.17). This assembly of contiguous sequences represented a 101,480 bp region in the N. caerulescens Saint Laurent Le Minier genome. Gene sequences flanking the 5' end of NcHMA4-3 were 100% syntenic with the orthologous region in Arabidopsis thaliana and A. halleri (Hanikenne et al., 2008). Gene sequences flanking the 3' end of NcHMA4-2, At2g19160 and At2g19170, were 100% syntenic with the equivalent orthologous regions in both A. thaliana and A. halleri. Closer to the 3' region of NcHMA4-2 however N. caerulescens did not contain At2g19120, At2g19130 and At2g19150 orthologues which were present in both Arabidopsis genomes and At2g19140 which was present in just A. thaliana.

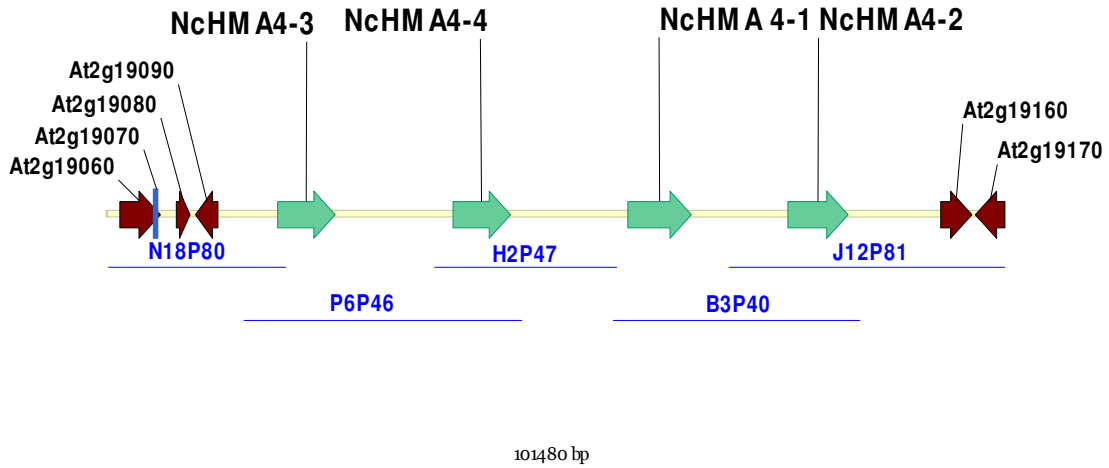
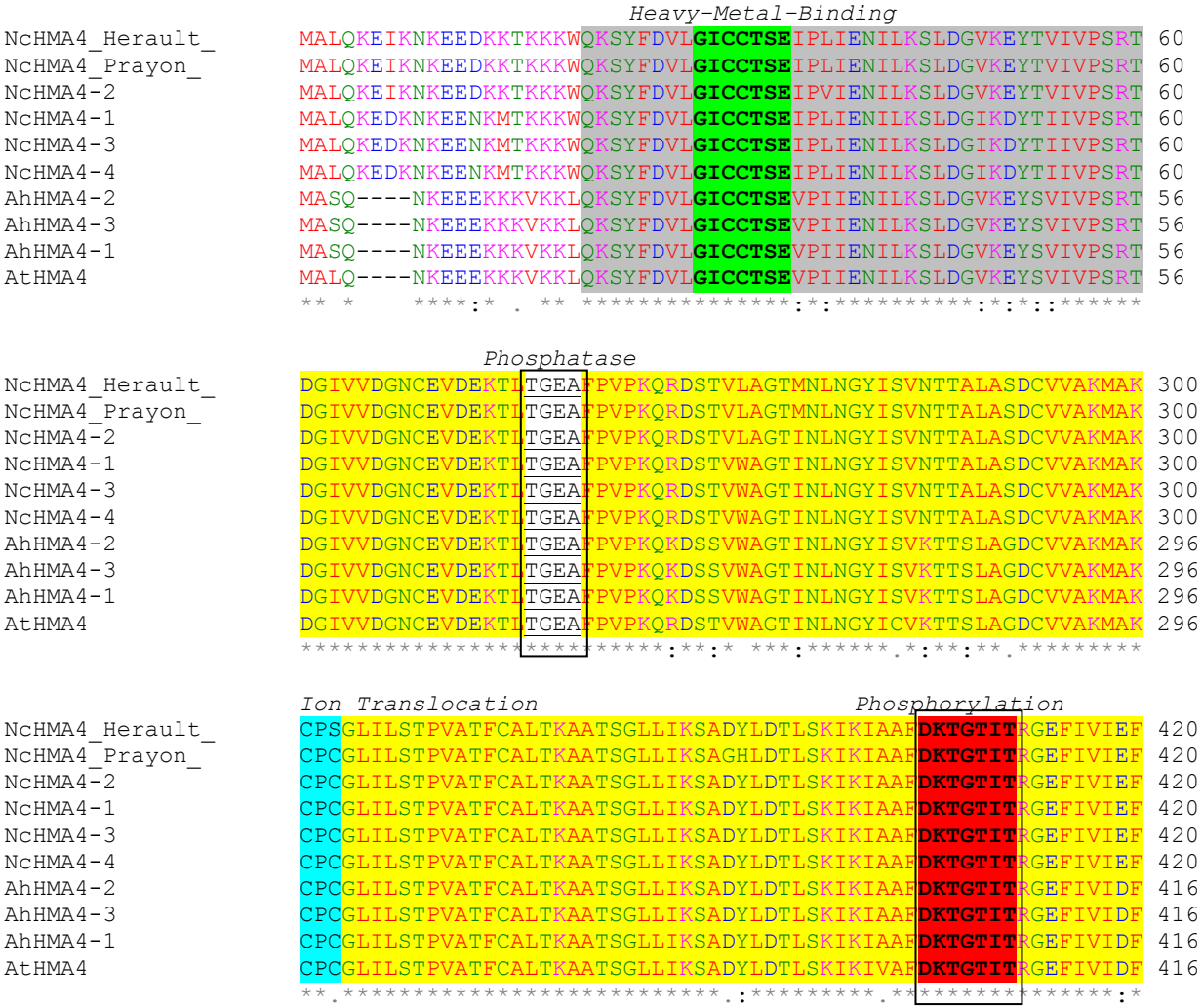


Figure 3.17 Consensus of the genomic illustration of the Noccaea caerulescens Saint Laurent Le Minier HMA4 locus. Yellow bar represents the entire 101,480 bp genomic sequence. Green arrows illustrate all four tandem repeats of NcHMA4 (NcHMA4-1, NcHMA4-2, NcHMA4-3 and NcHMA4-4) and their transcriptional direction. Brown arrows illustrate the transcriptional directions of genes flanking the NcHMA4 quadruplication; from the 5' end, At2g19060, At2g19070, At2g19080 and At2g19090 and from the 3' end, At2g19160 and At2g19170. Flanking genes are labelled according to their A. thaliana orthologues. Blue script and lines represent regions contained in fosmid clones from the N. caerulescens genomic fosmid library. Image created through Vector NTI 11 (Invitrogen, Paisley, UK).

3.5.12. Characterising Noccaea caerulescens HMA4 paralogues

Amino acid sequences for NcHMA4-1, 4-2, 4-3 and 4-4 were deduced using Vector NTI 11 translational software (Invitrogen, Paisley, UK). Amino acid sequences for N. caerulescens Herault (GenBank: CAD98808.1) and N. caerulescens Prayon (GenBank: AY486001.1), three HMA4 tandem repeat sequences from A. halleri (GenBank: EU382072.1 and EU382073.1) and the A. thaliana HMA4 sequence were sourced from NCBI (http://www.ncbi.nlm.nih.gov/guide/). All were aligned using ClustalW at default settings. The most informative regions were illustrated in Fig. 3.18. For all HMA4 sequences the least cross species protein conservation was recorded between AtHMA4 and NcHMA4-1 and NcHMA4_Prayon, sharing 72% sequence identity. Noccaea showed the highest protein conservation (84%) between NcHMA4-4 and all HMA4 proteins from both Arabidopsis species (Dot Matrix at default settings, Vector NTI 11) (Fig. 3.19). All sequences contained six identical predicted motifs found in P_{1B}-type ATPases including the E1-E2 ATPase phosphorylation site (DKTGTIT), the highly conserved CPx Motif involved in

ion translocation, a putative N-terminal heavy metal binding site (GICCTSE), a HP motif with unknown function, the ATP-binding motif (GDGxNDx) and a conserved region of the phosphatase domain (TGEA) (Papoyan and Kochian, 2004) (Fig. 3.18). All sequences, except NcHMA4-4 which has a truncation at amino acid 660, contained 8 predicted membrane-spanning domains and a large C-terminal tail concentrated with Cys pairs and multiple His residues.


Heavy-Metal-Binding
NcHMA4_Herault_ MALQKEIKNKEEDKKTKKKWQKSYFDVLGICCTSEIPLIENILKSLDGVKEYTVIVPSRT 60
NcHMA4_Prayon_ MALQKEIKNKEEDKKTKKKWQKSYFDVLGICCTSEIPLIENILKSLDGVKEYTVIVPSRT 60
NcHMA4-2 MALQKEIKNKEEDKKTKKKWQKSYFDVLGICCTSEIPVIENILKSLDGVKEYTVIVPSRT 60
NcHMA4-1 MALQKEDKNKEENKMTKKKWQKSYFDVLGICCTSEIPLIENILKSLDGIKDYTIIVPSRT 60
NcHMA4-3 MALQKEDKNKEENKMTKKKWQKSYFDVLGICCTSEIPLIENILKSLDGIKDYTIIVPSRT 60
NcHMA4-4 MALQKEDKNKEENKMTKKKWQKSYFDVLGICCTSEIPLIENILKSLDGIKDYTIIVPSRT 60
AhHMA4-2 MASQ----NKEEEKKKVKKLQKSYFDVLGICCTSEVPIIENILKSLDGVKEYSVIVPSRT 56
AhHMA4-3 MASQ----NKEEEKKKVKKLQKSYFDVLGICCTSEVPIIENILKSLDGVKEYSVIVPSRT 56
AhHMA4-1 MASQ----NKEEEKKKVKKLQKSYFDVLGICCTSEVPIIENILKSLDGVKEYSVIVPSRT 56
AtHMA4 MALQ----NKEEEKKKVKKLQKSYFDVLGICCTSEVPIIENILKSLDGVKEYSVIVPSRT 56
Phosphatase
NcHMA4_Herault_ DGIVVDGNCEVDEKTLTGEAFPVPKQRDSTVLAGTMNLNGYISVNTTALASDCVVAKMAK 300
NcHMA4_Prayon_ DGIVVDGNCEVDEKTLTGEAFPVPKQRDSTVLAGTMNLNGYISVNTTALASDCVVAKMAK 300
NcHMA4-2 DGIVVDGNCEVDEKTLTGEAFPVPKQRDSTVLAGTINLNGYISVNTTALASDCVVAKMAK 300
NcHMA4-1 DGIVVDGNCEVDEKTLTGEAFPVPKQRDSTVWAGTINLNGYISVNTTALASDCVVAKMAK 300
NcHMA4-3 DGIVVDGNCEVDEKTLTGEAFPVPKQRDSTVWAGTINLNGYISVNTTALASDCVVAKMAK 300
NcHMA4-4 DGIVVDGNCEVDEKTLTGEAFPVPKQRDSTVWAGTINLNGYISVNTTALASDCVVAKMAK 300
AhHMA4-2 DGIVVDGNCEVDEKTLTGEAFPVPKQKDSSVWAGTINLNGYISVKTTSLAGDCVVAKMAK 296
AhHMA4-3 DGIVVDGNCEVDEKTLTGEAFPVPKQKDSSVWAGTINLNGYISVKTTSLAGDCVVAKMAK 296
AhHMA4-1 DGIVVDGNCEVDEKTLTGEAFPVPKQKDSSVWAGTINLNGYISVKTTSLAGDCVVAKMAK 296
AtHMA4 DGIVVDGNCEVDEKTLTGEAFPVPKQRDSTVWAGTINLNGYICVKTTSLAGDCVVAKMAK 296
Ion Translocation
Phosphorylation
NcHMA4_Herault_ CPSGLILSTPVATFCALTKAATSGLLIKSADYLDTLSKIKIAAFDKTGTITRGEFIVIEF 420
NcHMA4_Prayon_ CPCGLILSTPVATFCALTKAATSGLLIKSAGHLDTLSKIKIAAFDKTGTITRGEFIVIEF 420
NcHMA4-2 CPCGLILSTPVATFCALTKAATSGLLIKSADYLDTLSKIKIAAFDKTGTITRGEFIVIEF 420
NcHMA4-1 CPCGLILSTPVATFCALTKAATSGLLIKSADYLDTLSKIKIAAFDKTGTITRGEFIVIEF 420
NcHMA4-3 CPCGLILSTPVATFCALTKAATSGLLIKSADYLDTLSKIKIAAFDKTGTITRGEFIVIEF 420
NcHMA4-4 CPCGLILSTPVATFCALTKAATSGLLIKSADYLDTLSKIKIAAFDKTGTITRGEFIVIEF 420
AhHMA4-2 CPCGLILSTPVATFCALTKAATSGLLIKSADYLDTLSKIKIAAFDKTGTITRGEFIVIDF 416
AhHMA4-3 CPCGLILSTPVATFCALTKAATSGLLIKSADYLDTLSKIKIAAFDKTGTITRGEFIVIDF 416
AhHMA4-1 CPCGLILSTPVATFCALTKAATSGLLIKSADYLDTLSKIKIAAFDKTGTITRGEFIVIDF 416
AtHMA4 CPCGLILSTPVATFCALTKAATSGLLIKSADYLDTLSKIKIVAFDKTGTITRGEFIVIDF 416

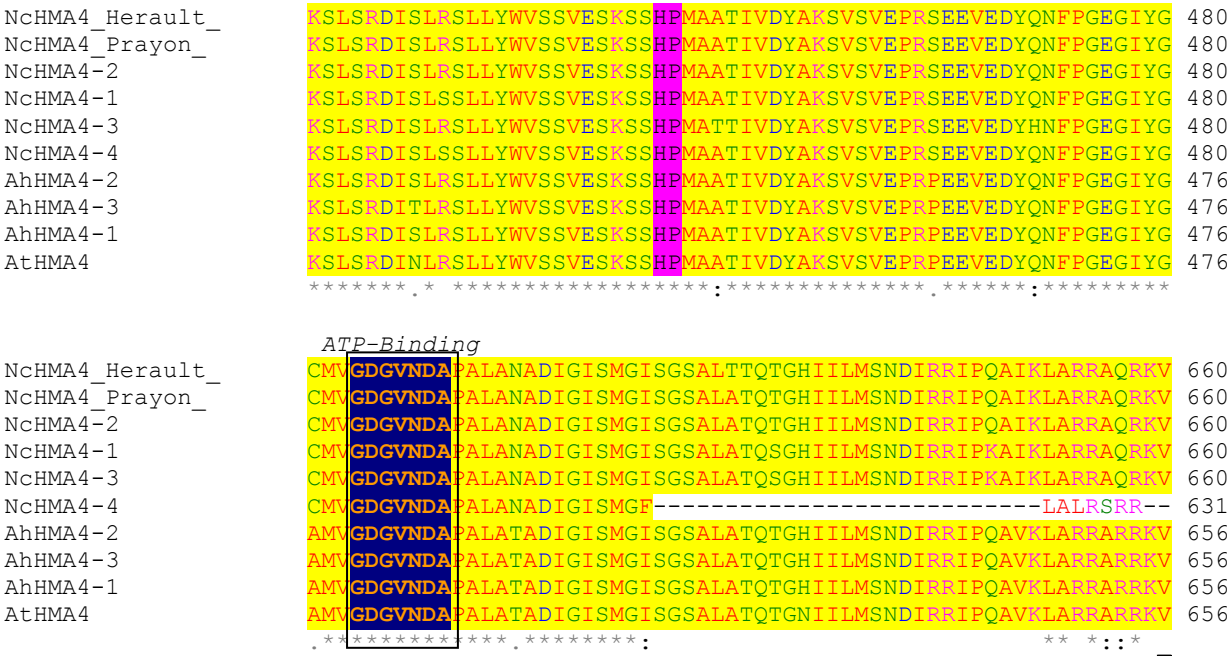


Figure 3.18 Multiple protein sequence alignment of all four deduced NcHMA4 tandem repeat protein sequences from N. caerulescens Saint Laurent Le Minier, two NcHMA4 sequences from ecotypes Herault (GenBank: CAD98808.1) and Prayon (GenBank: AY486001.1), three HMA4 tandem repeat sequences from A. halleri (GenBank: EU382072.1 and EU382073.1) and the A. thaliana HMA4 sequence.
Motifs common to P_{1B}-type ATPase 4 are indicated: DKTGTIT; the E1-E2 ATPase phosphorylation site (site of aspartyl phosphate formation), the highly conserved CPx Motif (ion translocation), GICCTSE; a putative N-terminal heavy metal binding site (underlined). HP motif (unknown function), GDGxNDx; ATP-binding motif and TGEA; part of the phosphatase domain.
"*": identical. ":": conserved substitutions (same colour group). ".": semi-conserved substitution (similar shapes).

Analysis of protein sequence parsimony of HMA4 orthologous sequences from A. halleri, A. thaliana and N. caerulescens using the Protpars software at default settings (Phylip version 3.68) (Section 2.8.2.2.) demonstrated that for both A. halleri and N. caerulescens, tandem gene repeats were grouped according to their genomic origin. Two groups containing two HMA4 repeats NcHMA4-1 and 4-2 and NcHMA4-3 and 4-4 were formed within N. caerulescens Saint Laurent Le Minier clade (Fig. 3.19 A). Published NcHMA4 sequences from N. caerulescens 'Prayon' and 'Herault' demonstrated a closer evolutionary relationship with NcHMA4-2 than all other N. caerulescens Saint Laurent Le Minier HMA4 sequences (Fig. 3.19 A).

Analysis of protein sequences using Dot Matrix at default settings (stringency of 30%, Vector NTI 11) revealed a closer phylogenetic relationship between HMA4 protein sequences within the Arabidopsis genus (> 90% sequence identity) than between the Arabidopsis and Noccaea genus (Fig. 3.19 B).

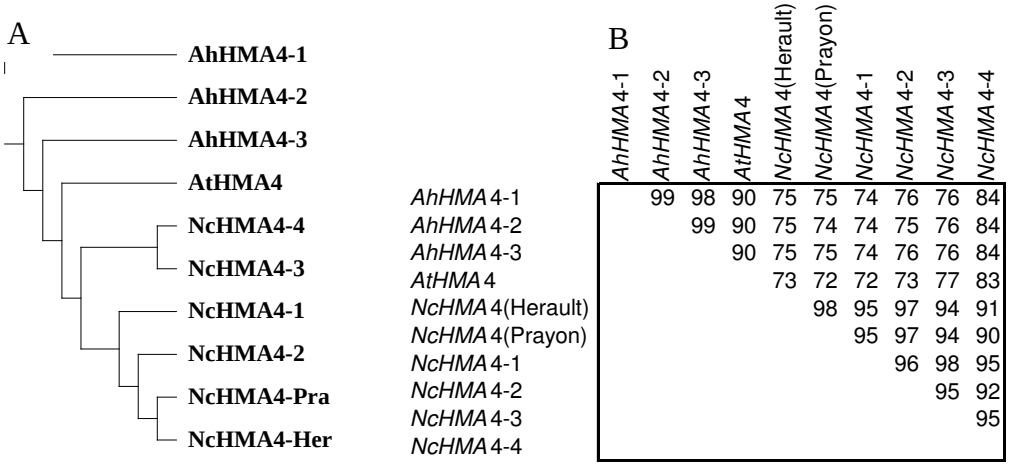


Figure 3.19 A) Cladogram and B) Dot Matrix comparison of protein sequences of HMA4 orthologues from Ah: Arabidopsis halleri, At: Arabidopsis thaliana and Nc: Noccaea caerulescens. Tandem repeats are highlighted by "-". "Pra" and "Her" refer to N. caerulescens ecotypes Prayon and Herault. The cladogram was created for nucleotide sequences through Protpars, Protein Sequence Parsimony Method and was run at default settings and supplied by Phylip version 3.68. The Dot Matrix program was run at default settings and supplied by Vector NTI 11.

3.6. DISCUSSION

This chapter has reported the identification of a novel quadruplication of HMA4 in the Zn and Cd hyperaccumulator plant Noccaea caerulescens Saint Laurent Le Minier. These results thus provide the first genomic evidence of HMA4 tandem duplication occurring independently in two distantly related Zn hyperaccumulating plants. Data from these sequences could be employed to perform functional and comparative analyses regarding novel genetic and cis regulatory roles within HMA4 in N. caerulescens. These gene sequences are being submitted to the GenBank database and represent an exceptional resource of novel genetic material which can be exploited to further develop scientific understanding of natural genetic variation and adaptation to numerous extreme environments.

Prior to this study there was no published evidence relating to HMA4 repeats within N. caerulescens. Physiological and classical genetic studies as well as recent molecular assays including transcriptome analysis identified possible molecular associations with Zn hyperaccumulation; however, progress was hindered due to a lack of clear genomic sequence (Broadley et al., 2007; Verbruggen et al., 2009). It was decided therefore, to create a publicly available genomic library of this species, thus providing direct access to reference genomic data.

3.6.1. Selecting Noccaea caerulescens Saint Laurent Le Minier to develop a genomic library

Noccaea caerulescens Saint Laurent Le Minier from Ganges (GA), southern France, was selected to develop a genomic DNA library due its rare Zn and Cd hyperaccumulation phenotype (Roosens et al., 2003; Liu et al., 2008). It is a small, short lived, nonmycorrhizal (Regvar et al., 2003), self compatible species thus making it ideal for molecular and functional plant based studies. Zinc hyperaccumulation and hypertolerance in N. caerulescens have been well documented, however functional genetic analyses of these traits have been severely hindered due to their constitutive nature throughout the genome and a lack of genomic sequence information (Broadley et al., 2007; Verbruggen et al., 2009). An F_2 population from an intraspecific cross between an ecotype from

Prayon (PR) (non-Cd hyperaccumulator) and (GA) (Zn and Cd hyperaccumulator) showed transgressive segregation of Zn accumulation, but not of Cd (Zha et al., 2004) and similarly for an F_2 cross between GA and a non-Cd hyperaccumulator ecotype La Calamine (LC) (Deniau et al., 2006). Genetic linkage maps created for the GA x LC population revealed 3 quantitative trait loci (QTL) involved in 44.5% of phenotypic variance in Zn concentrations in the shoot ($[Zn]_{shoot}$) and 3 QTL for phenotypic variance of $[Cd]_{shoot}$ (Deniau et al., 2006). All 3 Cd accumulation loci were derived from the GA ecotype. F_2 crosses between nickel (Ni) and Zn hyperaccumulator ecotypes demonstrated that Ni and Zn were pleiotropically controlled however Ni accumulation and tolerance appeared to be under nonpleiotropic control (Richau and Schat, 2008). Few molecular resources have been developed to aid functional research in N. caerulescens, with currently just 3 cDNA libraries published in the last decade (Lasat, 2000; Papoyan and Kochian 2004; Wei et al., 2008). Recent evidence in A. halleri demonstrated that cis factors were heavily involved in controlling Zn hyperaccumulation (Hanikenne et al., 2008). Providing a publicly available functionally robust genomic resource should therefore complement and further progress current studies into Zn hyperaccumulation within Noccaea caerulescens. It was therefore decided to produce a genomic library of N. caerulescens Saint Laurent Le Minier.

3.6.2. Creating a Genomic Fosmid Library

Functional and structural molecular genetic studies require accessibility to large insert libraries including bacterial artificial chromosome (BAC), yeast artificial chromosome (YAC), cosmid and fosmid libraries (Wang et al., 2009). Fosmid libraries have numerous advantages over most other systems including ease of manipulation, no bias for particular inserts as well as maintenance of highly stable cloned inserts (Kim et al., 1992). They have been successfully adapted to a wide variety of techniques including positional cloning, physical mapping and comparative analysis (Wang et al., 2009). A pCC1FOS™ Copy Control Fosmid Library was therefore constructed for Noccaea caerulescens Saint Laurent Le Minier. It contained 36,864 Escherichia coli EPI300-$T1^R$ host cells, each harbouring one ~40kb genomic insert and 8kb CopyControl™vector pCC1FOS™ resulting in a 5-fold Noccaea genomic coverage. The CopyControl

cloning system, permitted the manipulation of clones to either maintain low insert copy number and therefore maximum stability, or induce high copy number from an initial single copy to 10 – 200 copies to improve yield for downstream applications including sequencing, profiling and subcloning (Wild et al., 2002). Replication from the oriV (origin of replication) of the pCC1FOS™ vector was initiated by the trfA gene product of the phage T1-resistant EPI300-$T1^R$ cells (Epicentre Biotechnologies, 2010). Cells were stored in 96 x 384-well microtiter plates. DNA was arrayed evenly onto two nitrocellulose filters each organised into 6 sections, each of which contained DNA from 8 microtiter plates, each arranged into 16 rows (A – P) and 24 columns (1 – 24) producing a total of 384 'wells'. Each well contained two identical DNA fragments, 'spots', specifically arranged to indicate their plate of origin. Any positive probe hybridisations therefore, were specifically located to a particular 'plate' by the arrangement of the 'spot' pairs. Filters contained approximately 5% ribosomal and 15% chloroplast contamination. Since the primary aim of this thesis was to examine natural variation in Zn accumulation in the Brassicaceae, this library was employed to isolate candidate Zn transporter genes from N. caerulescens for both sequencing and later functional analyses

3.6.3. HMA4 as a candidate to test the library

Recent advances in gene and protein expression profiling (Assunção et al., 2001; Filatov et al., 2006; Hammond et al., 2006; Rigola et al., 2006; Talke et al., 2006; van de Mortel et al.,2006) as well as functional analyses of candidate proteins using heterologous expression systems including plants and yeast (Papoyan and Kochian, 2004, 2006; Hanikenne et al., 2008; Barabasz et al., 2010) have identified conserved candidate orthologues involved in Zn transport including members of the P_{1B}-type ATPase Heavy Metal Associated (HMA) family. By combining and re-analysing published global shoot transcriptome data using the full Affymetrix A. thaliana ATH1-121501 (ATH1) GeneChip array (Affymetrix Inc., Palo Alto, CA, USA), for N. caerulescens vs T. arvense (Hammond et al., 2006), A. halleri vs A. lyrata ssp. Petraea (Filatov et al., 2006) and A. thaliana (Talke et al., 2006), HMA homologues, were found to be significantly differentially expressed between hyperaccumulators and non-hyperaccumulators (Broadley et al., 2007). More specifically, in a cross species

transcriptome study comparing N. caerulescens to the non-hyperaccumulator Thlaspi arvense, HMA4 was shown to be significantly more highly expressed (2.19 fold) in Noccaea than Thlaspi (Hammond et al., 2006). On growth media supplemented with < 100 µM Zn ($[Zn]_{ex}$) RT-PCR investigations revealed that NcHMA4 expression remained unaltered (Xing et al., 2008) however upon addition of later studies by Hammond et al. (2006), showed that an addition of > 150 µM $[Zn]_{ext}$ increased NcHMA4 expression significantly but resulted in no negative growth affects in Noccaea, compared with Thlaspi, which showed reduced growth at > 30 µM $[Zn]_{ext}$. Moreover, increase in NcHMA4 expression was coincidental with a significant increase in the Zn concentration per g^{-1} f. wt. of shoot tissue ($[Zn]_{shoot}$) (Hammond et al., 2006). Functional analysis of an NcHMA4 homolog in yeast indicated that it could have a role in xylem loading of Zn (Papoyan and Kochian, 2004). This was supported by observations of a role for AtHMA2 and AtHMA4 in maintaining Zn homeostasis in A. thaliana through Zn transport into the root xylem (Hussain et al., 2004). Further overexpression of AtHMA4 in A. thaliana increased both Zn and Cd shoot concentrations (Verret et al., 2004). More recently, heterologous expression of AhHMA4cDNA under its endogenous A. halleri promoter in A. thaliana resulted in increased Zn concentrations in xylem parenchyma, resembling typical Zn distribution in A. halleri roots (Hanikenne et al., 2008). Expression of this construct in tobacco resulted in increased Zn concentrations in leaf tissue under Zn deficient conditions (Barabasz et al. 2010). Despite such convincing putative roles for HMA4 in Zn accumulation within plant species, a genomic sequence for N. caerulescens has not been made publicly available. No knowledge exists regarding the structure of this locus, its synteny or the degree of intergenic conservation which exists with other genomes. It was therefore decided to elucidate the full genomic structure of the HMA4 locus in Noccaea caerulescens Saint Laurent Le Minier using a data available from a genomic library.

A 421 bp probe was designed to a published full length 3561 bp N. caerulescens HMA4 cDNA sequence (GenBank: AY486001.1) (Papoyan and Kochian, 2004) and amplified from N. caerulescens Saint Laurent Le Minier DNA. The fragment was hybridized to the library and identified 36 clones, seven of which were confirmed as containing a HMA4 fragment through PCR amplification and

Sanger based sequencing. It was necessary to gain more detailed information regarding their full length genomic and promoter sequences before functional studies could be performed. Locus specificity, copy number and their level of synteny with both *A. thaliana* and *A. halleri* HMA4 regions could not be determined without full length sequencing of the fosmid insert (20 – 40 Kb). Therefore a sequencing approach which was more efficient than the traditional Sanger technique was required. It was thus decided to employ the highly efficient pyrosequencing technique to sequence these fosmids.

3.6.4. Sequencing Noccaea caerulescens HMA4

A combination of pyrosequencing using the Next Generation Genome Sequencer (GS) FLX 454 sequencer from Roche (Roche, Basel, Switzerland) (Nyrén, 2007; Pettersson et al., 2009; Lundin et al., 2010) and traditional Sanger sequencing (Sanger et al., 1977) using the ABI 3730 XL sequencer were employed to individually sequence the entire genomic DNA inserts in each of the five fosmid clones. The respective strengths of both technologies, pyrosequencing (GS FLX) and Sanger sequencing (ABI 3730 XL), were engaged to help improve accuracy and flexibility of data compilation. Individually both systems demonstrated important benefits to the sequencing process.

For the GS FLX system lengthy bacterial subcloning was replaced by a more efficient molecular titration and purified DNA amplification step. The library of sequences generated, presented no cloning bias and was well represented with difficult to clone AT-rich regions (Dr. Axel Strittmatter pers. comm.). Regarding throughput, the Standard-series chemistry which can produce over 100 million bases (Mb) within an 8 hour run, resulted in 400,000 reads with an average read length of 220 – 270 bases per run. The Titanium-series chemistry, which could produce between 400 – 500 Mb per run or 1 Gb per day, resulted in 1 million reads of an average length of 350 – 450 per run (Pettersson et al., 2009; Dr. Axel Strittmatter pers. comm.). Regions which are challenging for most sequencing chemistries to process include GC-rich templates, since unique secondary structures and melting temperatures, due to an increased proportion of three hydrogen bonds, result in stronger association between the double-stranded DNA (dsDNA). Relatively few irregularities were observed when sequencing these

areas with GS FLX Titanium (Dr. Axel Strittmatter, pers. comm.). Analyses of sequence quality reported >99.99% consensus accuracy at 20 fold sequence coverage, with single read accuracy >99.5% for over 200 bp reads and >99% for over 400 bp reads.

Unlike the GS FLX systems, the Sanger dideoxy sequencing chemistry using the ABI 3730 XL system could only produce 1.1 Mbp per day (Sanger et al., 1977; Pettersson et al., 2009), thus it was unsuitable as an efficient approach to whole (48 Kbp) fosmid sequencing. However, it could produce read lengths of 1.1 Kbp per run and so was employed in a robust primer walking strategy to complete sequencing in region which were too challenging for GS FLX, such as large (>1 Kbp) repetitive areas.

A suite of competitive Next Generation sequencing systems exist, each with their own benefits and limitations, such as greater numbers of reads per run including by the Illumina system (Illumina, 520 million reads), the Polonator system (Danaher Motion, 400 million reads) and the SOLiD (Sequencing by Oligonucleotide Ligation and Detection) system (Applied Biosystems, 750 million reads). Read lengths however remain far shorter than that of the FLX system with the Illumina, Polonator and SOLiD systems producing just 125, 26 and 75 bp respectively. For de novo sequencing of the Noccaea caerulescens fosmids, a combination of GS FLX and Sanger sequencing systems were therefore selected.

A pool of 13 fosmids from the fosmid library, 6 of which were positive for the NcHMA4 insert, were sequenced in a $1/16^{th}$ of a run using a GS FLX sequencer and Standard-series chemistry. This produced reads of an average of 220 -270 bp each representing 5 fold coverage however did not produce complete fosmid insert sequences. Subsequent GS FLX Titanium-series chemistry produced 350 – 450 bp per reads with an average of 20 fold fosmid coverage. In tandem with Sanger (ABI 3730 XL) sequencing, all five fosmids were completely sequenced. Two fosmids, P6P46 and B3P40, contained tandemly repeating gene sequences (>7.5 Kb), however both were completely sequenced using the GS FLX system Titanium-series. One fosmid, H2P47, required additional sanger sequencing due

to large difficult to sequence AT stretches. Using in silico techniques including commercial sequence alignment and contig assembly software, fosmid inserts were compared to publicly available A. halleri and A. thaliana full length HMA4 regions. Data indicated the presence of four NcHMA4 gene copies and that all were tandemly repeated in one region which was sytenic with orthologous regions in A. halleri and A. thaliana.

3.6.5. HMA4 is tandemly quadruplicated in Noccaea caerulescens Saint Laurent Le Minier

This chapter describes the identification of a previously unreported locus specific HMA4 quadruplication in Noccaea caerulescens Saint Laurent Le Minier. These findings provide vital evolutionary and molecular genetic information within this Zn hyperaccumulator and should provide the basis for future functional molecular research into Zn hyperaccumulation within Noccaea caerulescens. 36 fosmid clones isolated from the library following hybridization with the NcHMA4 probe were subjected to a colony PCR, using primers designed to amplify the NcHMA4 probe. Seven fosmids produced PCR products. Of the remainder at least 12 fosmids were shown to be the result of cross hybridisations with the closely related HMA2 and HMA3 and the remainder could have been the result of nonspecific hybridizations to sequences showing identity to HMA4. Sanger sequencing of HMA4 PCR products confirmed the presence of a HMA4 sequence. To elucidate whether these clones contained identical copies of a single HMA4 gene or whether there were genomic copies within the Noccaea, a DNA profile was performed for each, using a combination of endonuclease digestion and southern analysis. Results indicated the presence of multiple copies. Initial GS FLX (Standard chemistry) sequence results of these fosmids (in a pool of 13 fosmids) similarly indicated the presence of multiple copies, however contig assembly and sequence alignments (ContigExpress and AlignX software, Vector NTI 11, Invitrogen, Paisley, UK) could not identify complete genomic sequences for these copies. PCR analyses supported these results and further indicated that gene copies could be located in one locus within the Noccaea caerulescens genome. A combination of GS FLX (Titanium Series) and Sanger (ABI 3730 XL) sequencing on five selected fosmids revealed that two contained two genomic copies of NcHMA4 while the remainder contained one copy each. Following in

silico assemblage, by ContigExpress (VectorNTI 11, Invitrogen, Paisley, UK), a 101480 bp single locus containing four tandem repeats of NcHMA4 was identified. Sequences flanking the 5' and 3' ends of the tandem series were sytenic with the orthologous Arabidopsis thaliana and A. halleri orthologous loci. N. caerulescens HMA4 repeats demonstrated a moderate percentage identity with AtHMA4 (72 – 83%) and all three AhHMA4 repeats (74 – 84%) (Dot Matrix , VectorNTI 11, Invitrogen, Paisley, UK). NcHMA4 tandem repeats share between 93 – 99% genomic sequence identity compared with just 72 – 76% with HMA4 sequences from A. thaliana and A. halleri, suggesting that the quadruplication may have occurred relatively recently in the N. caerulescens lineage. NcHMA4-4 shared the lowest percentage identity among the Noccaea HMA4 repeats (92 – 95%) but the highest among inter species HMA protein sequences (83 – 84%). All NcHMA4 tandem repeats contained the majority of the predicted motifs found in AtHMA4 and AhHMA4 repeats and other heavy metal ATPases including, the E1-E2 ATPase phosphorylation site, the highly conserved CPx motif for ion translocation, the putative N-terminal heavy metal binding site, the ATP-binding motif and the TGEA phosphatase domain (Bernard et al., 2004; Papoyan and Kochian, 2004). All except NcHMA4-4 contained eight predicted membrane-spanning domains and a large C-terminal tail containing several putative heavy metal binding domains. The deduced protein for NcHMA4-4 displayed an early truncation at amino acid 660 and contains just six membrane-spanning domains. It is not known how this might affect gene function however evidence from heterologous studies in yeast suggest that the C terminus is extremely important for heavy metal, including Zn binding due to the presence of a large concentration of His and Cys residues (Papoyan and Kochian, 2004). Previous observations suggested a role for Cys dipeptides and His-rich domains in heavy metal binding in ATPases (Solioz and Vulpe, 1996; Williams et al., 2000). It was shown that the His-9 concatamer along with the high concentration of Cys pairs and single His residues in the C terminus of NcHMA4 were necessary for increased heavy metal tolerance and hyperaccumulation in yeast, and those that expressed a truncated peptide, without the poly-His tail, demonstrated significantly less tolerance and accumulation capabilities (Papoyan and Kochian, 2004). Within the Zn and Cd hyperaccumulator A. halleri a recent independent triplication of HMA4 has similarly occurred (Hanikenne et al., 2008). Heterologous expression of AhHMA4

cDNA under the control of its endogenous promoter in A. thaliana associated increased HMA4 transcripts with increased Zn concentrations in xylem root tissues compared to wild type A. thaliana which had highest Zn concentrations in the root pericycle cell layer (Hanikenne et al., 2008). It was concluded from these investigations that high HMA4 expression was specified in cis and amplified by the triplication of gene copies (Hanikenne et al., 2008). These conclusions were consistent with studies in Zea mays ssp. Mays (maize) where it was observed that up stream cis regulatory elements significantly altered the tb1 gene transcription, which generated the morphological differences that distinguish Zea mays ssp. Mays from its wild teosinte ancestor (Clark et al., 2007). Additionally, concerted evolution of duplicated genes in yeast was shown to be selected to attain higher gene dosages in response to environmental selection (Sugino and Innan, 2006). This was supported in studies involving Escherichia coli where adaptation to growth-limiting concentrations of lactulose was consistently achieved through duplications of lac operons, involved in the transport and metabolism of lactose (Zhong et al., 2004).

In conclusion, a Noccaea caerulescens Saint Laurent Le Minier genomic fosmid library has been functionally tested using comparative genomics to probe and identify P_{1B}-type ATPase HMA4 sequences orthologous to genes in both the A. thaliana and A. halleri genome. Four previously unreported tandem repeats of the NcHMA4 have thus been identified and fully sequenced through Next Generation Genome Sequencer FLX 454 sequencing.

The NcHMA4 genetic information was consistent with findings from a similar study in the Zn hyperaccumulator A. halleri where gene triplication and altered cis regulation were shown to regulate Zn hyperaccumulation (Hanikenne et al., 2008). To establish whether cis regulation was similarly altered in N. caerulescens, a comparison of HMA4 expression from N. caerulescens, A. halleri and A. thaliana promoter sequences was initiated. A. thaliana was employed as a heterologous expression system. These experiments are described in detail in Chapter 4.

3.7. SUMMARY

- A genomic fosmid library was constructed for Noccaea caerulescens Saint Laurent Le Minier composed of 36,864 Escherichia coli EPI300-T1R clones, containing 40kb of randomly sheered genomic DNA ligated to 8kb CopyControl™vector pCC1FOS™, giving a 5-fold genomic coverage.
- A PCR generated probe, designed using a published Noccaea caerulescens 'Prayon' HMA4 cDNA sequence, identified seven HMA4 clones from the Noccaea genomic fosmid library.
- Multiple independent HMA4 copies were identified through DNA fingerprinting (restriction endonuclease digestions and southern hybridisations) using six fosmid clones.
- Five clones were individually sequenced through the Genome Sequencer FLX 454 pyrosequencing platform and assembled into one contiguous sequence using Vector NTI software.
- A tandem quadruplication of HMA4 was isolated from Noccaea caerulescens Saint Laurent Le Minier.
- Regions flanking the entire N. caerulescens HMA4 tandem repeat locus were highly collinear with both Arabidopsis thaliana and A. halleri.
- Coding and protein sequences were substantially conserved among all gene copies among Noccaea and both Arabidopsis species, with Noccaea demonstrating 72 – 83% identity at a protein level.
- Noccaea HMA4 promoter regions have diverged significantly from both A. thaliana and A. halleri showing a maximum of 51% sequence identity and may present altered cis regulation of NcHMA4.

CHAPTER 4 EXPRESSION ANALYSIS OF HMA4

4.1. INTRODUCTION

A previously unreported tandem quadruplication of HMA4 in the zinc (Zn) hyperaccumulator Noccaea caerulescens was described in Chapter 3 and was consistent with the published HMA4 triplication in the Zn hyperaccumulator Arabidopsis halleri (Hanikenne et al., 2008). For both genomes, these regions were highly syntenic with the HMA4 locus in Arabidopsis thaliana (Chapter 3). It is remarkable that both genomes, which hyperaccumulate Zn in leaf tissue, could have increased copy numbers of a gene which has been shown to be involved in Zn transport from roots to shoots in plants (Mills et al., 2003, 2005; Hussain et al., 2004; Verret et al., 2004, 2005, Hanikenne et al., 2008; Barabasz et al., 2010; Siemianowski et al., 2010). Even more astounding, is that these tandem replications of HMA4 orthologues associate with Zn hyperaccumulation traits which have evolved independently i.e. once in the Arabidopsis halleri species and once in the distantly related Noccaea clade within the Brassicaceae (Broadley et al., 2007; Verbruggen et al., 2008). It could be therefore, that replication of HMA4 in plants forms part of a specialist mechanism to hyperaccumulate heavy metals including Zn into aerial shoot tissue. However, at a sequence level N. caerulescens HMA4 (NcHMA4) proteins demonstrated similarly high sequence identities with orthologues from both A. halleri and the Zn non-hyperaccumulator A. thaliana (between 72 – 84%) (Chapter 3). For A. halleri and A. thaliana, HMA4 orthologues shared 90% sequence identity at the protein level (Chapter 3). Conversely, sequences flanking the 5' regions of A. halleri HMA4 (AhHMA4) repeats were shown to be highly divergent from A. thaliana (Hanikenne et al., 2008). This group similarly demonstrated that these sequence divergences resulted in regulatory innovations in A. halleri and that enhanced AhHMA4 expression levels of individual repeats was a result of this altered cis regulation. A. thaliana and A. halleri transformed with constructs containing fusions of the promoters of AtHMA4 and all three AhHMA4 repeats with the *β*-glucuronidase (GUS) reporter gene, exhibited significant expression variation between species. Reporter activity under the regulation of the AtHMA4 promoter (AtHMAp) was significantly lower than that observed for each of the three A. halleri HMA4 promoters which showed similar activity to that of the

strong constitutive cauliflower mosaic virus (CaMV) 35S promoter (Hanikenne et al., 2008).

Spatially, expression of all three AhHMA4p repeats were similar, showing high levels of GUS activity in the root pericycle and xylem parenchyma for both transformed plant species. Within the shoots, expression under AhHMA4 promoters was significantly more widely distributed throughout the lamina in both A. halleri and A. thaliana than for the AtHMA4 promoter, which was localised to the stem and shoot vascular tissue in agreement with findings by Hussain et al. (2004). This HMA4 expression pattern was confirmed by in situ hybridisation studies and was consistent with putative roles for HMA4 in metal distribution throughout the lamina and exclusion of metals from specific cell types (Papoyan and Kochian et al., 2004; Hussain et al., 2004; Mills et al., 2005; Hanikenne et al. 2008; Barabasz et al., 2010; Siemianowski et al., 2010). Collectively this data confirmed that triplication of AhHMA4 and altered cis regulation resulted in high transcript levels in A. halleri.

It was hypothesised therefore that a similar regulatory mechanism could have evolved in Noccaea caerulescens and that altered expression of NcHMA4 repeats was specified in cis. To test this, sequences were first compared between all HMA4 promoter regions in A. thaliana, A. halleri and N. caerulescens to examine sequence divergence. The expression pattern of the NcHMA4-2, AhHMA4-3 and AtHMA4 promoters were then compared using A. thaliana as a heterologous expression system to demonstrate similarities between both Zn hyperaccumulator species.

4.1.1. The expression vector pGWB3 for promoter expression analysis

For all promoter expression analyses, the plant expression vector pGWB3 was employed as a basis for promoter fused with GUS (promoter::GUS) constructs. The pGWB series of plant vectors were created as a result of modifications imposed on traditional vector systems to facilitate comprehensive adaption to the efficient Gateway™ cloning system (Nakagawa et al., 2007). The original binary vector employed as the initial framework for this pGWB system was pABH-Hm1 (Mita et al., 1995), which itself was developed from pBI101 through the

introduction of the beta amylase promoter and HPT unit (Nakagawa et al., 2007). The HindIII-SacI region in pABH-Hm1, which contained the β-amylase promoter::GUS, was replaced by the Gateway™ cassette and tag (Nakagawa et al., 2007). This system was demonstrated to be compatible with a wide variety of marker genes for promoter expression analyses including *β*-glucuronidase (GUS), luciferase (LUC), synthetic green fluorescent protein with the S65T mutation (sGFP), enhanced yellow fluorescent protein (EYFP), and enhanced cyan fluorescent protein (ECFP) (Nakagawa et al., 2007). The applicability of this pGWB3 vector to promoter::GUS analyses was confirmed by the authors where expression levels from a phosphate induced promoter (PHT1) promoter::GUS showed similar qualitative and quantitative GUS activity to previous independent observations (Nakagawa et al., 2007). The pGWB3 contained the NPTII and HPT genes which conferred kanamycin and hygromycin resistance respectively in transformed plants. Bacteria transformed with pGWB3 prior to the Gateway™ cloning reaction were conferred with kanamycin, hygromycin and chloramphenicol resistance while those bearing vectors created following this reaction were both kanamycin and hygromycin resistant (Nakagawa et al., 2007).

4.1.2. Agrobacterium tumefaciens GV3101 as a plant transformation system

The Agrobacterium tumefaciens strain GV3101 (Koncz and Schell, 1986) was selected for Arabidopsis thaliana transformation for all pGWB3 promoter::GUS fusion constructs since it was widely considered a highly efficient floral dip-based T-DNA insertion vector. It carried chromosomal resistance to rifampicin and had a Ti-plasmid coding virulence genes and gentamycin resistance. Since this strain was sensitive to kanamycin, it was considered an ideal system to select for transformed strains containing pGWB3 vectors which conferred resistance to the antibiotic. To maintain selection for this strain of A. tumefaciens, rifampicin and gentamycicn were supplemented in growth media at concentrations of 25 µg ml^{-1}.

4.1.3. Arabidopsis thaliana Col-0 as a heterologous expression system

Efficient reproducible transformation systems are currently under investigation for both Zn hyperaccumulators Noccaea caerulescens and Arabidopsis halleri. For N. caerulescens, a floral dip based approach involving a modified approach to that reported by Clough and Bent (1998) was published (Peer et al., 2003). Subsequently a tissue culture-based system was reported to have regenerated transgenic N. caerulescens plants (Guan et al., 2008). No succeeding transformations have since been reported in this genus, suggesting that current Agrobacterium tumefaciens mediated transformation techniques may be inefficient for certain ecotypes of this species. Within A. halleri one successful series of A. tumefaciens mediated transformations have been reported by Hanikenne et al. (2008). This was based on the inefficient tissue culture based approach and no subsequent transformation events have been reported. Heterologous expression systems have been successfully employed for such recalcitrant species including the use of yeast (Saccharomyces cerevisae) to elucidate the functional role of HMA4 in N. caerulescens (Papoyan and Kochian, 2004). Plant based systems represent a more favoured approach with, A. thaliana established as one of the most efficient and successful transformation systems (Clough and Bent, 1998) and therefore suitable for heterologous promoter expression assays. It was therefore decided to use A. thaliana as a heterologous system to analyse promoter expression for all HMA4 promoters.

4.1.4. Reporter Genes

The UIDA gene encodes the *β*-glucuronidase (GUS) enzyme from E. coli and has been extensively as a histochemically and fluorometrically detectable marker gene in transgenic systems (Jefferson et al., 1987; Fior et al., 2009). One of the foremost advantages of the GUS reporter gene system throughout plants lies in the stability of its expression within all cells under all plant physiological conditions (Fior et al., 2009). Its activity can be routinely detected by application of a range of differently coloured chromogenic or fluorogenic enzyme substrates. Unlike the majority of available marker genes, very little background signal is encountered since intrinsically, there is no significant endogenous expression of *β*-glucuronidase (GUS) in plants or within the taxon Eubacteria (Jefferson et al., 1987). Additionaly the presence of this E. coli based enzyme bears no effect on

plant metabolism. For some marker genes, such as the green fluorescent protein (GFP) from the jelly fish Aequorea victoria, chlorophyll often conceals expression levels and so specialist filters are required to correctly view protein activity. For the GUS gene however, the easily detectable and insoluble colours, highlighted in regions of protein expression, become clearly apparent as the enzymes hydrolyze the glycosidic bond between the applied chromogenic substrate (glucuronic acid) such as X-gluc (5-bromo-4-chloro-3-indolyl β-D-glucuronic acid) and molecules of endogenous and exogenous compounds within living tissue (Jefferson et al., 1987; Fior et al., 2009). In addition, a wide range of β-glucuronic acid substrates are available for detection of GUS expression. Moreover, the expression of GUS can be quantified using a fluorometric assay and thus requires no DNA extraction, electrophoresis or autoradiography to detect expression levels in response to external stimuli. This assay is commonly performed using the fluorogenic substrate 4-methyl-umbelliferyl-β-d-glucuronide (MUG) through a discontinuous measurement system (Fior et al., 2009). It was considered appropriate for this analysis since it was relatively easy to employ; it was highly sensitive, relatively inexpensive, extremely reliable, stable and safe, and no specialist equipment was required to view the final product (Jefferson et al. 1987; Fior et al., 2009).

This chapter describes the use of Arabidopsis thaliana Col-0 as a heterologous expression system to investigate the variation in expression patterns between HMA4 promoters isolated from two Zn hyperaccumulators, A. halleri and Noccaea caerulescens, and the Zn non-hyperaccumulator A. thaliana, which were each fused to the *β*-glucuronidase (GUS) reporter gene.

4.2. AIMS

The aim of the work in this chapter was to test for functionality and altered spatial expression of HMA4 cis sequences isolated from A. thaliana, A. halleri and Noccaea caerulescens using a heterologous transformation approach in A. thaliana.

4.3. OBJECTIVES

- To clone the HMA4-2 promoter sequence from Noccaea and develop promoter::GUS fusion constructs to identify spatial patterns and observe expression.
- To clone HMA4-1 and HMA4-3 promoter sequences from Arabidopsis halleri and develop promoter::GUS fusion constructs to identify spatial patterns and observe expression.
- To clone the HMA4 promoter sequence from Arabidopsis thaliana and develop promoter::GUS fusion constructs to test for spatial patterns and observe expression.
- To generate T_2 A. thaliana transformants and test for expression of all promoter::GUS constructs in whole plantlets.

4.4. MATERIALS AND METHODS

4.4.1. Creating promoter::GUS fusion constructs

Primers specific for the promoter regions of AtHMA4, AhHMA4-1, AhHMA4-3 and NcHMA4-2 were designed using the internet package Primer 3 Version 0.4.0 (Section). Promoter fragments were PCR amplified from genomic plant DNA by Phusion® proofreading polymerase and were A-Tailed (Section 2.5.6.1.) before ligating into the pCR8®/GW/TOPO® entry vector using the TA cloning system (Section 2.5.6.2.). Entry vector clones were then selected on solid LB overnight and transformed colonies were grown for 16 hrs in liquid LB for plasmid extraction (Section 2.5.7.). Plasmid concentrations were quantified using Nanodrop software (Section 2.5.7.) and molar ratios of plasmids were calculated as per manufacturer's guidelines for the LR reaction protocol (Invitrogen, Paisley, UK). The reaction mixture consisted of 150 ng of entry vector (pCR8®/GW/TOPO® harboring the promoter of interest) and 150 ng of destination vector (pGWB3) and sterile distilled water (SDW) to a final volume of 8 µl. To each reaction 2 µl of the enzyme clonase™ was added to initiate the LR reaction. The solutions were mixed by vortexing twice for 2 sec and centrifuged for 4 sec. Samples were incubated for 16 hrs at room temperature (RT). To terminate the reaction 1 µl of proteinase K was added, vortexed for 1 sec and incubated at 37°C for 10 min. The process was summarised in Fig. 4.1.

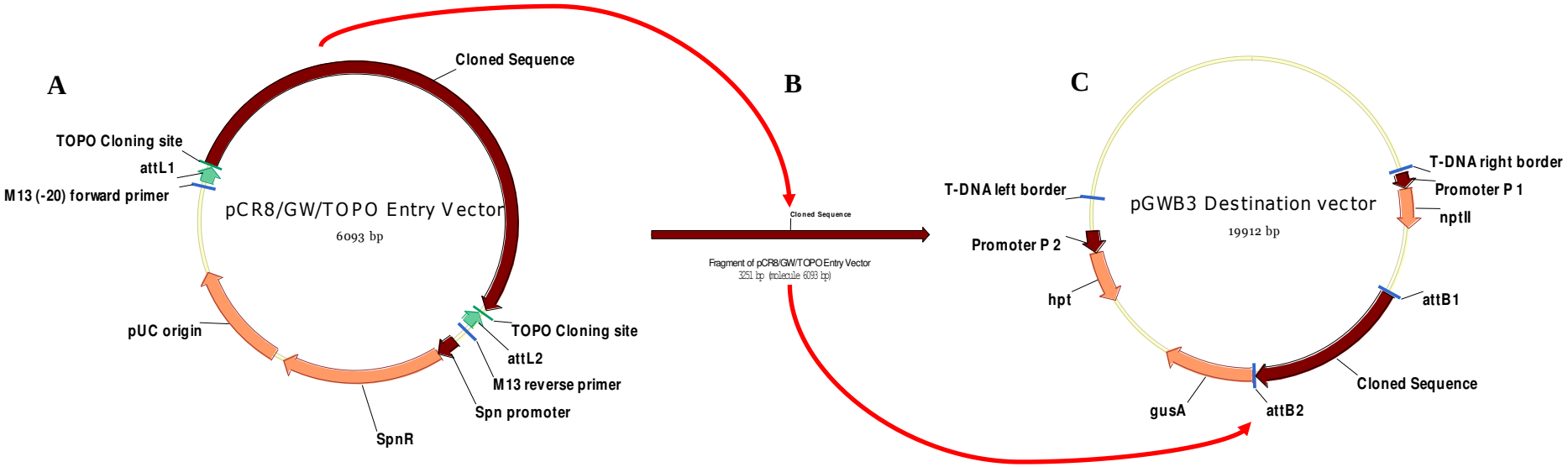


Figure 4.1 The LR reaction, was a recombination reaction between attL sites (green arrows) on the entry vector (A) and the attR sites on the destination vector in response to clonase ™. The reaction generated an expression clone (C) which was catalyzed by recombinant proteins. The entry vector (pCR8/GW/TOPO) was mixed with the appropriate Gateway® destination vector pGWB3 and Gateway® Clonase enzyme to initiate the LR reaction which transferred the cloned sequence (B) (brown arrow) to the destination vector. Image created using Vector NTI 11 software (Invitrogen, Paisley, UK).

4.4.2. Bacterial transformations

Chemically competent Escherichia coli DH5α cells were transformed with 4 µl of these reactions and transformed colonies selected for 16 hrs, at 37°C, on solid LB supplemented with kanamycin (50 µg ml^{-1}) and hygromycin (50 µg ml^{-1}) (Section 2.5.6.4.). Plasmid transformation was identified by PCR amplification using primers specific for GUS, NPTII and the promoter insert of interest. Identified plasmids were sequenced using M13 primers (Section 2.8.1.) to confirm insertion of the promoter in the correct orientation with the GUS gene. Plasmids were extracted from confirmed transformed E. coli and transformed into electrocompetent Agrobacterium tumefaciens strain GV3101 (Section 2.5.6.5.). A. tumefaciens were selected for transformants for 3 d at 29°C in darkened conditions on solid LB supplemented with kanamycin (50 µg ml^{-1}), hygromycin (50 µg ml^{-1}), gentamycin (50 µg ml^{-1}), tetracycline (5 µg ml^{-1}) and rifampicin (50 µg ml^{-1}).

4.4.3. Analysis of GUS expression in T_2 transgenic Arabidopsis thaliana

The protocol used for the histochemical detection of GUS activity was based on the method described by Jefferson et al. (1987), and was performed on T_2 segregating Arabidopsis thaliana plants transformed using the floral dip method (Section 2.9.) with promoter::GUS fusion constructs AhHMA4-3p::GUS, AtHMA4::GUS and NcHMA4-2p::GUS. Transformed plants were germinated in vitro on agar-based medium supplemented with 50 µg ml^{-1} kanamycin sulphate. Seven d after sowing (DAS), healthy, green actively growing plants were transferred under axenic conditions to translucent polycarbonate growth boxes containing 75 ml agar-based media (10 g l^{-1} sucrose, 8 g l^{-1} agar and 2.1 g l^{-1} MS and with no kanamycin sulphate supplementation (Section 2.9.1.). Plants were cultured for a further 14 days (21 DAS).

For each replicate all samples from each line were placed into individual sterile glass universals (3 plants per bottle) containing 10ml of GUS assay solution containing (5mg of X-Gluc (5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid; Melford Laboratories Ltd, Ipswich, UK) dissolved in 100µl of dimethyl formamide (DMF), phosphate buffer (0.2M NaH_2PO_4 plus 0.2M Na_2HPO_4, pH 7.0), 0.5M Na_2EDTA, 10mM $K_3Fe(CN)_6$, 10mM $K_4Fe(CN)_6.3H_2O$ and 0.1%

(v/v) Triton-X-100 (Sigma-Aldrich Co., Steinheim, Germany). The explants were incubated in the dark at 37°C for 16 h.

Chlorophyll was removed from each sample to aid later imaging of GUS staining from the histochemical treatment. Initially samples were subjected to acidified methanol (2ml Conc HCl + 10 ml methanol + 38 ml Water) for 15 min at 50°C with intermittent gentle shaking by hand. Solutions were discarded and retained samples were suspended in a neutralisation solution (7% NaOH in 60% ethanol) for 15 min at RT. The solution was discarded and the retained samples were rehydrated using a series of decreasing concentrations of ethanol (from 40, 20 and 10% v/v). Once fully rehydrated in sterile distilled H_2O (SDW) samples were mounted in 50% glycerol (v/v) to be viewed under a stereo microscope for traces of indigo staining to confirm the presence of GUS activity. As a control, non-transformed A. thaliana plants which were not selected on kanamycin were subjected to identical histochemical treatment. Images of individually assayed samples were then recorded.

4.4.4. Primers employed

Sequences of primers employed to isolate HMA4 promoter sequences from A. thaliana, A. halleri and Noccaea caerulescens were from 5' to 3':

NcHMA4-2 promoter_Fwd	CTCCTTCTGTAACGCCATTTCTGTA
NcHMA4-2 promoter_Rev	CTCCTTCTGTAACGCCATTTCTGTA
AtHMA4 promoter_Fwd	ACTTACCGATCGGGTATGCCATG
AtHMA4 promoter_Rev	TTTCTCTTCTTCTTTGTTTTGTAACGCC
AhHMA4-1 promoter_Fwd	CACGGGACTGGTTATATTTCGGAAATGA
AhHMA4-1 promoter_Rev	TTTCTCTTCTTCTTTGTTTTGTGACGCC
AhHMA4-3 promoter_Fwd	GTGTTTGCTGGTGCTACTGTCTGA
AhHMA4-3 promoter_Rev	TTTCTCTTCTTCTTTGTTTTGTGACGCC

4.5. RESULTS

4.5.1. Sequence analysis of HMA4 promoters in Noccaea caerulescens

DNA parsimony of the initial 2000 bp promoter sequence from the transcriptional start site of HMA4 orthologous sequences from A. halleri, A. thaliana and N. caerulescens demonstrated closer sequence identities between both Arabidopsis genomes than with N. caerulescens Saint Laurent Le Minier (DNApars software at default settings, Phylip version 3.68) (Fig. 4.2 A). Two groups, one containing promoter sequences from AhHMA4-3 and AhHMA4-2 and one containing those of AhHMA4-1 and AtHMA4 were formed within this Arabidopsis spp. clade (Fig. 4.2 A). All N. caerulescens tandem gene promoter sequences were grouped together, and shared low sequence identity (< 51%), with all orthologous sequences from the Arabidopsis spp (Dot Matrix at default settings stringency of 30%, Vector NTI 11) (Fig. 4.2 Aand B). AhHMA4-1 promoter revealed a closer phylogentic relationship with the AtHMA4 promoter (88%) than with either AhHMA4-2 or AhHMA4-3 promoters (57%) (Fig 4.2 A and B). Inter-species analysis revealed that the AtHMA4 promoter sequence shared its lowest identity with AhHMA4-3 (53%), which shared its lowest sequence identity with the NcHMA4-2 promoter region (40%). In order to investigate the functional effects of variation in cis sequences a transgenic in planta functional assay was attempted.

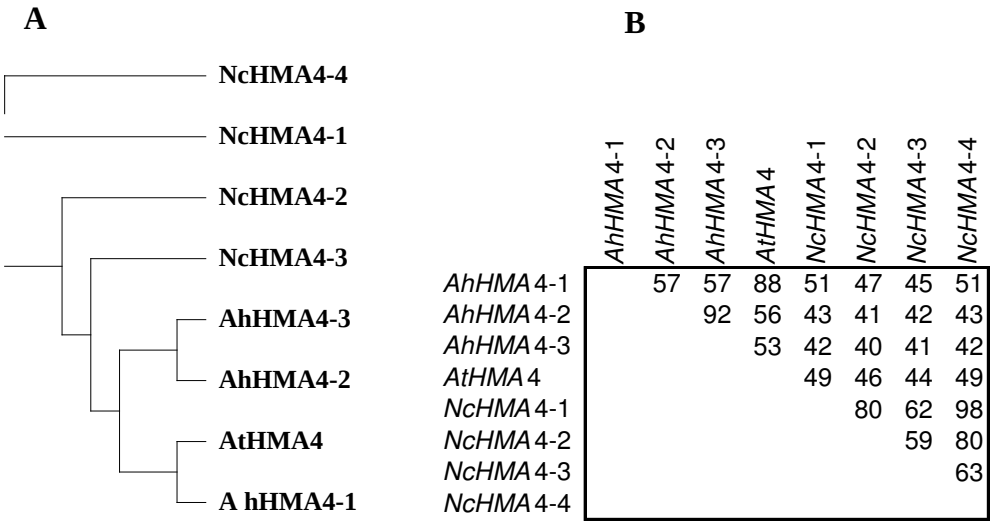


Figure 4.2 A) Cladogram and B) Dot Matrix comparisons of promoter sequences 2000 bp upstream from the transcriptional start site of HMA4 orthologues from Ah: Arabidopsis halleri, At: Arabidopsis thaliana and Nc: Noccaea caerulescens. Tandem repeats are differentiated by "-". The cladogram was created for nucleotide sequences by the DNA Sequence Parsimony Method (DNApars), and was run at default settings in Phylip version 3.68. The Dot Matrix program was run at default settings and supplied by Vector NTI 11. Numbers represent percentage sequence identities.

4.5.2. Functional analysis of promoter sequences

Primers were designed to amplify regions upstream of the transcriptional start site of HMA4 orthologues AhHMA4-1, AhHMA4-3, NcHMA4-2 and AtHMA4. For both A. thaliana and A. halleri, primers employed were identical to those used previously by Hanikenne et al. (2008) (Section 4.4.4.). Sequence data for A. halleri was isolated from bacterial artificial chromosome (BAC) clones deposited in GenBank (GenBank: EU382072.1 and EU382073.1). Sequence information for the NcHMA4-2 promoter was obtained from fragmented sequence data from the Noccaea caerulescens Saint Laurent Le Minier fosmid library through an initial pooled GS FLX 454 sequencing attempt (Chapter 3). For A. thaliana primers were designed to amplify a 2014 bp region upstream from the transcriptional start site of AtHMA4. For A. halleri, primers for the AhHMA4-1 promoter were designed to amplify a 2330 bp region upstream of the transcriptional start site. Primers for the AhHMA4-3 promoter were designed to amplify a region of 1352 bp upstream from the transcriptional start site.

For N. caerulescens the internet package Primer 3 Version 0.4.0 (Section 2.7.1) was employed to design primers to amplify a 3276 bp region upstream from the transcriptional start site using default settings but with both "Max Self Complementarity" and "Max 3' Self Complementarity" settings adjusted to 2, to reduce the likelihood of hairpin loops and dimerisation. Genomic DNA was extracted from leaf tissue (Section 2.6.1.) from Noccaea caerulescens (J and C. Presl.) FK Mey Saint Laurent Le Minier, A. thaliana L. (Heynh) Col-0 and A. halleri (L. O'Kane and Al-Shehbaz subsp. halleri). HMA4 orthologous promoter specific PCR fragments were amplified using HMA4 promoter primers (Section 4.4.4.) and a proofreading polymerase (Phusion, Fisher Scientific, UK) under the conditions described (Section 2.7.3.). Fragments were purified by gel electrophoresis and isolated by gel extraction techniques (Section 2.6.3.) for subsequent cloning.

4.5.3. Cloning the AtHMA4 promoter sequence

Following DNA gel purification and isolation, the 2014 bp fragment was adapted for downstream cloning into the pCR8®/GW/TOPO® TA Cloning® system (Invitrogen, Paisley, UK). This involved adding an adenine base to the 3' end of each strand of the proofread PCR product (A-tailing) (Section 2.5.6.1.). A 2 µl aliquot of the pCR8® compatible AtHMA4 promoter fragment was ligated into a pCR8® entry vector as described (Section 2.5.6.2.). Transformed E. coli cells were selected on solid LB media supplemented with 50 µg ml^{-1} spectinomycin overnight at 37°C. Colony PCR with AtHMA4 promoter primers tested for presence of the promoter in the bacterial plasmid (Section 2.5.6.4.) (Fig. 4.3). The positive "Genomic" control contained A. thaliana genomic DNA as a template (Fig. 4.3). Three colonies produced amplicons of the correct 2014 bp size were grown overnight in conical flasks containing 50 ml of Luria-Bertani (LB) broth, supplemented with 50 µg ml^{-1} spectinomycin, at 37°C on a horizontal shaker (250 – 300 rpm). Plasmids were isolated from bacterial cultures using the Qiaprep Miniprep kit (Qiagen, Crawley, UK) (Section 2.5.7.) and inserts sequenced using the Value Read, Single Read Service by Sanger sequencing with the ABI 3730 XL capillary sequencer using BigDye v.3.1 dye-terminator chemistry (Eurofins MWG GmbH, Ebersberg, Germany) (Section 2.8.1.).

Glycerol stocks were made for all three confirmed E. coli DH5α cells harbouring the AtHMA4 promoter sequence.

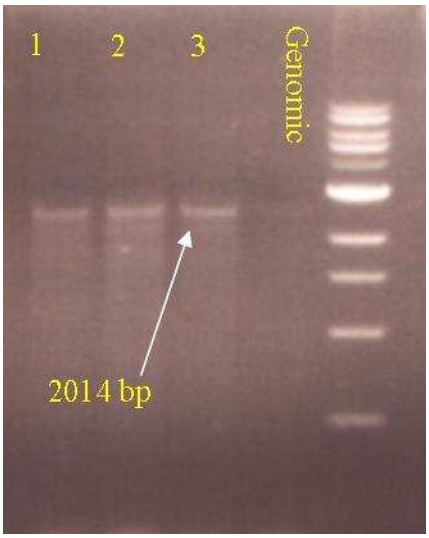


Figure 4.3 Single colony PCR on Escherichia coli DH5α cells transformed with the PCR8® entry vector (Invitrogen) containing 2014 bp promoter region upstream from the ATG transcriptional start site of A. thaliana HMA4.Lane labelled 'Genomic' contain PCR amplified A. thaliana genomic DNA. Arrow represents bp according to the 1 Kb DNA ladder (New England Biolabs). Gel contained 1% (w/v) agarose.

4.5.4. Cloning the AhHMA4-1 and AhHMA4-3 promoter sequences

Both the 2330 bp AhHMA4-1 and 1352 bp AhHMA4-3 promoter sequences were purified and isolated by gel electrophoresis and A-tail adaption for downstream cloning into the pCR8®/GW/TOPO® TA Cloning® system (Invitrogen, Paisley, UK) (Section 2.5.6.1.). A 2 µl aliquot of the pCR8® compatible AhHMA4-1 and AhHMA4-3 promoter fragments were individually ligated into a pCR8® entry vector as described in Section 2.5.6.2. Transformed E. coli colonies were selected on solid LB media and subjected to colony PCR with either AhHMA4-1 or AhHMA4-3 specific promoter primers to test for their presence in the bacterial plasmid (Section 2.5.6.4.) (Fig. 4.4). The positive "Genomic" control contained A. halleri genomic DNA as a template (Fig. 4.4). Ten colonies containing AhHMA4-1p and two containing AhHMA4-3p were grown overnight in 50 ml of LB, supplemented with 50 µg ml^{-1} spectinomycin and following plasmid isolation (Section 2.5.7.), inserts were sequenced using the Value Read, Single Read Service by Sanger sequencing (Section 2.8.1.). Glycerol stocks were made

for all confirmed E. coli DH5α cells bearing the both AhHMA4-1 and AhHMA4-3 promoter sequences.

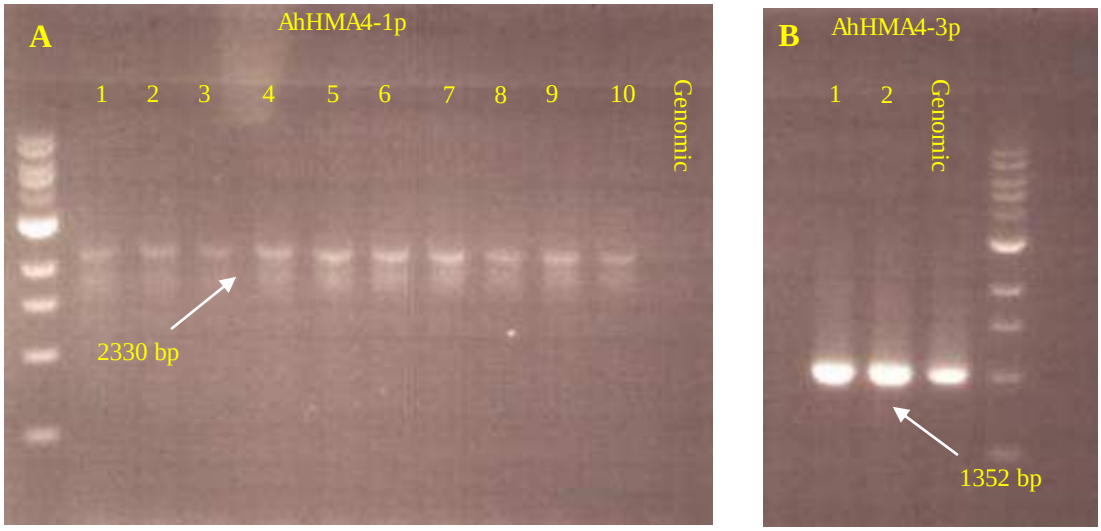


Figure 4.4 Single colony PCR on Escherichia coli DH5α cells transformed with the PCR8® entry vector (Invitrogen) containing **A;** 2330 bp promoter region upstream from the ATG transcriptional start site of A. halleri HMA4-1 and, **B;** 1352 bp promoter region upstream from the ATG transcription start site of AhHMA4-3. Lanes labelled 'Genomic' contain PCR amplified A. halleri genomic DNA. Arrows represent bp according to the 1 Kb DNA ladder (New England Biolabs). Gel contained 1% (w/v) agarose.

4.5.5. Cloning the NcHMA4-2 promoter sequence

The NcHMA4-2 promoter sequence (3276 bp) was purified and isolated by gel electrophoresis and A-tail adapted for downstream cloning into the pCR8®/GW/TOPO® TA Cloning® system (Section 2.5.6.2.). A 2 µl aliquot of this product was ligated into a pCR8® entry vector and transformed into competent E. coli DH5α cells. Following overnight selection on solid LB media, colony PCR (Section 2.5.6.4.) identified seven E. coli DH5α clones bearing the NcHMA4-2 promoter using NcHMA4-2 promoter specific primers (Section 4.4.4.) (Fig. 4.5). The positive "Genomic" control contained N. caerulescens genomic DNA as a template (Fig. 4.5). All seven transformed colonies' plasmids were extracted as described and sequenced using the Sanger-based Value Read, Single Read Service (Section 2.8.1). Glycerol stocks were made for all confirmed E. coli DH5α cells bearing the NcHMA4-2 promoter sequence.

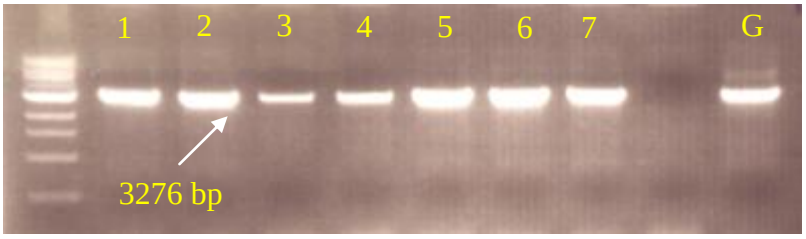


Figure 4.5 Single colony PCR on Escherichia coli DH5α cells transformed with the PCR8® entry vector (Invitrogen) containing 3276 bp promoter region upstream from the ATG transcription start site of N. caerulescens HMA4-2. Lane labelled 'G' contained PCR amplified N. caerulescens genomic DNA. Arrows represent bp according to the 1 Kb DNA ladder (New England Biolabs). Gel contained 1% agarose.

4.5.6. Creating an AtHMA4 promoter fused with the *β*-glucuronidase (GUS) reporter gene construct in Agrobacterium tumefaciens GV3101

AtHMA4 promoter inserts of 2014 bp previously cloned into pCR8® entry vectors were transferred into pGWB3 destination vectors which were compatible with the Gateway cloning system (Section 4.4.1.). The pGWB3 vector contained both the HPT and the NPTII genes conferring both hygromycin and kanamycin resistance respectively to plants (Section 4.4.1.) (Nakagawa et al., 2007). Equimolar concentrations of both pCR8® and pGWB3 vectors (150 ng) were mixed and incubated as described in Section 4.4.1. and 4 µl aliquots were employed to transform E. coli using chemical and heat shock treatments (Section 2.5.6.3). Following overnight selection on solid LB media supplemented with and 50 µg ml^{-1} kanamycin, and colony PCR (Section 2.5.6.4.) clones bearing the AtHMA4 promoter were identified using AtHMA4 promoter specific primers (Section 4.4.4.). Presence of the NPTII and UIDA (GUS) genes were also confirmed through gene specific primers (Section 2.10.).

Plasmids (AtHMA4p::GUS) were extracted and sequenced (Section 2.8.1.) which confirmed the presence and correct orientation of the promoter sequence (Fig. 4.6).

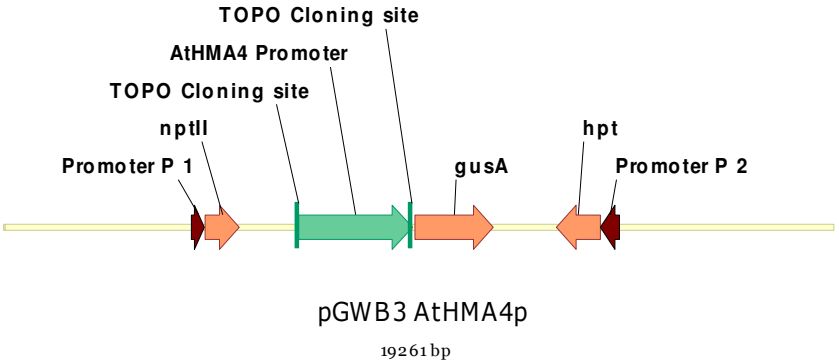


Figure 4.6 Linear illustration of the pGWB3 plasmid containing the AtHMA4 promoter (AtHMA4p::GUS construct). Yellow bar represented the entire 19261 bp pGWB3 AtHMA4p plasmid. Green arrow depicted the genomic direction of the 2014 bp AtHMA4 promoter insert which regulated the transcription of the UIDA (GUS) gene. Orange arrows represented the transcriptional direction of genes specific to the pGWB3 vector, NPTII, GUS and HPT. Brown arrows represented promoters P1 and P2 which regulated NPTII and HPT transcription respectively. The image was created through Vector NTI 11 (Invitrogen, Paisley, UK).

Extracted AtHMA4p::GUS plasmids were subsequently transformed into Agrobacterium tumefaciens strain GV3101 to facilitate future transgene integration assays into plant genomes. An aliquot of 2 µl (~250 ng) of plasmid DNA was mixed with 50 µl of freshly thawed untransformed competent A. tumefaciens GV3101 cells. Bacteria was transformed by electroporation as described (Section 2.5.6.5.) and selected for 5 d on solid LB supplemented with kanamycin, gentamycin, hygromycin and rifampicin (50 µg ml^{-1}) and tetracycline (5 µg ml^{-1}). Colony PCR was performed using primers specific for the AtHMA4 promoter, GUS and NPTII sequences to verify plasmid integrity in the A. tumefaciens transformed cells (Fig. 4.7). The positive "Genomic" control contained A. thaliana genomic DNA as a template (Fig. 4.7). Two colonies produced amplicons for GUS (1029 bp), NPTII (690 bp) and the AtHMA4 promoter (2014 bp) and were therefore grown and their plasmids were isolated for sequencing by Sanger Value Read, Single Read Service (Eurofins MWG GmbH, Ebersberg, Germany) (Section 2.8.1.). Upon sequence confirmation of correct insert orientation, glycerol stocks were made.

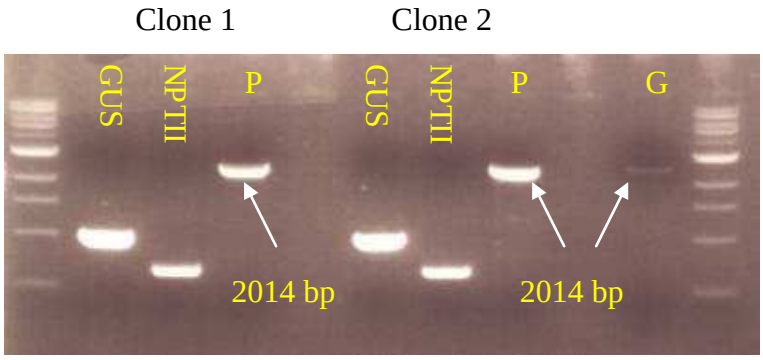


Figure 4.7 Single colony PCR on Agrobacterium tumefaciens GV3101 cells transformed with the pGWB3 destination vector containing a 2014 bp promoter region upstream from the ATG transcriptional start site of A. thaliana HMA4 (AtHMA4p::GUS constructs). Constructs were also tested for the presence of GUS and NPTII. Lane labelled 'G' contained PCR amplified A. thaliana genomic DNA. Arrows represented bp according to the 1 Kb DNA ladder (New England Biolabs). Gel contained 1% (w/v) agarose.

4.5.7. Creating an AhHMA4-1 and AhHMA4-3 promoter fused with *β*-glucuronidase (GUS) reporter gene constructs in Agrobacterium tumefaciens strain GV3101

Both AhHMA4-1 and AhHMA4-3 promoter sequences contained in pCR8® entry vectors (Section 4.5.4.) were transferred into the plant expression vector pGWB3 via the Gateway mediated LR cloning system (Section 4.4.1.). Following quantification of DNA using Nanodrop (Section 2.5.7.), 50 µl of E. coli DH5α were transformed with 4 µl aliquots from each of these solutions as described in Section 4.5.4. Colony PCR for AhHMA4-1 and AhHMA4-3 promoters as well as NPTII and UIDA (GUS) genes on overnight cultures selected on solid LB media supplemented with 50 µg ml^{-1} kanamycin (Section 2.5.6.4.) verified the introduction of plasmids into the E. coli DH5α cells. Extracted AhHMA4-1p::GUS and AhHMA4-3p::GUS plasmids were sequenced and their respective promoters were confirmed in the correct orientation (Fig. 4.8).

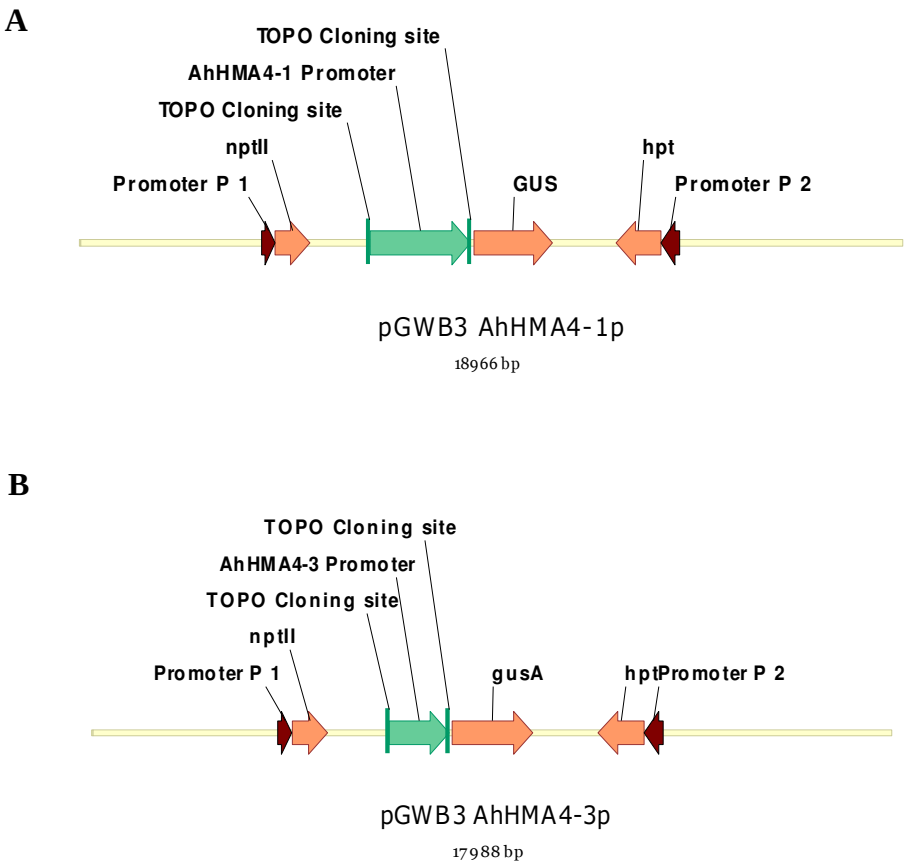


Figure 4.8 Linear illustrations of the pGWB3 plasmids containing either (A) the AhHMA4-1 promoter or (B) the AhHMA4-3 promoter. Yellow bars represented the entire (A) 18966 bp pGWB3 AhHMA4-1p (AhHMA4-1p::GUS) or (B) 17988 bp pGWB3 AhHMA4-3p (AhHMA4-3p::GUS) plasmids. Green arrows depicted the genomic direction of the (A) 2330 bp AhHMA4-1 and (B) 1352 bp AhHMA4-3 promoters which regulated the transcription of the UIDA (GUS) genes. Orange arrows represented the transcriptional direction of genes specific to the pGWB3 vector, NPTII, GUS and HPT. Brown arrows represented promoters P1 and P2 which regulated NPTII and HPT transcription respectively. The images were created through Vector NTI 11 (Invitrogen, Paisley, UK).

Plant expression vectors pGWB3 containing either AhHMA4-1 promoters (AhHMA4-1p::GUS) or AhHMA4-3 promoters (AhHMA4-3p::GUS) were transformed into Agrobacterium tumefaciens strain GV3101 for downstream plant based functional molecular analyses. Aliquots of 2 µl (~250 ng) of plasmid DNA was introduced into bacterial cells using electroporation techniques as described in Section 4.5.6. Following selection on solid LB medium for 5 d, colonies were subjected to PCR using primers specific to AhHMA4-1 or AhHMA4-3 promoters and GUS and NPTII genes. A. halleri genomic DNA was employed as a positive control "G" (Fig. 4.9 A and B). For each construct, two colonies produced amplicons for GUS (1029 bp), NPTII (690 bp) and the AhHMA4-1 (2330 bp) or AhHMA4-3 (1352 bp) promoters (Fig 4.7 A and B).

Colonies were grown, plasmids isolated and sequenced and all clones were shown to contain respective promoter sequences in correct orientations.

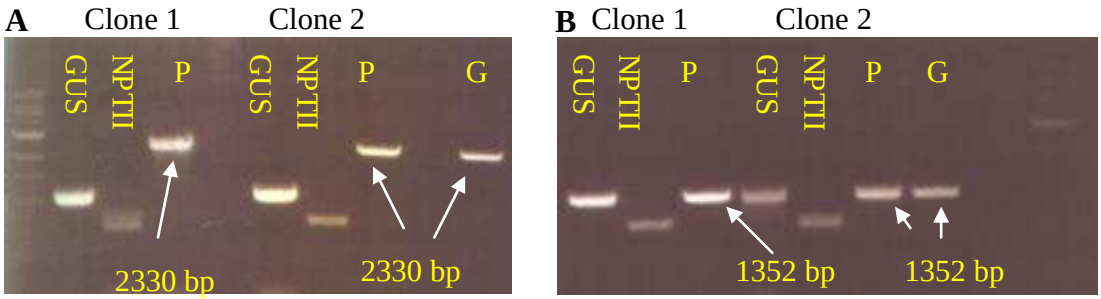


Figure 4.9 Single colony PCR on Agrobacterium tumefaciens GV3101 cells transformed with the pGWB3 destination vector containing **A;** 2330 bp promoter region upstream from the ATG transcriptional start site of A. halleri HMA4-1 (AhHMA4-1p::GUS) and, **B;** 1352 bp promoter region upstream from the ATG transcription start site of AhHMA4-3 (AhHMA4-3p::GUS). Lanes labelled 'Genomic' contained PCR amplified A. halleri genomic DNA. Arrows represented bp according to the 1 Kb DNA ladder (New England Biolabs). Gel contained 1% (w/v) agarose.

4.5.8. Creating a NcHMA4-2 promoter fused with the *β*-glucuronidase (GUS) reporter gene construct in Agrobacterium tumefaciens strain GV3101

Promoter DNA from NcHMA4-2 cloned into the pCR8® entry vector (Section 4.5.5.) was employed to create pGWB3 expression vectors of NcHMA4-2 fused to the UIDA (GUS) gene through the Gateway mediated LR cloning system (Section 4.4.1.). Equimolar vector concentrations were added to the cloning reaction and 4 µl aliquots were used to transform E. coli DH5α cells as described in Section 4.5.6. Following selection, colony PCR for NcHMA4-2 promoter and NPTII and UIDA (GUS) genes and sequencing of these NcHMA4-2p::GUS constructs verified transformation and NcHMA4-2 promoter orientation (Section 4.5.3.) (Fig. 4.10).

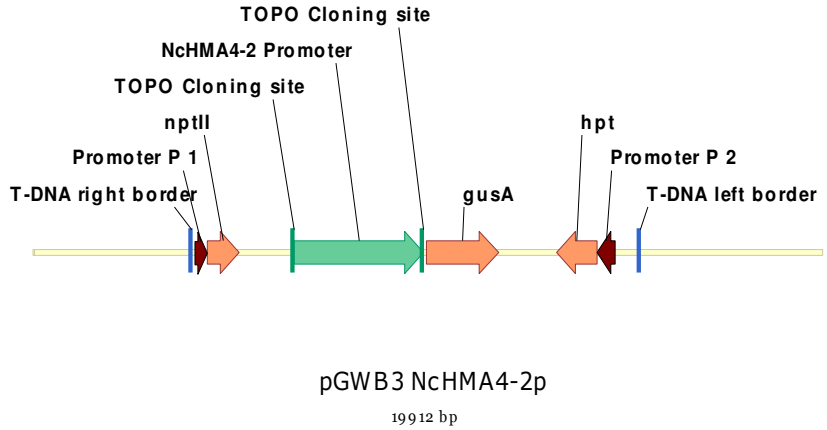


Figure 4.10 Linear illustration of the pGWB3 plasmid containing the NcHMA4-2 promoter. Yellow bar represented the entire 19261 bp pGWB3 NcHMA4-2p plasmid (NcHMA4-2p::GUS). Green arrow depicted the genomic direction of the 3276 bp NcHMA4-2 promoter insert which regulated the transcription of the UIDA (GUS) gene. Orange arrows represented the transcriptional direction of genes specific to the pGWB3 vector, NPTII, GUS and HPT. Brown arrows represented promoters P1 and P2 which regulated NPTII and HPT transcription respectively. The image was created through Vector NTI 11 (Invitrogen, Paisley, UK).

The pGWB3 plant expression vector containing the NcHMA4-2 promoter (NcHMA4-2p::GUS) was introduced into the Agrobacterium tumefaciens strain GV3101 for downstream functional molecular analyses in planta. Aliquots of 2 µl (~250 ng) of plasmid DNA was introduced into bacterial cells using electroporation techniques as described in Section 4.5.6. Colony PCR was performed with N. caerulescens genomic DNA as a positive control "G" (Fig. 4.11). Two colonies produced amplicons for GUS (1029 bp), NPTII (690 bp), however only one produced a NcHMA4-2 promoter amplicon (3276 bp) (Fig 4.11). This plasmid was isolated and sequenced and the promoter was sequence confirmed to be correctly orientated.

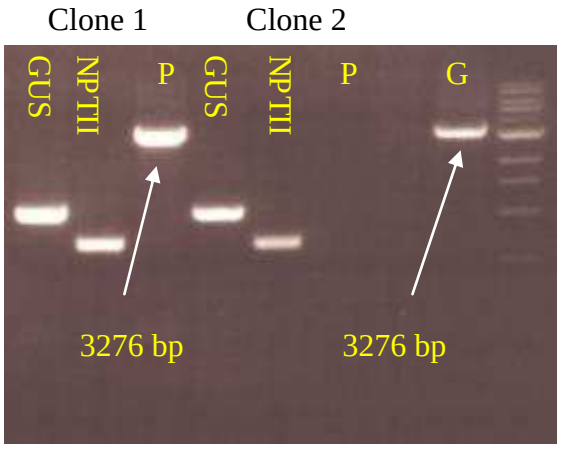


Figure 4.11 Single colony PCR on Agrobacterium tumefaciens GV3101 cells transformed with the pGWB3 destination vector containing a 3276 bp promoter region upstream from the ATG transcription start site of N. caerulescens HMA4-2 (Clone 1) (NcHMA4-2p::GUS construct). Clone 2 did not contain NcHMA4-2p. Lane labelled 'G' contained PCR amplified N. caerulescens genomic DNA. Arrows represented bp according to the 1 Kb DNA ladder (New England Biolabs). Gel contained 1% (w/v) agarose.

4.5.9. Transforming Arabidopsis thaliana

Arabidopsis thaliana Col-0 were grown for three weeks under glasshouse conditions (22.3°C and 13.3°C mean day and night temperatures respectively) at a 16 hr photoperiod, in 0.32 l pots containing Levington M_3 compost (Section 2.3). Plants were considered optimal for transformation when there was an abundance of immature flower clusters and relatively few fertilized siliques. Agrobacterium tumefaciens strain GV3101, harbouring the constructs AtHMA4p::GUS (Section 4.5.6.), AhHMA4-3p::GUS (Section 4.5.7.) or NcHMA4-2p::GUS (Section 4.5.8.) was selected to transform these sequences into the A. thaliana genome. A single colony of A. tumefaciens harbouring these constructs was grown under antibiotic selection in a flask of liquid LB on a horizontal shaker (250 – 300 rpm) for 16 h at 28 °C in darkened conditions (Section 2.5.4.). Bacterial growth was examined through spectrophotometry until an optical density of 0.8 was recorded at 600 nm light wave lengths, whereby cells were centrifuged as described (Section 2.9.) and resuspended in a solution of H_2O and sucrose (50 g l^{-1}). Before transformation was attempted, the surfactant (Silwet L-77) was added at a concentration of 0.05% (v/v). Five pots containing plants were transformed per construct. All inflorescence were placed into 200 ml of bacterial solution for 5 – 6 sec, removed and placed into closed

plastic sleeves for 24 hrs (Section 2.9.). Three weeks following this floral dip event, seeds were harvested and stored for downstream analysis.

4.5.10. Selecting primary T_1 Arabidopsis thaliana transformants

All seeds returned from plants subjected to Agrobacterium tumefaciens mediated transformation were selected in vitro. Seeds were surface sterilised (Section 2.9.1.) and sown by evenly distributing approximately 30 - 40 onto Petri dishes containing 20 ml growth media (agar 8 g l^{-1}, sucrose 10 g l^{-1} and 2.1 g l^{-1} MS salts) supplemented with kanamycin sulphate (50 µg ml^{-1}). Plants surviving under these conditions were transplanted to ex vitro conditions in the glasshouse (22.3°C and 13.3°C mean day and night temperatures respectively; 16 h photoperiod) (Section 2.3.). Plant genomic DNA was extracted and subjected to PCR analysis using primers specific for the NPTII and UIDA genes. Approximately 0.5% of plants screened were shown to be transformed with the construct of interest. These plants were maintained under glasshouse conditions and selfed to create segregating populations of 1:2:1 (homozygous for the insert : heterozygous for the insert : homozygous wild type) plants. Three independent T_1 A. thaliana transformants were identified bearing the AtHMA4p::GUS construct, five bearing the AhHMA4-3p::GUS construct, one bearing the NcHMA4-1p::GUS construct and four bearing the NcHMA4-2p::GUS construct. For all, three lines were selected for further functional promoter analysis.

4.5.11. Promoter functional analysis

All T_2 and one line of wild type A. thaliana Col-0 plants were initially sown in polycarbonate growth boxes in 75 ml of growth media containing 10 g l^{-1} sucrose, 8 g l^{-1} agar and 2.1 g l^{-1} MS growth minerals equivalent to half the strength of the original recipe published by Murashige and Skoog (1962) (Section 2.4.2.). This supplied 15 µM Zn to plants in the form $ZnSO_4$. Growth media was supplemented with 50 µg ml^{-1} kanamycin sulphate as a selective agent for both transformed A. thaliana plants and three replicates of WT plants. A further three growth boxes containing WT plants were not supplemented with kanamycin sulphate. All plants were grown in growth boxes randomly distributed in three blocks at 21 °C under a 16 hr photoperiod supplying a light intensity of 50 – 80 µmol m^{-2} s^{-1}. Seven d after sowing (DAS), plants displaying

healthy green leaves were selected as T_2 transformants and transplanted under axenic conditions in an airflow cabinet onto fresh growth medium (75 ml) containing no kanamycin sulphate supplementation. For all transformed lines a 3 to 1 ratio of surviving plants to dying plants was observed indicating the insertion of a single construct into the genome. All wild type *A. thaliana* plants used as negative controls died under kanamycin sulphate supplementation. Three plants were transplanted into a total of three boxes for each line. A randomised block design comprising three replicates was employed, with three independent transformant lines for each of the three promoter constructs and one wild type (WT) line allocated at random within each replicate (N=30). All boxes contained a single concentration of 15 μM $ZnSO_4$. Plants were grown under identical growth conditions as previous (16 hr photoperiod, 21°C) for a further 14 days. 21 d old plants were then removed from these conditions and processed for histochemical analysis.

4.5.12. Histochemical GUS analysis

Plants were harvested from growth boxes 21 DAS and were subjected to histochemical detection of GUS activity as described by Jefferson et al. (1987) (Section 4.4.3.). All plants growing in each growth box were selected for the histochemical GUS assay, therefore 30 independent assays, each containing 3 plants were performed. All plants were incubated in darkened conditions for 16 hrs in a solution of 0.5 g l^{-1} 5-bromo-4-chloro-3-indolyl β-D-glucopyranoside. (Section 4.4.3.). To clearly visualise the blue product, chlorophyll was removed from leaves by stepwise replacement of the staining solution with ethanol before mounting in solutions of 50% glycerol. All plants were observed under a stereomicroscope for distribution patterns of blue product throughout leaves, stems and shoots. Images were taken of single representative samples among all transformants for each construct (Fig. 4.12).

As a negative control a representative WT plant subjected to the histochemical assay was photographed and demonstrated no blue product in leaves, stems or roots. Lines bearing the AtHMA4p::GUS construct contained a significant proportion of blue staining in root and some stem tissue, however no staining was observed in leaf tissues. This indicated that the UIDA gene expression,

under the control of the AtHMA4 promoter was restricted to roots and some stem tissues. For both NcHMA4-2p::GUS and AhHMA4-3p::GUS constructs, transformed lines displayed blue staining in most plant tissue including roots, shoots and stems. The UIDA gene appeared to be expressed throughout the plant when driven by either the NcHMA4-2 or the AhHMA4-3 promoters (Fig. 4.12).

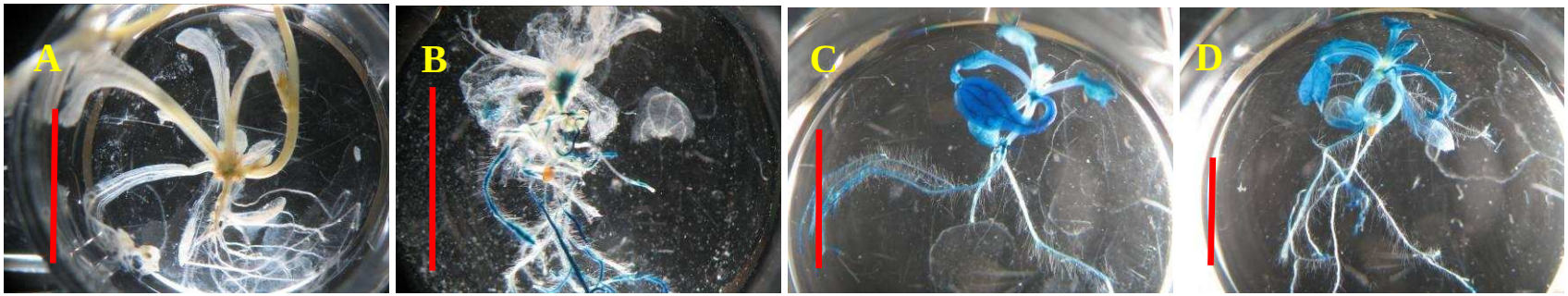


Figure 4.12 The activity of *β*-glucuronidase (GUS) in whole Arabidopsis thaliana plants. **A;** wild type, untransformed control, then T_2 transformants bearing pGWB3 constructs containing the GUS marker gene under the control of promoter sequences from **B;** AtHMA4, **C;** AhHMA4-3 and **D;** NcHMA4-2. Transgenic plants were first selected on agar (0.8 g l^{-1}) supplemented with 50 µg ml^{-1} of kanamycin sulphate for 7 days before being transferred to fresh agar (0.8 g l^{-1}) supplemented with 10 g l^{-1} sucrose and half strength MS salts for 21 days. Plants were incubated in a staining solution of 0.5 g l^{-1} 5-bromo-4-chloro-3-indolyl β-D-glucopyranoside and incubated overnight at 37 °C. To clearly visualise the blue product, chlorophyll was removed from leaves by stepwise replacement of the staining solution with ethanol before mounting in solutions of 50% glycerol. Red bars represent 1 cm.

Upon closer examination of GUS staining in each line it was observed that blue staining appeared more intense in the roots of transformants harbouring the UIDA gene driven by the promoters of either AhHMA4-3 or NcHMA4-2 than those driven by the AtHMA4 promoter (Fig. 4.13). This suggested that the UIDA genes under the control of either NcHMA4-2 or AhHMA4-3 promoters were more highly expressed in the roots than those under the control of the AtHMA4 promoter under these in vitro mineral replete conditions.

Spatially, GUS activity appeared to localise more strongly to the root vasculature when UIDA was under the control of either AtHMA4 or NcHMA4-2 promoters. Lines containing the AhHMA4-3p::GUS construct revealed GUS staining throughout the root tissue suggesting that UIDA expression was not restricted to the root vasculature when driven by the AhHMA4-3 promoter.

Staining was not observed in the leaves of plants bearing the UIDA gene under the control of the AtHMA4 promoter. Conversely leaves of plants containing both AhHMA4-3p::GUS and NcHMA4-2p::GUS constructs demonstrated blue staining throughout leaf tissue. This indicated that under the control of either AhHMA4-3p or NcHMA4-2p, UIDA expression was not restricted to shoot vasculature but was also represented in leaf trichomes and parenchyma cells (Fig 4.13).

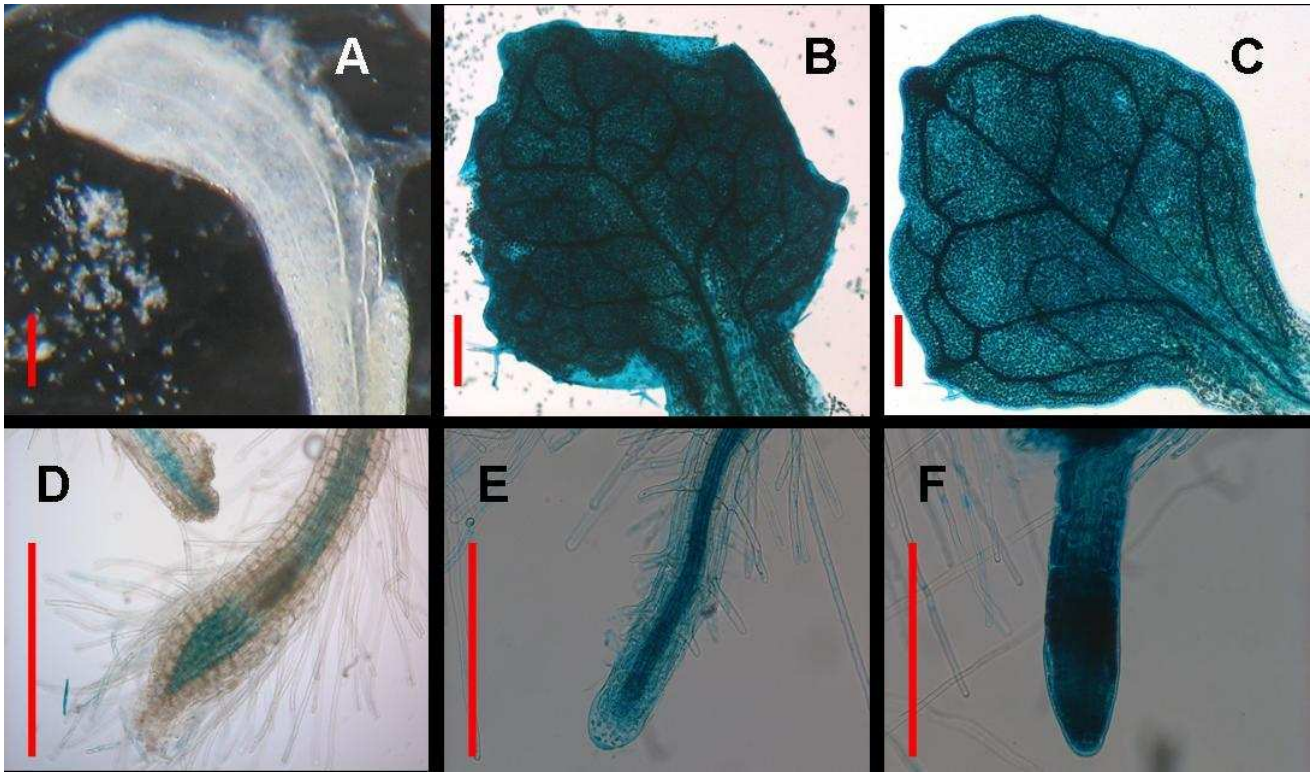


Figure 4.13 The activity of *β*-glucuronidase (GUS) in whole leaves A - C and roots D – F from Arabidopsis thaliana T_2 transformants bearing pGWB3 constructs containing the GUS marker gene under the control of promoter sequences from **A, D;** AtHMA4, **B, E;** NcHMA4 and **C, F;** AhHMA4-3. Transgenic plants were first selected on agar (0.8 g l^{-1}) supplemented with 50 μg ml^{-1} of kanamycin sulphate for 7 days before being transferred to fresh agar (0.8 g l^{-1}) supplemented with 10 g l^{-1} sucrose and half strength MS salts for 21 days. Plants were incubated in a staining solution of 0.5 g l^{-1} 5-bromo-4-chloro-3-indolyl β-D-glucopyranoside and incubated overnight at 37 °C. To clearly visualise the blue product, chlorophyll was removed from leaves by stepwise replacement of the staining solution with ethanol before mounting in solutions of 50% glycerol. Red bars represent 2 mm.

4.6. DISCUSSION

In this chapter heterologous expression of HMA4 promoters from Arabidopsis thaliana, A. halleri and Noccaea caerulescens were compared in A. thaliana T_2 transformants. Promoter activity from both hyperaccumulators showed remarkable similarities but were remarkably distinctive from that of A. thaliana. Moreover, in contrast to the shared high protein sequence identities between all three genomes (Chapter 3), N. caerulescens HMA4 promoters (NcHMA4p) exhibited low nucleotide identities (40 - 51%) with promoters from both A. thaliana and A. halleri. Despite this, both NcHMA4-2p and AhHMA4-3p exhibited generally higher expression throughout the plant than AtHMA4p which supported previous reports of higher HMA4 expression in N. caerulescens (Bernard et al., 2004; Papoyan and Kochian, 2004; Hammond et al., 2006) and A. halleri (Talke et al., 2006; Hanikenne et al., 2008) than in A. thaliana. Throughout leaf tissue, both hyperaccumulator promoters demonstrated enhanced activity compared to AtHMA4p, which was consistent with their putative roles in Zn exclusion from certain cell types (Papoyan and Kochian, 2004; Hanikenne et al., 2008) and associated with reported high Zn concentrations in leaf epidermal and mesophyll cell walls in N. caerulescens (Frey et al., 2000; Ma et al., 2005) and mesophyll vacuoles in A. halleri (Küpper et al., 2000). It has been recently demonstrated in A. halleri that HMA4 expression levels were significantly higher due to cis-regulatory changes and gene triplication (Hanikenne et al., 2008). It is therefore conceivable that novel cis regulatory elements in N. caerulescens, could contribute to increased NcHMA4 gene expression. Therefore elucidating these cis regulatory regions could enable the manipulation of HMA4 expression which may be further developed and exploited for use within other plant systems to enhance Zn leaf accumulation.

4.6.1. Noccaea caerulescens promoter regions are significantly diverged from Arabidopsis thaliana and Arabidopsis halleri.

Sequence divergence between A. thaliana, A. halleri and N. caerulescens was found to be most pronounced in the 5' flanking regions of HMA4 sequences. Similar observations were recorded by Hanikenne et al. (2008) when comparing cis sequences between A. thaliana and A. halleri. Dot Matrix sequence comparisons (Vector NTI, stringency > 30%) for all HMA4 promoter regions for all three genomes illustrated that the four NcHMA4 promoter sequences displayed between 44

– 49% sequence identity with AtHMA4 and between 40 – 51% sequence identity with the three A. halleri HMA4 promoter sequences. Interestingly both promoter sequences from AhHMA4-2 and AhHMA4-3 shared 99% sequence identity but only 57% with that of the AhHMA4-1 promoter, which demonstrated the highest inter specific sequence similarity (88%) with the AtHMA4 promoter. Within N. caerulescens, promoter sequences from NcHMA4-1, NcHMA4-2 and NcHMA4-4 shared between 80 – 98% sequence identity but just 62, 59 and 63% respectively with the NcHMA4-3 promoter. For both A. halleri and N. caerulescens, HMA promoter regions exhibiting these relatively high levels of sequence identity suggests gene replication could have occurred relatively recently within these genomes. However, in contrast with the high degree of protein and cDNA conservation between these genomes (Chapter 3), promoter sequences illustrate vast heterogeneity. Understanding the regulatory control on gene expression within this heterogeneity could help elucidate promoter specific roles in Zn transport within N. caerulescens. Comparative analysis of these promoter sequences in planta under similar environmental conditions is an approach which may help to further understanding of these mechanisms which have independently evolved in plants. Therefore, an in planta promoter assay was initiated to compare cis-induced expression profiles of HMA4 promoters from A. thaliana, A. halleri and N. caerulescens.

4.6.2. Developing an efficient heterologous promoter expression system

Promoter expression profiling through promoter::GUS fusion experiments remains one of the most widely applied methods to qualification and quantification of promoter regulation. In order to gain an understanding of the function and spatial patterning of endogenous promoters, sequences are fused to marker genes and either stably introgressed into the plant genome, through transgenic recombination technology or alternatively through transient transformation techniques (Agius et al., 2005). While uniformity and heritable stable transformation events are desired, regeneration of transformed cells from recalcitrant species remains a limiting factor. Due to its unstable nature and limitations associated with genomic integration, transient expression is applied primarily to verify function of transgenic constructs and expression analyses within slow generating plants or recalcitrant fruits such as strawberry (Agius et al., 2005). Noccaea caerulescens has proved immensely

challenging to stably transform with just one reported successful floral dip based approach (Peer et al., 2003) which has thus far never been repeated. More recently Guan et al. (2008) reported a tissue culture based transformation approach however this method is unlikely to favour a system for in planta molecular analyses since tissue culture is slow and can introduce epigenetic effects and somaclonal variation (Vazquez and Linacero, 2010). Recently a stable transformation protocol had been developed for A. halleri, (Hanikenne et al., 2008), however this system too, relies on inefficient tissue culture techniques. In addition, A. halleri is an obligate out-breeder, and so remains unsuitable for efficient generation of homozygous transgenic plants.

Heterologous transformation events within yeast have helped elucidate gene function and localisation at a unicellular level however, results often vary according to the strain of yeast employed (Mills et al., 2005; Verret et al., 2005). Coupled with this, spatial expression patterns, tissue specific regulation, and complex protein interactions in response to exogenous metal concentration cannot be determined in unicellular organisms (Siemianowski et al., 2010). Unlike N. caerulescens and A. halleri, A. thaliana is rapid cycling and can be relatively simply and extremely efficiently transformed using the floral dip approach (Clough and Bent, 1998). For these reasons it is routinely employed for heterologous promoter expression studies. Tittarelli et al., (2007) described a comparative analysis of the promoter activity of the wheat phosphate transporter TaPT2 promoter fused to GUS which was transformed into both wheat and A. thaliana. Under analogous growth conditions the GUS reporter was expressed at similar levels for both species. The authors concluded that conserved cis-acting elements and trans-acting factors enabled the promoter to be regulated in a tissue-specific and exogenous phosphate dependent manner for both monocots and dicots. Within closely related dicots, spatial patterns of GUS reporter activity for all AhHMA4 and AtHMA4 promoter constructs were highly similar when examined in both A. halleri and A. thaliana (Hanikenne et al., 2008). A. thaliana was therefore employed as a heterologous system to examine and compare N. caerulescens HMA4-2 promoter activity with that of AhHMA4-3 and AtHMA4.

4.6.3. *A. thaliana* is a suitable system for heterologous promoter expression analysis

To test whether variation in promoter sequences resulted in cis regulatory modifications, fusions of the promoters of AtHMA4, AhHMA4-1, AhHMA4-3 and NcHMA4-2 gene copies to the *β*-glucuronidase (GUS) were generated and ligated into the pGWB3 Gateway® compatible destination vector through LR reactions (Invitrogen, Paisley UK). These pGWB3 constructs were then transformed into Agrobacterium tumefaciens GV3101 using electroporation techniques. *A. tumefaciens* were subsequently grown to mid log phase (OD_{600} 0.8) and resuspended in solutions of 5% sucrose. Using the floral dip technique (Clough and Bent, 1998), Arabidopsis thaliana Col 0 wild type flowering plants were transformed by floral submergence in this liquid for 20 s to allow bacterial mediated T-DNA integration. T_1 seeds harvested from these plants were selected in vitro on agar growth media supplemented with kanamycin sulphate (50 µg ml^{-1}) and hygromycin sulphate (50 µg ml^{-1}). Plants which were actively growing one week after germination were transferred to ex vitro glasshouse conditions and selfed. PCR analysis confirmed construct introgression by employing primers specific for both GUS and the promoter of interest. Three independent transgenic T_2 lines were identified for three promoter::GUS fusion constructs i.e. bearing the GUS gene under the control of promoters from NcHMA4-2, AhHMA4-3 or AtHMA4 genes. Three replicates of each line were grown in vitro for 21 days on agar growth media supplemented with half strength MS, under a 16 hr photoperiod at 20 - 22°C. Plants were then removed from in vitro conditions and incubated in a staining solution of 0.52 g l^{-1} 5-bromo-4-chloro-3-indolyl β-D-glucopyranoside overnight at 37°C. Stepwise replacement of the staining solution with ethanol to remove chlorophyll improved visualisation of the blue product. Before viewing using a stereo microscope, plants were mounted in solutions of 50% glycerol. GUS activity was identified in all transformed T_2 plants with little variation between replicates for each or between independent transgenic lines for each construct. Wild type untransformed *A. thaliana* was maintained as a negative treatment control and demonstrated no GUS activity throughout.

4.6.4. GUS regulation is promoter specific

GUS activity throughout entire transgenic A. thaliana plants showed remarkable spatial variation depending on the regulating promoter. For T_2 lines transformed with the AtHMA4 promoter fused with GUS (AtHMA4p::GUS), GUS expression was localised in the root and shoot vasculature. Similar results were observed by Hussain et al. (2004) where transgenic A. thaliana plants containing HMA4p::GUS fusion constructs demonstrated GUS activity in the vascular tissue of the roots and stems, including the endodermal and pericycle cell layers. These findings were later supported by Sinclair et al. (2007) who examined the distribution of zinc (Zn) within 7 – 8 d old seedlings of Arabidopsis thaliana using the (Zn)-fluorophore, Zinpyr-1. Following immersion in 20 μM of Zinpyr-1 and examination using a confocal laser-scanning microscope, (excitation at 488 nm with a 100mW Ar ion laser and a ×60 Plan Apo water immersion lens with fluorescein isothiocyanate and Texas Red filters), it was shown that wild type roots exhibited the highest fluorescence signal in the xylem while relatively little fluorescence was identified in the pericycle cells. Conversely, in a mutant transgenic line which demonstrates a Zn deficient phenotype as a result of a double knock out of both HMA2 and HMA4 (hma2, hma4 knock out), fluorescence predominated in the endodermal and pericycle cell layer. This was replicated in HMA4 but not in HMA2 single knock out lines, suggesting that HMA4 may have a more significant role in root to shoot translocation of Zn. Using similar AtHMA4 promoter::GUS fusions, Hanikenne et al. (2008) demonstrated comparable spatial patterns for reporter activity in both A. thaliana and A. halleri transgenic lines.

GUS expression in A. thaliana bearing the NcHMA4-2p::GUS construct was spatially significantly different to that of AtHMA4p::GUS lines. As with AtHMA4p::GUS, expression was distributed both in the root and shoot vasculature, however strong GUS expression was similarly represented in leaf lamina. Similar expression distribution was observed in A. thaliana transgenic lines bearing the AhHMA4-3p::GUS construct. Examination of GUS expression in A. thaliana and A. halleri lines bearing a similar AhHMA4-3p::GUS construct by Hanikenne et al. (2008) revealed comparable spatial distributions within both genomes. Further inspection of transverse sections of roots from their A. thaliana and A. halleri transformants revealed that GUS expression was localised in the xylem parenchyma and pericycle cell layers. In situ HMA4 messenger RNA hybridisations in A. halleri

roots confirmed HMA4 specific mRNA accumulation in these cell types, thus supporting previous evidence for a role of HMA4 in xylem loading of Zn (Bernard et al., 2004; Hussain et al., 2004; Papoyan and Kochian, 2004; Verret et al., 2004, 2005; Hanikenne et al., 2008). Similar analysis performed on transverse sections of leaf tissue from A. thaliana and A. halleri identified GUS expression in the xylem parenchyma and cambium cell layers, which endorses proposed roles for HMA4 in Zn exclusion from specific cell types and distribution within the leaf (Frey et al., 2000; Hussain et al., 2004; Talke et al., 2006; Hanikenne et al., 2008).

In the absence of an efficient stable transformation system for N. caerulescens, NcHMA4 promoter expression and localisation can be observed through heterologous expression in Arabidopsis thaliana. Previous comparisons between heterologous promoter::GUS fusion expression profiles in A. thaliana and in endogenous systems have frequently demonstrated that spatial patterns of reporter activities were highly similar between species, thus indicating that A. thaliana was a robust heterologous expression system (Tittarelli et al., 2007; Hanikenne et al., 2008). Results from the present study demonstrated that NcHMA4-2p and AhHMA4-3p presented very similar spatial expression patterns when fused with GUS. Unlike the AtHMA4 promoter, NcHMA4-2p and AhHMA4-3p showed remarkable expression throughout the leaf lamina which was consistent with previous evidence associating up regulation of HMA4 with increased Zn in leaf tissues among Zn hyperaccumulators (Verret et al., 2004; Hammond et al., 2006; Talke et al., 2006; Broadley et al., 2007; Hanikenne et al., 2008; Barabasz et al., 2010; Siemianowski et al., 2010). Moreover, such extensive expression patterns throughout the leaf supported highly concentrated Zn sequestration patterns reported in mesophyl vacuoles and cell walls of epidermal cells in N. caerulescens (Frey et al., 2000; Ma et al., 2005). Further investigations will be required to gain a thorough appreciation of these associations.

However, to gain a comprehensive understanding of the role of HMA4 in N. caerulescens an efficient direct mutagenic approach in a fast cycling system is required. Attempts to transform this species have been hampered by poor reproducibility and a recalcitrance from some ecotypes to Agrobacterium tumefaciens mediated protocols (Peer et al., 2003). In addition N. caerulescens

requires an extensive nine month culture period to produce seed including a 7 – 12 week period of short-day vernalisation (Chapter 5). With this in mind efforts to establish a rapid cycling system for efficient transformation through A. tumefaciens mediated approaches were described in Chapter 5.

4.7. SUMMARY

- Promoter regions for AhHMA4-1, AhHMA4-3, AtHMA4 and NcHMA4-2 were ligated into pGWB3 promoter::GUS destination expression vectors.
- All pGWB3 vectors were cloned into Agrobacterium tumefaciens strain GV3101 through electroporation and employed to transform Arabidopsis thaliana through the 'floral dip' approach.
- T_2 segregating A. thaliana transformants were generated containing promoters from AhHMA4-1, AhHMA4-3, AtHMA4 or NcHMA4-2 fused to the GUS marker gene.
- Histochemical analysis of T_2 transformants illustrated that GUS activity for plants bearing NcHMA4-2 and AhHMA4-3 promoters demonstrated remarkable similarities in GUS spatial expression.
- GUS expression for both AhHMA4-3 and NcHMA4-2 transformants was evident in both root and shoot tissue whereas expression in the AtHMA4 transformants was localised in root vasculature.

CHAPTER 5 ADAPTING NOCCAEA CAERULESCENS FOR EFFICIENT MOLECULAR GENETIC ANALYSES

5.1. INTRODUCTION

Elucidating the functions of Noccaea caerulescens HMA4 genes and their regulatory sequences in planta is essential to fully exploit their associated roles in Zn hyperaccumulation and develop an appreciation of novel protein interactions within this genome. Prior to this study little information existed regarding the genomic structure of the NcHMA4 locus and none regarding its in cis regulatory control. In Chapter 3 it was demonstrated for the first time that Noccaea caerulescens Saint Laurent Le Minier contained a tandem HMA4 quadruplication which was remarkably similar to an evolutionarily independent HMA4 triplication event in the distantly related Zn hyperaccumulator Arabidopsis halleri (Hanikenne et al., 2008). This discovery was consistent with previous reports of higher expression levels for HMA4 orthologues from both species when compared to several closely related non-hyperaccumulating species (Bernard et al., 2004; Hammond et al., 2006; Talke et al., 2006; van de Mortel et al., 2006). Hanikenne et al. (2008) demonstrated that higher HMA4 transcript levels in A. halleri were the result of both gene copy number and in cis gene regulation. A novel HMA4 promoter sequence from N. caerulescens was shown to have similar activity to an A. halleri HMA4 promoter under heterologous conditions (Chapter 4). Both demonstrated spatially similar expression patterns in both roots and shoots and differed significantly to the A. thaliana HMA4 promoter. Results from these chapters provided strong additional support for a functional role of NcHMA4 in Zn hyperaccumulation. Although heterologous expression in A. thaliana provided preliminary data for the regulatory role of the NcHMA4-2 promoter, a definitive characterisation within N. caerulescens was required. Direct mutagenesis of these genes and promoter sequences would provide researchers with the tools necessary to fully characterise this hyperaccumulation phenomenon. In planta mutagenesis would offer a more sophisticated understanding of these genes and their putative interactions within the genome. Comparative analysis of the molecular mechanisms involved in Zn hyperaccumulation between species where this trait has independently occurred would provide invaluable data in terms of the development of plant evolutionary biology. At an applied level, information regarding natural genetic variation in Zn accumulation should provide the molecular

data necessary to develop Zn biofortification strategies for a sustainable reduction to global Zn malnutrition, which currently affects over 2 billion people (White and Broadley, 2005; Palmgren et al., 2008; White and Broadley, 2009). It was therefore necessary to develop an efficient direct mutagenic system to functionally analyse this natural genetic variation. However, dissecting these molecular mechanisms in N. caerulescens has proved challenging due to a number of significant obstacles. Within the species, many ecotypes display an apparent recalcitrance to Agrobacterium tumefaciens mediated transformation and currently only two solitary reports of transformation events, using different techniques, within Noccaea caerulescens have been published (Peer et al., 2003; Guan et al., 2008). Moreover, this species requires a lengthy growth period and has an obligate vernalisation requirement to initiate flowering. Thus, ecotypes cultivated to date require up to 32 weeks to flower, including a 7-12 week period of short-day vernalisation (5°C and 8 hr photoperiod), with an additional 4 weeks culture to achieve fully ripened seeds (Peer et al., 2003, 2006). In addition to this extensive culture period, growing plants for up to 9 months in controlled environments represents a significant challenge (and cost) in terms of husbandry, including maintaining plants in a disease-free state. A more rapid cycling system which required no vernalisation to initiate flowering was therefore required to provide a basis for these crucial future molecular genetic studies. Previous assays have shown that the removal of vernalisation requirements to induce flowering led to the development of rapid cycling populations in several important model Brassicaceae species including crop Brassica ssp., which has greatly benefited current molecular genetic analyses (Williams and Hill, 1986; Iniguez-Luy et al., 2009). In late flowering ecotypes of Arabidopsis thaliana, the vernalisation requirement had been removed through fast neutron induced mutations in either FLOWERING LOCUS C (FLC) and FRIGIDA (FRI) which interacted synergistically to repress flowering (Michaels & Amasino, 1999; Sung & Amasino, 2004). Recent expression analysis had similarly identified conserved roles for FLC homologues in vernalisation responses in Brassica rapa (Zhao et al., 2010) and Beta vulgaris (Reeves et al., 2007) and in the perennial species Arabis alpina (Wang et al., 2009).

It was hypothesised therefore that the vernalisation requirement of Noccaea caerulescens Saint Laurent Le Minier could be removed through fast neutron mutagenesis and that this would induce rapid cycling, vernalisation independent

plants, to support future forward and reverse molecular genetic approaches. To test this, plants were subjected to fast neutron bombardment and selected for early flowering phenotypes independent of vernalisation. Growth responses to Zn concentrations were examined for in vitro growth media to identify deficient, replete and toxic levels which could be employed for future comparative physiological and molecular genetic assays in the model Zn hyperaccumulator N. caerulescens, and the model Zn non-hyperaccumulator A. thaliana. Furthermore, root and shoot biomass were compared for each treatment to identify variances between optimal exogenous Zn levels for roots and shoots in both the Zn hyperaccumulator and the Zn non-hyperaccumulator species. Finally Agrobacterium tumefaciens mediated transformation approaches adapted from Peer et al. (2003) were examined for functionality in Noccaea caerulescens Saint Laurent Le Minier.

5.1.1. Fast Neutron Mutagenesis of Noccaea caerulescens

Mutation breeding using fast neutron bombardment of seed has been shown to create random deletions ranging from one base to greater than 100 kb and has been commonly employed in mutating plant genomes, representing a rapid approach to obtain large mutant pools (Kodym & Afza, 2003; Salt et al., 2008; Bruce et al., 2009). The technique represents a relatively inexpensive method of producing large mutant populations in species whose genomes are not amenable to T-DNA transformation, generating genome-wide saturation in relatively small populations (Bruce et al., 2009). Seeds are subjected to fast neutron exposure (bombardment) for a specified period and subsequently grown to create mutant populations for screening (Salt et al., 2008; Gilchrist and Haughn, 2010). It has been most commonly used for mutagenesis in A. thaliana with induced mutations typically including chromosomal deletions and rearrangements although some point mutations have been recorded (Salt et al., 2008). Due to the breadth of these deletions and rearrangement, this form of mutation has proven most suitable to experiments requiring complete loss-of-function alleles. In contrast, chemical mutagens such as ethyl methanesulfonate (EMS) induce point mutations which typically result in a broad range of genetic effects including including hypomorphic, hypermorphic and neomorphic effects (i.e. alleles of reduced, enhanced or novel gene function, respectively) (Stephenson et al., 2010). Similarly T-DNA insertional mutagenesis can often lead to hypomorphic alleles rather than the desired loss-of-function

mutation and requires a much larger population size than either chemical or physical mutagenesis techniques (Gilchrist and Haughn, 2010). Moreover, it has been shown in other genomes that complete loss-of-function of either the FLC or FRI orthologues resulted in alterations to vernalisation responses (Reeves et al., 2007; Wang et al., 2009; Zhao et al., 2010). Therefore, a fast neutron mutagenic approach was deemed more suitable to induce alterations in the vernalisation requirement of Noccaea caerulescens.

5.1.2. Transformation and mutagenesis of Noccaea caerulescens

Efficient Agrobacterium tumefaciens mediated transformation of the model plant Arabidopsis thaliana, as well as crops such as tomato and tobacco, has revolutionised current understanding of the plant genome. However, despite being regarded by most as a model hyperaccumulator species (Reeves et al., 2001; Broadley et al., 2007; Verbruggen et al., 2008; Krämer et al., 2010), transformation in Noccaea caerulescens has been challenging (Peer et al., 2003; Guan et al., 2008). It shares approximately 88% nucleotide sequence identity within the coding regions with A. thaliana and like the model plant species, it is a small (0.7 pg) diploid 2n = 2x = 14 genome and self compatible (Peer et al., 2003, 2006). Moreover, it is distinguished from most plants by hyperaccumulating Zn, Cd and Ni and contained the highest ever reported leaf Zn concentration of 53,450 μg g^{-1} dry biomass (Reeves et al., 2001). Despite this, current understanding of the molecular genetics surrounding this phenomenon remains inadequate, primarily due to a lack of direct evidence from functional reverse genetic analyses (Peer et al., 2003; Broadley et al., 2007; Verbruggen et al., 2008; Krämer et al., 2010). Alternatives to transformation approaches exist; however, they are often inefficient and offer a limited range of functional molecular genetic assays. Techniques such as virus-induced gene silencing (VIGS) which degrades plant targeted gene transcripts through the creation of small interfering RNA (siRNA) molecules, via viral transfection, can result in gene knockout or knockdown phenotypes. However non target genes can be affected, gene silencing levels are often variable, missense mutations cannot be created and induced mutations remain transient (Gilchrist and Haughn, 2010). Although fast neutron mutagenesis is heritable and can create gene knockouts, it randomly generates deletions and chromosomal rearrangements in the genome which can impede efficient analysis (Salt et al., 2008; Bruce et al., 2009; Gilchrist and

Haughn, 2010). Introducing point mutations in DNA through ethyl nitrosourea (ENU) or ethylmethane sulphonate (EMS) mutagenesis is heritable and can induce both loss of function and gain of function mutants leading to null, hypomorphic, neomorphic or hypermorphic phenotypes. However, this approach is relatively expensive, labour intensive and mutations are randomly distributed in the genome (Colbert et al., 2001; Till et al., 2003; Stephenson et al., 2010). Despite their benefits these approaches have a number of limitations in terms of both reverse genetics and functional analyses e.g. promoter::GUS fusion assays.

This chapter therefore describes the process of mutagenising Noccaea caerulescens Saint Laurent Le Minier using fast neutrons and the subsequent selection for fast cycling lines, flowering independently of vernalisation. An Agrobacterium tumefaciens mediated T-DNA transformation technique is described using published protocols for N. caerulescens (Peer et al., 2003). Finally the establishment of an in vitro growth system for robust molecular genetic analysis of Zn supply to N. caerulescens and Arabidopsis thaliana was described.

5.2. AIMS

The aim of this chapter was to characterise growth responses of Arabidopsis thaliana Colombia and Noccaea caerulescens Saint Laurent Le Minier to ensure culture conditions were optimised for downstream molecular genetic analyses. Both Noccaea and Arabidopsis were grown in vitro on media supplemented with increasing Zn concentrations to identify levels suitable for future comparative physiological and molecular genetic analyses to Zn deficiency, repletion and toxicity. Noccaea culture was characterised following i) physical mutagenesis ii) Zn and Cd ex vitro media supplementation and iii) 'floral dip' transformation applications.

5.3. OBJECTIVES

- To develop faster cycling lines of Noccaea caerulescens through fast neutron physical mutagenesis techniques.
- To determine the response of Noccaea caerulescens growth to different external Zn concentrations in vitro.
- To determine the response of Arabidopsis thaliana growth to different external Zn concentrations in vitro.
- To determine fecundity in Noccaea in response to different external Zn and Cd concentrations supplied to peat-based growth media ex vitro.
- To test the floral dip transformation system in Noccaea caerulescens.

5.4. MATERIALS AND METHODS

5.4.1. Ex vitro culture of Noccaea caerulescens

5.4.1.1. Vegetative growth of wild type Noccaea caerulescens

Unless stated otherwise, all Noccaea caerulescens (J. & C. Presl) F.K.Mey. ecotype Saint Laurent Le Minier plants were cultured under glasshouse conditions (16 h photoperiod using supplementary sodium lighting at 22.3 ±4°C and 13.3 ±2°C mean day and night temperatures respectively) (Section 2.3.2.). Unless otherwise stated all plant containers were placed in impermeable growth tray (length 97 cm; width 38 cm; height 5 cm) (Giant Plant Grobag Tray, Sankey, UK) and maintained by basal irrigation, thus preventing aerial humidity and mineral leaching. Seeds were sown in plug trays (2 cm^2 plugs) containing Levington M3 high nutrient peat based compost (pH 5.3–5.7) (Monro Group, Wisbech, Cambridgeshire, UK). 14 days after sowing (DAS) seedlings were transplanted into 0.32 l pots (height 7.9 cm; diameter 9 cm) containing a compost mix (Levington M3, sand (<1 mm) and grit (1-3 mm) at a ratio of 2:1:1 respectively v:v:v) (Monro Group) (Section 2.3.1.) which was supplemented with a single dose of 2000 ppm $ZnSO_4$ (455 mg Zn kg^{-1} dry compost) and grown for a further 11 weeks before transportation to vernalisation conditions to initiate flowering (Section 5.4.1.2.) (Fig. 5.1).

5.4.1.2. Floral initiation of wild type Noccaea caerulescens by vernalisation

Noccaea caerulescens were transferred to a controlled environment (CE) conditions for 10 wks subjected to 5 ± 1°C. An 8 hr photoperiod was maintained with lighting levels of 137-147 µmol m^{-2} s^{-1} from 8 metal halide lamps (Osram Powerstar HQI-BT 400W/D Osram, Berkshire, UK) and mean relative humidity (RH) at 83% (±10%). Phenotypic development was recorded daily while temperature and relative humidity were monitored every 10 minutes using a data logger (Tinytag Plus 2, Gemini Data Loggers Ltd., Chichester, UK). After 10 wks plants were returned to glasshouse conditions (Section 2.3.2.) to flower and form seeds during a 4 - 6 week period (Fig. 5.1).

Figure 5.1 Growth stages of wild type Noccaea caerulescens Saint Laurent Le Minier under glasshouse conditions (16 h photoperiod using supplementary sodium lighting at 22.3 ±4°C and 13.3 ±2°C mean day and night temperatures respectively) (Section 2.3.2.). **1,** 12 week old plants prior to vernalisation treatment. **2**, 25 week old plants following 10 wks vernalisation (5 ± 1°C at an 8 hr photoperiod) (Section 5.4.1.2.).

5.4.1.3. Characterising the effects of Zn and cadmium supply on seed fecundity in Noccaea caerulescens

Noccaea caerulescens were grown in 0.32 l pots containing Levington M3 compost mix as described in Section 2.3.1. but were subjected to a matrix of four Zn and four cadmium (Cd) growth media supplementation treatments (16 treatments in total), after one month of vegetative growth. A randomised block design comprising three replicates was employed, with 10 individual pots for each of 16 treatments allocated at random within each replicate within a growth tray (n=480) (length 48 cm; width 30 cm; height 5 cm) (Multi Tray, Sankey, UK) (Section 2.3.2.). To each growth tray, a single Zn and Cd solution was applied. Treatment concentrations were calculated on a parts per million (ppm) basis of treatment ($ZnSO_4$ or $CdSO_4$) to dry compost. Treatments consisted of 0, 2,000, 4,000 and 10,000 ppm $ZnSO_4.7H_2O$ and 0, 250, 500, 1,000 ppm $CdSO_4$ $8/3H_2O$. Seeds produced were weighed and counted for each condition.

5.4.2. Growth media preparations for in vitro culture of Noccaea caerulescens and Arabidopsis thaliana

For both Noccaea caerulescens Saint Laurent Le Minier and Arabidopsis thaliana Col-0, plants were cultured on media containing 8 g l^{-1} agar (A1296, Sigma-Aldrich, Poole, UK) 10 g l^{-1} sucrose (Fisher Scientific, Loughborough, UK), no added growth regulators, and the equivalent to half strength (MS) for all basal medium nutrients except for Zn, which was adjusted independently (Murashige and Skoog, 1962)

(Table 5.1). Prior to the addition of nutrient elements, 8 g of agar was suspended in 500 ml of 10% (v/v) HCL and mixed in a 1 l Duran bottle on a magnetic stirrer for 16 h at 5°C (acid wash). The acidic solution was decanted and the residual agar resuspended in 500 ml milli-Q water (18.2 MΩ cm) and mixed on a stirrer as previous, for 10 min at 5°C before allowing to settle for 30 min. The milli-Q water was decanted and the whole rinsing process repeated three times before a final resuspension with added plant nutrients to a total volume of 1 l. Solutions were adjusted to pH 5.6 with 0.1 M KOH, and autoclaved at 121 ºC under 104 kPa for 20 mins. All elements were purchased from Sigma-Aldrich (Poole, UK) and were > 90% pure on a trace element basis.

Table 5.1 Stock nutrient solution compositions for in vitro agar-based growth media supplementation.

Compound	Molecular Weight	100 X Stock solutions	Volume Required for 1 L of Final Agar Growth Media	Concentration in Final Agar Media (Half strength MS)
Macronutrients	**MW**	**mM**	**ml**	**mM**
NH_4NO_3	80.04	2061.47	5	10.30735
KNO_3	101.11	1879.33	5	9.39665
$CaCl_2$	147.02	225.99	5	1.12995
$MgSO_4$.7H2O	246.48	150.08	5	0.7504
KH_2PO_4	136.09	124.91	5	0.62455
Micronutrients	**MW**	**µM**	**ml**	**µM**
H_3BO_3	61.83	10030	5	50.15
$Na_2EDTA.2H_2O$	372.24	10010	5	50.05
FeNaEDTA	367.05	10000	5	50
$MnSO_4.H_2O$	151	10000	5	50
$ZnSO_4.7H_2O$	287.54	2990	5	14.95
KI	166.01	500	5	2.5
$Na_2MoO_4.2H_2O$	241.95	100	5	0.5
$CoCl.6H_20$	237.93	10	5	0.05
$CuSO_4.5H_2O$	249.68	10	5	0.05

5.4.2.1. Characterising the in vitro growth effects of zinc on Noccaea caerulescens and Arabidopsis thaliana

Arabidopsis thaliana Col-0 and Noccaea caerulescens Saint Laurent Le Minier were employed to establish growth responses to external Zn concentrations in vitro. The Col-0 accession was selected since it was employed for previous heterologous promoter::GUS fusion studies, it was amenable to efficient transformation

techniques, and was a representative of non Zn hyperaccumulator species. N. caerulescens Saint Laurent Le Minier was selected for analysis since it was a model self compatible heavy metal hyperaccumulator.

Arabidopsis thaliana were grown on agar-based growth media (Section 2.4.2.) supplied with nine Zn concentrations between 0 and 300 μM Zn for 21 days before harvesting. N. caerulescens were grown similarly but supplied with 13 Zn concentrations between 0 and 6000 μM and harvested 45 DAS. Zinc sulphate ($ZnSO_4$) was used to supply Zn to all plants in all experiments. Agar containing no added Zn was mixed with agar containing 1000 or 6000 μM Zn in order to obtain a range of Zn concentrations. A. thaliana were grown on nutrient agar supplemented with 0, 0.1, 0.3, 1, 3, 10, 30, 100 and 300 μM Zn. N. caerulescens were grown on nutrient agar supplemented with 0, 0.1, 0.3, 1, 3, 10, 30, 100, 300, 600, 1000, 3000 and 6000 μM Zn. Sixteen seeds for each species were sown in un-vented, polycarbonate culture boxes on 75 ml agar-based media (Section 5.4.2.) for each Zn treatment and maintained in growth room conditions (20 ±2 °C for a 16h photoperiod, at a light intensity of 50 – 80 μmol photons m^{-2} s^{-1}). Fresh and dry weights of shoot and root material were then determined for individual plants.

5.4.3. In planta mutagenic techniques in Noccaea caerulescens

5.4.3.1. Fast neutron mutagenesis of Noccaea caerulescens Saint Laurent Le Minier

Approximately 5500 M_0 Noccaea caerulescens Saint Laurent Le Minier seeds were irradiated at the Biological Irradiation Facility (BIF) of the Budapest Research Reactor (BRR) (Budapest Neutron Centre (BNC), Hungary) in order to produce a population of fast neutron mutants. The BRR was a tank type reactor, moderated and cooled by light water with a nominal thermal power of 10 MW. For seed irradiation with fast neutrons the filter/absorber arrangement number 2Y was used, and the order of the filters used to decrease gamma and neutron intensity, modify the neutron spectrum, and the neutron-gamma ratio, were, from the core towards the irradiation cavity; internally 143.6 mm Al + 18 mm Pb + 15 mm Al, and at the beam stop behind the sample, 30 mm Fe + 45 mm Pb + 8 mm Al + 20 mm B_4C. Seeds were irradiated inside a rotating Cd capsule (16 rpm) of 2 mm wall thickness. Irradiation temperature was less than 30˚C, at normal air pressure and humidity less than 70%.

The mean neutron dose rate (water kerma ~ absorbed dose in water) was monitored in real time by U-235 and Th-232 fission chambers and a Geiger-Müller counter and was 438 mGy min^{-1} (±3.0%) (milliGrays per minute) at 10 MW. During real time dose monitoring the irradiation was terminated when the required dose was delivered. The samples were re-packed to avoid surface contamination and the activation of the original container tested. The measured surface gamma dose from the seeds was 130*background dose. Upon dispatch, this had decreased to <2*background dose, where (background dose is ~90 nGy h^{-1}).

5.4.3.2. Selection of faster cycling fast neutron mutagenised Noccaea caerulescens

M_1 plants were grown under glasshouse conditions (GC) set to 16 hr photoperiod (Section 2.4.2.). Seeds were sown in plug trays (2 cm^2 plugs) containing Levington M3 high nutrient peat based compost (pH 5.3–5.7) (Monro Group, Wisbech, Cambridgeshire, UK). Seven day old seedlings were then transplanted into 0.32 l pots in a compost mix as described (Section 2.3.1.). After 12 weeks, plants were vernalised as described (Section 5.4.1.2.) for 10 weeks, before returning to GC. Subsequently, 80,000 M_2 lines were grown under GC in 2 cm^2 plugs and after 2 weeks transplanted into 0.32 l pots to select for non-vernalised early flowering (within 16 weeks) individuals. Subsequent faster cycling M_3 progeny were selfed and selected under similar conditions as M_2 lines. Two M_4 lines, observed to have consistently high fecundity and early flowering phenotypes (A2 and A7), were selected for further characterisation.

5.4.3.3. Characterising early flowering mutants

Seeds of two M_4 early flowering lines (A2 and A7) and one S_2 wild type were grown for 123 d under controlled environment (CE) conditions. The CE conditions were 16 hr photoperiod ,19°C (±2°C), lighting levels of 137-147 µmol m^{-2} s^{-1} from 8 metal halide lamps (Osram Powerstar HQI-BT 400W/D Osram, Berkshire, UK) and relative humidity (RH) 83% (±10%). Two seeds were germinated in 1.05 l pots (height 11.3 cm; diameter 13 cm) containing 1 l of compost mix 2:1:1 (v:v:v) Levington M3 high nutrient compost:perlite (2-5 mm):vermiculite (2-5 mm) (Monro Group) (Section 2.3.1.). One plant from each pot was later harvested for mineral analysis, and the remainder left to fruit. A randomised block design comprising six

replicates was used, with six individual pots of each of three lines allocated at random within each replicate within a growth tray (length 97 cm; width 38 cm; height 5 cm) (Giant Plant Grobag Tray, Sankey, UK). A single Zn solution was applied to all pots on a single occasion after one month of growth, supplying 455 mg Zn kg^{-1} compost. Phenotypic development was recorded daily while temperature and relative humidity were monitored every 10 minutes using a data logger (Tinytag Plus 2, Gemini Data Loggers Ltd., Chichester, UK).

5.4.3.4. Inductively coupled plasma-mass spectrometry (ICP-MS) to characterise mineral composition of early flowering mutants

Aerial tissue was harvested, dried at 60°C for 2 d and homogenized manually to ensure particle size uniformity. Approximately 300 mg of dried tissue was digested under closed-vessel microwave heating (45 min, 20 bar) in 2 ml of 70% trace analysis grade (TAG) HNO_3, 1 ml H_2O_2 (Fisher Scientific UK Ltd, Loughborough, Leicestershire, UK) and 1 ml milli-Q water (18.2 MΩ cm). The microwave system comprised a Multiwave 3000 platform with a 48-vessel 48MF50 rotor (Anton Paar GmbH, Graz, Austria). Samples were digested in perfluoroalkoxy (PFA) liners inserted into polyethylethylketone (PEEK) pressure jackets (Anton Paar, Hertford, UK). Digested samples were diluted to 15 ml with milli-Q water and stored at room temperature. Mineral analysis was conducted using inductively coupled plasma-mass spectrometry (ICP-MS) as described previously (Broadley et al., 2010). Samples were further diluted 1-in-10 with milli-Q water and analysed using an ICP-MS (X-SeriesII, Thermo Fisher Scientific Inc., Waltham, MA, USA). Internal standards included Sc (50 ng ml^{-1}), Rh (10 ng ml^{-1}) and Ir (5 ng ml^{-1}) in 2% TAG HNO_3. External multi-element calibration standards (Claritas-PPT grade CLMS-2, SPEX Certi-Prep Ltd, Stanmore, Middlesex, UK) included Al, As, Ba, Bi, Cd, Co, Cr, Cs, Cu, Fe, Mn, Mo, Ni, Pb, Rb, Se, Sr, U, V, and Zn, in the range 0-100 µg l^{-1}, and Ca, Mg, K and Na in the range 0-100 mg l^{-1}. Data were corrected using blank digestions.

5.4.3.5. In planta Agrobacterium –mediated transformation of N. caerulescens

Single Agrobacterium tumefaciens colonies containing the constructs of interest were grown at 29˚C for approximately 16 hrs, in 100 ml liquid LB with selective antibiotics to an OD_{600nm} of 0.7 – 0.8. Since antibiotics were phytotoxic, the cultures were centrifuged at 3000 g for 20 mins. Cells were resuspended in a 5% sucrose and

0.02% v/v Silwet L-77 surfactant (Vac-In-Stuff, Lehle Seeds, USA) maintaining an OD_{600nm} of approximately 0.8 – 1. Plants which produced their first fully opened flower were subjected to bacterial solutions. Individual flowers and early emerging flower primordia were submerged in these solutions for 10 sec (Peer et al., 2003; Zhang et al., 2006). Humidity was maintained by placing individual plants into closed plastic sleeves (Zwapak, Alsmeer, Netherlands). This was assumed to optimise bacterial transformation conditions and prevent desiccation as plants began cuticle regeneration, damaged by dipping culture surfactants (Clough and Bent, 1998: Dr. Sean May, pers. comm.). Plant sleeves were gradually reopened over a 72 h period until complete removal after 96 hrs. T_1 seeds were harvested 4 – 6 weeks after floral dipping.

5.4.3.6. In vitro selection of T_1 transgenic Noccaea caerulescens

T_1 Noccaea caerulescens were selected in vitro on agar-based medium containing half MS using the protocol described in (Section 2.4.2.) but supplemented with 50 µM $ZnSO_4$ and ticarcillin disodium:potassium clavunalate (15:1) solution (Melford Laboratories, UK) to a final concentration of 125 mg l^{-1} to eradicate any residual Agrobacterium tumefaciens remaining on seeds. All T_1 seeds were surface sterilised as described in Section 2.4.1. with final axenic seed suspensions decanted onto sterile filter paper in the laminar air flow cabinet and left to dry for 4 - 5 hrs. Once adequately dried, seeds were evenly sprinkled onto agar-based media in each growth box. Box lids were secured using Nescofilm (Bando Chemical Co., Kombe, Japan), before placing in growth rooms at 20 ±2 °C for a 16h photoperiod, at a light intensity of 50 – 80 µmol photons m^{-2} s^{-1}.

5.4.3.7. Floral dip using GUS::NPTII and NcHMA4RNAi gene constructs

The GUS::NPTII (kanamycin resistance) and NcHMA4RNAi (spectinomycin resistance) constructs in Agrobacterium tumefaciens strains 1065 and C58 respectively were employed for floral dipping.

The GUS::NPTII construct (donated by Dr. Paul Anthony, University of Nottingham), which was cloned into the supervirulent A. tumefaciens strain 1065 (Curtis et al., 1994), was formed through the introduction of the TOK47 Ti plasmid and a binary vector pMOG23 (Sijmons et al., 1990) into A. tumefaciens strain

LBA4404. The pMOG23 was derived from pBIN19 and contained a NOS.NPTII.NOS between the left and right border sequences and a CaMV35S.GUS-intron reporter gene (Vancanneyt et al., 1990) inserted in a multiple cloning site between these border sequences (Jin et al., 1987). Using this construct, transformed plant cells degraded the substrate X-Gluc (5-bromo-4-chloro-3-indoyl *β*-D-glucuronic acid) resulting in a blue colouration, permitting their identification (Curtis et al., 1994). Transgenic plants were conferred with kanamycin resistance.

The NcHMA4RNAi construct (donated by Dr. Victoria Mills, University of Nottingham) was cloned into A. tumefaciens C58 and was created using the Gateway™ LR Clonase™ reaction between the pCR8®/GW/TOPO® entry vector (Section 4.4.1.) harbouring a 400 bp N. caerulescens HMA4 cDNA fragment and the destination vector pK7GW1WG2 (II) (Section 2.5.3.2.). This resulted in a construct that produced double stranded RNA (hairpin RNA) triggering post transcriptional gene silencing of HMA4 in planta (Vanhecke and Janitz, 2005; Karimi et al., 2002) and conferred in planta kanamycin resistance.

5.5. RESULTS

5.5.1. Establishing an in vitro growth environment to observe Zn dependent plant growth responses

The effects of Zn nutrition on both Arabidopsis thaliana and Noccaea caerulescens plant growth were determined by growing plants in vitro on agar-based growth media. Plants were subjected to Zn concentrations ranging from 0 to 300 μM for A. thaliana and 0 – 6000 μM for N. caerulescens before harvesting 21 and 45 days after sowing (DAS) respectively. The growth traits determined for both plants were shoot and root dry weights expressed as (g plant^{-1}) to reduce errors caused by variations in shoot water content due to potential differences in water availability between different treatments. Previous attempts to attain Zn deficiency responses for A. thaliana on agar containing no Zn supplementation were unsuccessful (data not shown). To reduce background Zn contamination in agar, two assays were attempted. Agar was stirred overnight in a beaker containing either milli-Q water (18.2 MΩ cm) or dilute HCL (10% v/v) before decanting and rinsing with milli-Q water ahead of the mineral salt supplementation and autoclaving (Section 5.4.2.). To test if these water and HCL (10% v/v) 'wash' treatments could permit an observable Zn-dependent growth response, plants were grown in respectively treated agars supplemented with 0 and 15 μM $ZnSO_4$ for 14 d in translucent polycarbonate boxes (Bibby Sterilin, Stone, UK). The growth environment was maintained at 21 °C under a 16 hr photoperiod supplying a light intensity of 50 – 80 μmol m^{-2} s^{-1}. A randomised block design comprising three replicates was employed, with four treatments (two Zn concentrations under two agar wash treatments) allocated at random within each replicate (n = 12). Mean shoot dry weights (d. wt.) of A. thaliana plants growing on agar washed with deionised H_2O showed no significant variation for both 0 and 15 μM $[Zn]_{agar}$. Plants grown on agar washed in diluted HCL (10% v/v) demonstrated a significant reduction in mean shoot d. wt. under 0 μM $[Zn]_{agar}$ ($P < 0.05$) (Fig. 5.2). Plants growing on 15 μM $[Zn]_{agar}$ washed in either H_2O or dilute HCL (10% v/v) demonstrated no significant difference in mean shoot dry weights (Fig. 5.2). Phenotypically plants growing on 15 μM $[Zn]_{agar}$ for both wash treatments were highly similar with good root and shoot development (Fig. 5.2). For both agar wash treatments, plants uniquely growing on agar treated with HCL (10%) demonstrated significantly altered phenotypes in response to altered $[Zn]_{agar}$. Plants growing on 0

μM $[Zn]_{agar,}$ washed with HCL (10%) were chlorotic and had more stunted roots and shoots than those growing on 15 μM $[Zn]_{agar}$ (Fig. 5.2). Together both phenotypic and dry weight data demonstrated that a Zn deficient growth response for Arabidopsis thaliana was evident uniquely on agar treated with HCL (10% v/v). To further elucidate in vitro growth responses of plants to increasing exogenous supplies of Zn, trials were initiated in Arabidopsis thaliana Col-0 and Noccaea caerulescens Saint-Laurent-Le-Minier.

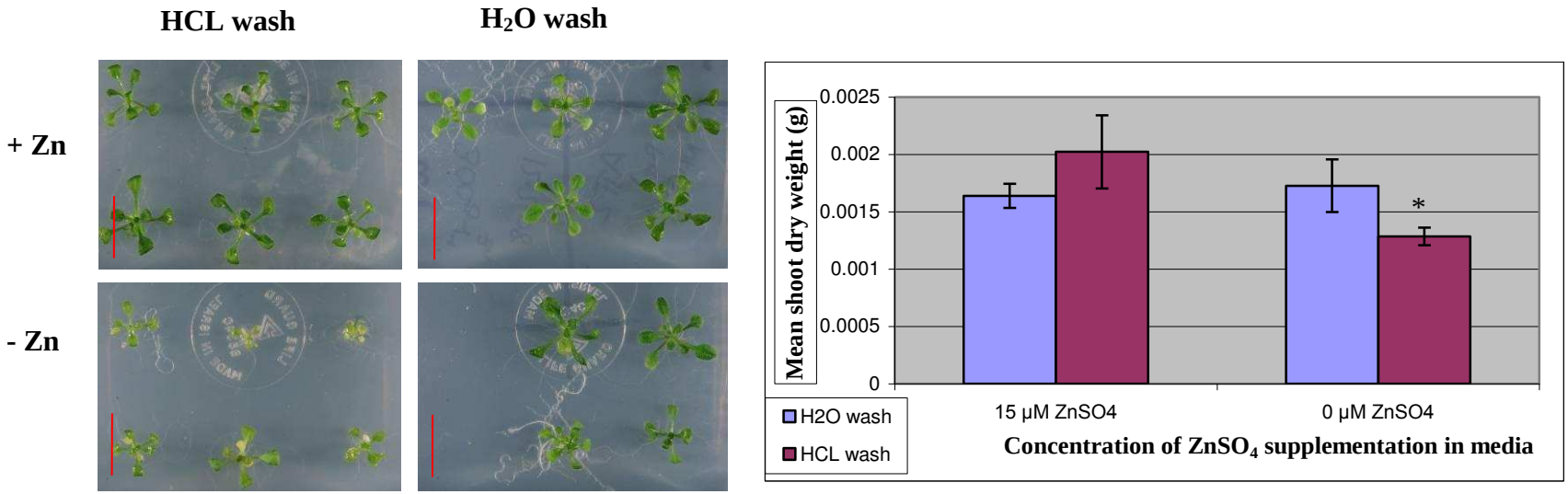


Figure 5.2 Images of 2 week old A. thaliana growing on Agar (0.8 g l^{-1}) based growth media supplemented with half strength Murashige and Skoog (MS) nutrient salts containing either 15 µM (top row +Zn) or 0 uM (bottom row –Zn) $ZnSO_4$. Prior to dissolving in deionised H_2O, Agar was washed over night in a solution of 10 % HCL (left column HCL wash) or deionised H_2O (right column H_2O wash). Red bars represent 1 cm. Mean shoot dry weight taken from these plants is shown in the bar chart. Blue and maroon bars represent samples growing in Agar washed with H_2O or HCL (10% v/v) respectively. Error bars represent standard error of the mean dry weights. Asterisk * indicates significant differences ($P < 0.05$) according to the unpaired Student's t test ($n = 3$ per Zn treatment).

5.5.2. The growth responses of Arabidopsis thaliana to increasing levels of Zn supply in agar

Shoot and root dry weights were determined for A. thaliana plants grown on 75 ml of agar based growth media (10 g l^{-1} sucrose, 8 g l^{-1} agar, no growth regulators) with appropriate replete concentrations of all mineral elements, except Zn, as described (Section 5.4.2.). Zinc ($ZnSO_4$) was subsequently supplemented at 9 different concentrations (Section 5.4.2.1.). Plants were grown for 21 d in translucent polycarbonate boxes (Bibby Sterilin, Stone, UK) at 21 °C under a 16 hr photoperiod supplying a light intensity of 50 – 80 µmol m^{-2} s^{-1}. A randomised block design comprising six replicates was employed, with the nine Zn supplementation treatments allocated at random within each replicate (n = 54). Mean shoot dry weights (d. wt.) of A. thaliana plants increased significantly ($P < 0.0001$) upon supplementation with 0.1 µM $ZnSO_4$ but did not significantly change for plants growing on agar containing between 0.1 and 30 µM Zn. Throughout, mean shoot d. wt. ranged from 0.1 mg to 5 mg $plant^{-1}$ (Fig. 5.4 a). Mean root dry weights increased steadily as Zn concentrations were increased between 0 and 30 µM before reducing sharply thereafter (Fig. 5.4 b). Upon supplementation with 1 µM Zn, root d. wt. was significantly greater than the control (0 µM Zn) ($P < 0.005$) and increased significantly further upon addition of 10 µM Zn ($P < 0.0001$) (Fig. 5.4 b). Phenotypically, plants growing on agar supplemented with 0, 0.1 and 0.3 µM Zn had chlorotic yellow shoot tissue and poorly developed root systems (Fig 5.3). Those growing on 300 µM Zn developed the smallest roots and shoots among all treatments. Those growing on 1, 3, 10 and 30 developed healthy green shoots and well developed root systems (Fig. 5.3). Based on both shoot and root dry weights and phenotypic observations, plants growing on agar supplemented with 30 µM Zn grew optimally, whereas those on agar containing concentrations up to 0.3 µM showed Zn deficiency and those on 300 µM demonstrated Zn toxicity.

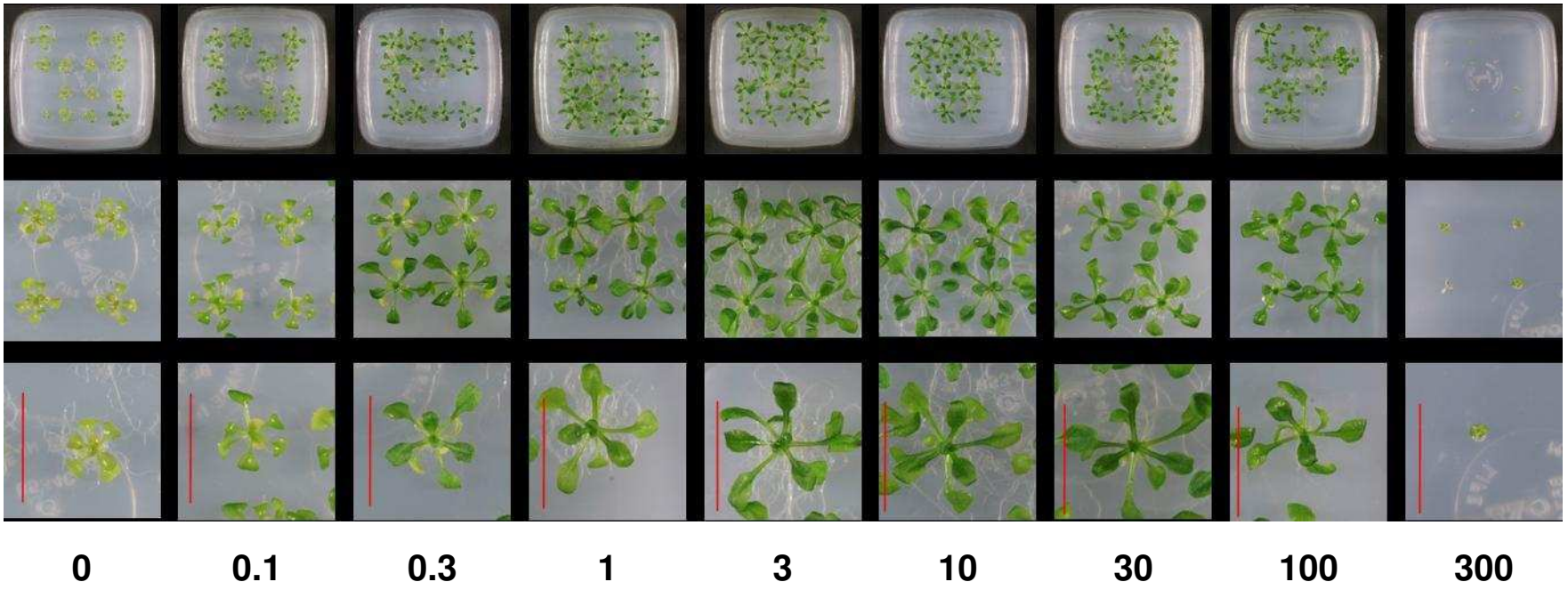


Figure 5.3 21 day old A. thaliana samples growing on HCL (10% (v/v)) washed agar (0.8 g l^{-1}) based growth media supplemented with half strength Murashige and Skoog (MS) nutrient salts and increasing concentrations of $ZnSO_4$. Numbers below images represent concentration of supplemented $ZnSO_4$ in μM. Top row; whole growth boxes, middle row; groups of representative plants, bottom row; single plants. Red bars represent 1 cm.

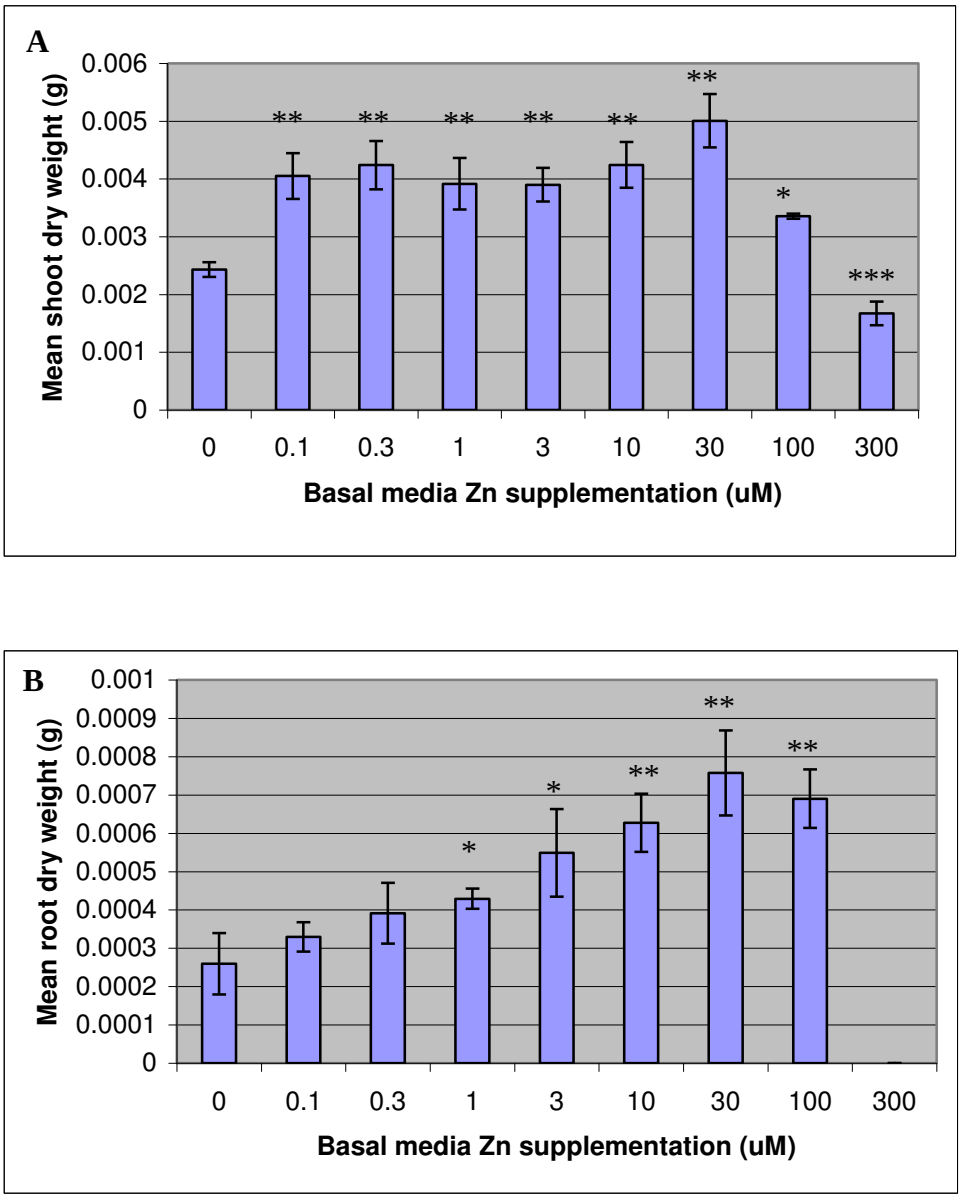


Figure 5.4 Changes in A) mean shoot dry weight (d. wt), and B) mean root d. wt with increasing initial agar Zn concentration. A. thaliana plants were grown for 21 days on solidified agar supplemented with MS salts. Zn concentrations were obtained by mixing relative amounts of agar containing 300 µM $ZnSO_4$ and agar containing no $ZnSO_4$. Data were obtained for 20 – 30 plants per treatment for each repeat experiment. Two experiments were performed. Error bars represent standard error of the mean dry weights. Asterisks indicate significant differences: *; $P < 0.005$, **; $P < 0.0001$, ***; $P < 0.05$ according to unpaired Student's t tests (n = 6 per treatment).

5.5.3. The growth response of Noccaea caerulescens to increasing levels of Zn supply in agar

Shoot and root dry weights were determined for N. caerulescens plants grown on 75 ml growth media containing 10 g l^{-1} sucrose, 8 g l^{-1} agar, no growth regulators and appropriate replete concentrations of all mineral elements, except Zn, as described in

Section 5.4.2.1. Zinc ($ZnSO_4$) was supplemented at 13 different concentrations (Section 5.4.2.1.). Plants were grown for 45 d in translucent polycarbonate boxes at 21 °C under a 16 hr photoperiod supplying a light intensity of 50 – 80 µmol m^{-2} s^{-1}. A randomised block design comprising six replicates was employed, with the 13 Zn supplementation treatments allocated at random within each replicate (n = 78). Mean shoot dry weights (d. wt.) of N. caerulescens plants increased gradually as Zn concentrations increased from 0 to 3 µM (Fig 5.6 A). Mean shoot d. wt. did not significantly change for plants growing on Zn agar concentrations ($[Zn]_{agar}$) between 3 and 30 µM. but were significantly greater ($P < 0.001$) than plants on 0 µM $[Zn]_{agar}$. Mean shoot d. wt. increased steadily for plants on agar containing between 30 µM and 300 µM Zn and were significantly ($P < 0.0001$) greater than plants on 0 µM Zn agar. Mean shoot d. wt. decreased significantly ($P < 0.0001$) upon subsequent $[Zn]_{agar}$ supplementations until no significant d. wt. difference was observed between plants grown on 0 and 6000 µM $[Zn]_{agar}$ (Fig. 5.6 A). Mean root d. wt. similarly increased gradually as $[Zn]_{agar}$ increased to 3 µM and remained significantly unchanged until 30 µM $[Zn]_{agar}$ (Fig. 5.6 B). Mean root d. wt. rose significantly under 100 and 300 µM $[Zn]_{agar}$ ($P < 0.0001$) and thereafter reduced significantly until no significant difference was observed between plants grown on 0 and 6000 µM $[Zn]_{agar}$.

Phenotypically, plant development improved gradually as agar was supplemented with Zn from 0 to 600 µM. Plant size decreased marginally upon supplementation with 1000 µM $ZnSO_4$ and drastically thereafter becoming yellow and developing chlorotic leaves (Fig. 5.5). N. caerulescens developed stunted growth with darkened and bleached leaves under 0 – 0.3 µM $[Zn]_{agar}$. From 1 – 30 µM root and shoot growth increased considerably and leaf chlorosis and bleaching was reduced (Fig. 5.5). Between 100 and 600 µM $[Zn]_{agar}$ root and shoot tissue growth was optimal and resembled phenotypically normal N. caerulescens plants. Based on all dry weight and phenotypic observations, 300 µM $[Zn]_{agar}$ produced optimal healthy N. caerulescens plants. Zinc deficiency symptoms were evident between 0 and 30 uM $[Zn]_{agar}$ while Zn toxicity became significant upon supplementation with 1000 µM $[Zn]_{agar}$.

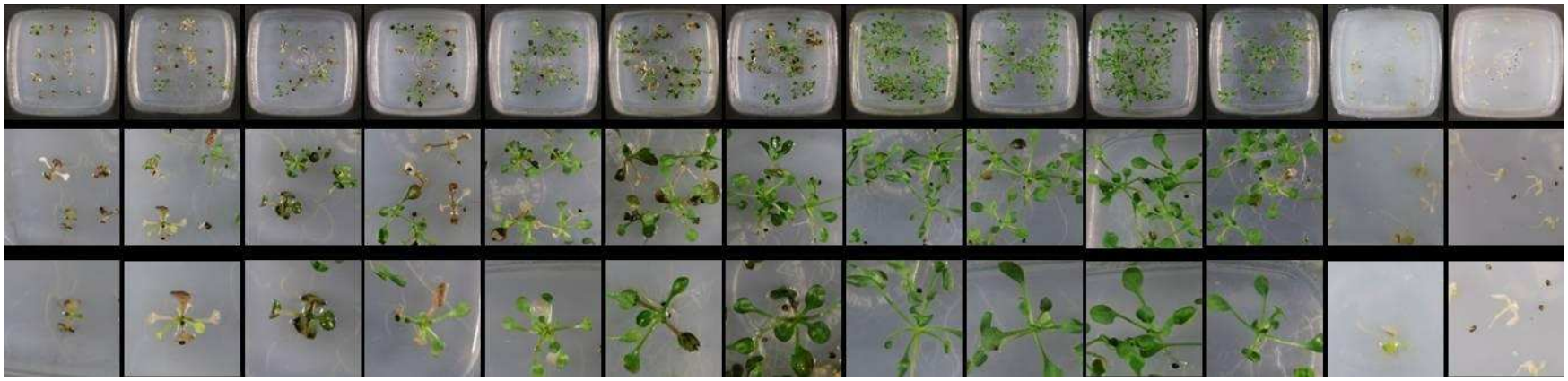


Figure 5.5 45 day old N. caerulescens plants growing on HCL (10%) washed Agar (0.8 g l^{-1}) based growth media supplemented with half strength Murashige and Skoog (MS) nutrient salts and increasing concentrations of $ZnSO_4$. Numbers below images represent concentration of supplemented $ZnSO_4$ in μM. Top row; whole growth boxes, middle row; groups of representative plants, bottom row; single plants. Red bars represent 1 cm.

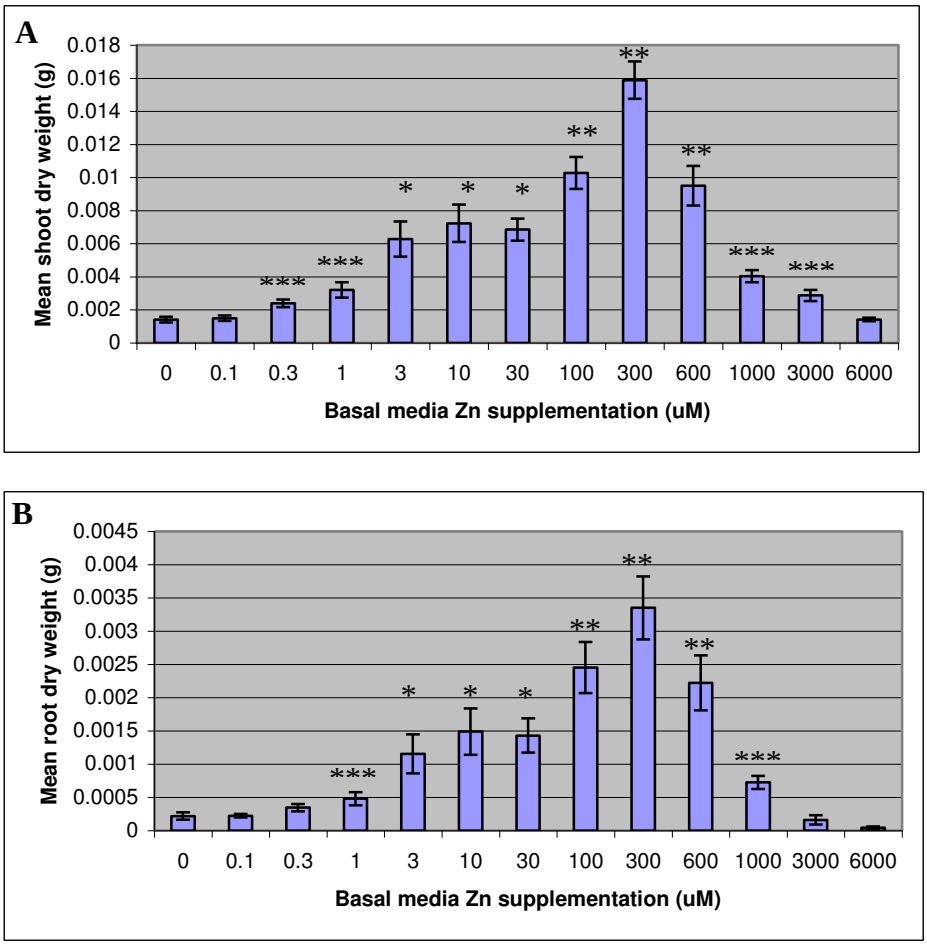


Figure 5.6 Changes in A) mean shoot dry weight (d. wt), and B) mean root d. wt with increasing initial agar Zn concentration. N. caerulescens plants were grown for 45 days on solidified agar supplemented with MS salts. Zn concentrations were obtained by mixing relative amounts of agar containing 6000 µM $ZnSO_4$ and agar containing no $ZnSO_4$. Data were obtained for 15 – 25 plants per treatment for each repeat experiment. Two experiments were performed. Error bars represent standard error of the mean dry weights. Asterisks indicate significant differences: *; $P < 0.001$, **; $P < 0.0001$, ***; $P < 0.05$ according to unpaired Student's t tests (n = 6 per treatment).

5.5.4. Fecundity of Noccaea caerulescens in response to different levels of Zn and cadmium in compost

To efficiently adapt to molecular genetic research, plants, such as Arabidopsis thaliana, must have a standardised reproducible growth system that produces appreciable quantities of seeds. Previous reports have observed that Cd (Lombi et al., 2000) and Zn (Baker et al., 2004) supplementation may have beneficial effects on some N. caerulescens ecotypes' growth and development, however none examined their effects on seed production. To test this, plants grown ex vitro in 0.32 l pots containing Levington M3 compost mix (Section 2.3.1.) were subjected to four Zn (0, 2000, 4000 and 10000 parts per million (ppm) $ZnSO_4$) and four Cd (0, 250,

500, 1000 ppm $CdSO_4$) treatments (Section 5.4.1.3.). A randomised block design comprising three replicates was employed, with 10 individual pots for each of 16 treatments allocated at random within each replicate within a growth tray (n=480) (length 48 cm; width 30 cm; height 5 cm) (Multi Tray, Sankey, UK) (Section 2.3.2.). All Zn and Cd solutions were applied to respective pots on a single occasion after one month of growth.

Three months after sowing, plants were vernalised for 10 weeks (at 5°C, 8 hr photoperiod) under controlled environment (CE) conditions to induce flowering (Section 5.4.1.2.). All plants appeared relatively healthy, however some phenotypic variations were observed. Plants grown on compost not supplemented with Zn or Cd appeared smaller and were more susceptible to disease than other treatments. Those grown on a compost supplementation range from (0 Zn, 500 Cd) ppm – (2000 Zn, 500 Cd) ppm appeared healthy and phenotypically similar to wild type plants under their natural habitat conditions. All plants grown on media supplemented with 4000 ppm $ZnSO_4$ were more stunted and produced more leaves than other treatments. Those supplemented with 10,000 ppm $ZnSO_4$ produced more chlorotic and longer lanceolate leaves. Leaf yellowing was evident for all $ZnSO_4$ treatments supplemented with 1000 ppm $CdSO_4$. For all treatments which contained $ZnSO_4$, mean seed production decreased steadily upon increasing supplementation with $CdSO_4$ (Fig 5.7). Seed production was significantly reduced for plants growing on media without $ZnSO_4$ supplementation than for all $ZnSO_4$ supplemented treatments containing up to 500 ppm $CdSO_4$ ($P < 0.05$). Plants grown under 2000 ppm $ZnSO_4$, but with no additional $CdSO_4$, produced the highest mean number of seeds per plant (3000) (Fig. 5.7). It was evident that under these ex vitro conditions, Noccaea caerulescens reached optimal seed production when supplemented with 2000 ppm $ZnSO_4$, supplying 455 mg Zn kg^{-1} compost.

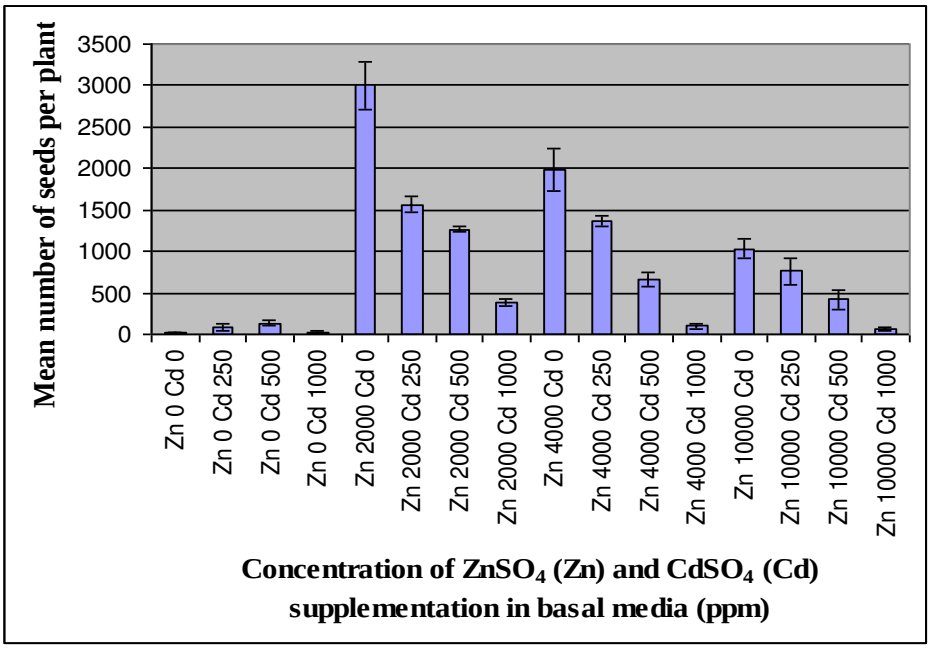


Figure 5.7 Mean number of seeds produced per plants growing on Levington peat based compost. Growth media was supplemented with a matrix of increasing $ZnSO_4$ (Zn) and $CdSO_4$ (Cd) concentrations calculated in parts per million (ppm). Error bars represent standard error of the mean seeds per plant. Asterisks indicates significant differences: *; $P < 0.001$, **; $P < 0.05$ according to unpaired Student's t tests (n = 19 - 45 per treatment).

5.5.5. Fecundity of Noccaea caerulescens in response to increasing concentrations of surfactants in a bacterial floral dip solution

A successful floral dip based transformation approach was described for N. caerulescens by Peer et al. (2003) but has never since been reproduced. Their observations showed that increasing concentrations of the surfactant, Silwet L-77, significantly reduced fecundity and transgenic seed yields in N. caerulescens, however, it was necessary for transformation to take place. To maximise recovery of potentially transformed seeds in N. caerulescens Saint Laurent Le Minier, seed production was recorded from 6 month old flowering plants subjected to floral dipping solutions containing Agrobacterium tumefaciens C58 (resuspended to an optical density (OD_{600nm}) of 0.8) and increasing concentrations of Silwet L-77.

A randomised block design comprising three replicates was employed, with three individual pots for each of 6 treatments allocated at random within each replicate (Section 5.4.3.5.). Five treatments contained Agrobacterium tumefaciens C58 resuspended in 5% (w/v) sucrose solutions supplemented with 0, 0.005, 0.01, 0.02 and 0.05 % (v/v) Silwet L-77, while the negative control contained uniquely 5% sucrose (w/v). Mean seed production per plant was significantly higher for control

plants (dipped in a 5% sucrose solution containing no A. tumefaciens) than for all other treatments ($P < 0.0001$) (Fig. 5.8). Plants dipped in a solution of A. tumefaciens with no Silwet L-77 produced significantly more seed ($P<0.001$) than all plants treated with the surfactant. There was no significant difference ($P<0.05$) in seed production between plants dipped in bacterial solutions containing 0.01% and 0.02% (v/v) Silwet L-77 however both produced significantly more seeds (> 3 fold) than plants dipped in 0.05% Silwet L-77 (Fig. 5.8). Bacterial solutions containing 0.02% (v/v) Silwet L-77 were therefore selected as an optimal concentration for further attempts to transform Noccaea caerulescens.

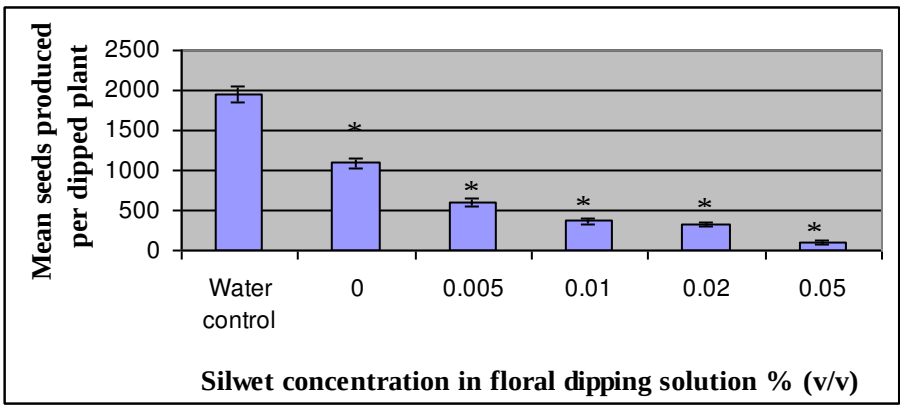


Figure 5.8 Mean number of seeds produced per plant following floral dipping in solutions 5% sucrose (w/v), resuspended Agrobacterium tumefaciens (optical density (OD_{600nm}) of 0.8) and supplemented with increasing concentrations of Silwet surfactant. As negative controls plants were dipped in either a solution of deionised H_2O and sucrose (5% w/v) without (water control) or with (0) resuspended A. tumefaciens (OD 0.8). Error bars represent standard error of the mean seeds per plant. Asterisks indicate significant differences: *; $P < 0.0001$, according to the unpaired Student's t test ($n = 9$ per treatment).

5.5.6. Agrobacterium tumefaciens-mediated transformation of Noccaea caerulescens through floral dip

Floral dipping of approximately 4000 individual N. caerulescens Saint Laurent Le Minier plants took place under darkened conditions with individual flowers from each plant dipped in a solution of Agrobacterium tumefaciens resuspended in 5% (w/v) sucrose solutions supplemented with 0.02% (v/v) Silwet L-77 (Section 5.4.3.5.). Two strains of Agrobacterium tumefaciens harbouring two independent constructs were employed. The first was a supervirulent A. tumefaciens strain 1065 (Curtis et al., 1994) containing a pBIN19 derived plasmid which carried a CaMV35S.gus-intron reporter gene (Vancanneyt et al., 1990) and conferred in

planta kanamycin resistance through a NOS.NPTII.NOS sequence (the GUS::NPTII construct) (donated by Paul Anthony, University of Nottingham) (Section 5.4.3.7.). The second was A. tumefaciens strain C58, containing the RNA interference (RNAi) destination vector Pk7GW1WG2 (II), constructed using a 400 bp fragment from NcHMA4 cDNA and conferring in planta kanamycin resistance (the HMA4RNAi construct) (donated by Victoria Mills, University of Nottingham) (Section 5.4.3.7.). Plants were selected in vitro on agar (10 g l^{-1} sucrose, 0.8 g l^{-1} agar, 2.1 g l^{-1} MS salts and 50 μM $ZnSO_4$) and supplemented with kanamycin sulphate (75 μg ml^{-1}) (Section 5.4.3.6.). No transgenic Noccaea caerulescens Saint Laurent Le Minier were isolated for all T_1 screens for both constructs.

5.5.7. Agrobacterium tumefaciens-mediated transformation of Arabidopsis thaliana through floral dip

To test functionality of both constructs A. thaliana was transformed using the floral dip technique (Section 2.9.). From a total of ~1000 A. thaliana T_1 seeds from T_0 parents dipped with the GUS::NPTII construct, ten plants survived kanamycin screening and 8 demonstrated an amplicon following PCR using primers specific for GUS (Section 2.10.) indicating approximately 0.8% efficiency (Fig. 5.9). From a total of ~1000 T_1 seeds from T_0 parents transformed with the HMA4RNAi construct, 13 survived kanamycin screening and three amplified NPTII fragments using NPTII specific primers (Section 2.10.) demonstrating 0.3% transformation efficiency (Fig. 5.10). Both lines were selfed to homozygosity in T_3 (Fig. 5.8). Both constructs therefore were shown to be functional, thus it was concluded that further refinements were necessary to establish an efficient system of Agrobacterium tumefaciens-mediated transformation in Noccaea caerulescens. Moreover, other methods of mutagenesis were available to recalcitrant genomes including fast neutron bombardment.

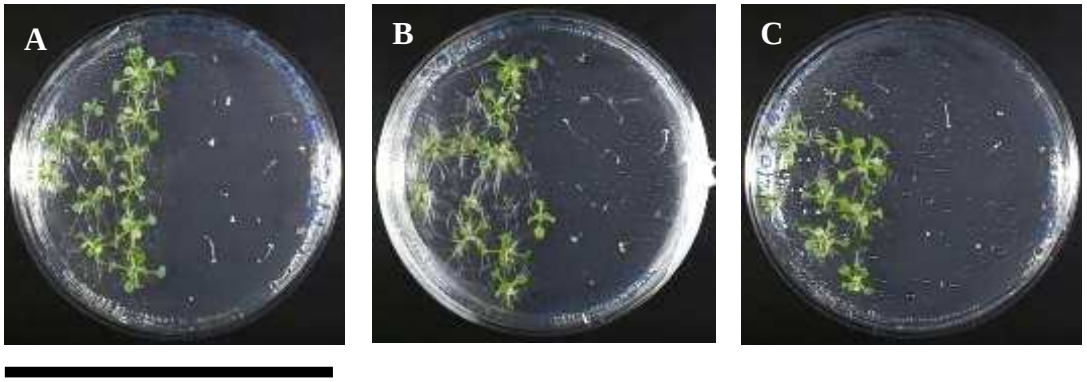


Figure 5.9 Kanamycin tolerance screen for homozygous A. thaliana T_3 lines A, B, and C transformed with Nc-Ganges-HMA4RNAi construct. Each Petri dish contained transformed lines either **A**, **B** or **C** (left half) and WT accession (right half). Each dish contained 20ml of agar based growth medium supplemented with 50 µg ml^{-1} Kanamycin sulphate. Scale bar represents 10cm.

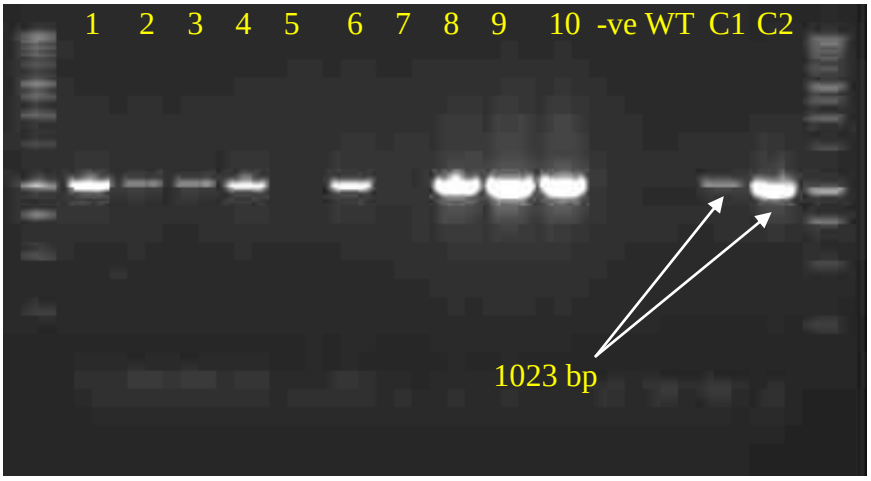


Figure 5.10 PCR amplified products from individual samples of putative transgenic A. thaliana 'Colombia'. Numbers indicate lanes. Lanes 1-10 contained individual PCR amplifications from T_2 A. thaliana 'Colombia' which survived in vitro kanamycin screening. Negative controls were Wild type Arabidopsis DNA, (WT) and no template (-ve). C1 and C2 contained DNA from and two positive controls for GUS (pVDH65 and pAHC25) respectively. Ladder = 1kbp. Arrows highlighted GUS fragment = 1023 bp. Agarose = 1% (w/v).

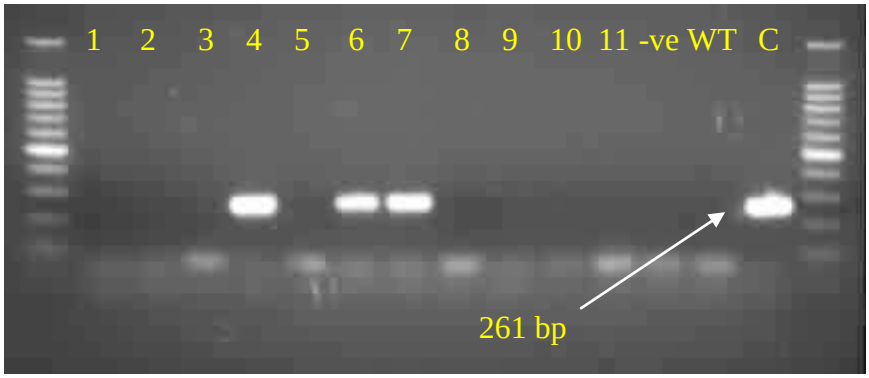


Figure 5.11 PCR amplified products from individual samples of T_2 Nc-Ganges-HMA4RNAi::NPTII transgenic A. thaliana 'Colombia'. Lanes 1-11 contained individual PCR products from A. thaliana 'Colombia' that survived in vitro kanamycin screening. Negative controls were Wild type Arabidopsis DNA, (WT) and no template (-ve). 'C'contained DNA from Pk7GW1WG2 (II) and was a positive NPTII control. Ladder = 1kbp. Arrow highlighted NPTII fragment = 261 bp. Agarose = 1% (w/v).

5.5.8. Fast neutron mutagenesis of Noccaea caerulescens Saint Laurent Le Minier

A population of 5500 M_0 seeds were subjected to fast neutron mutagenesis in order to create faster cycling lines of Noccaea caerulescens Saint Laurent Le Minier (Ó Lochlainn et al., 2011). Seeds were irradiated at the Budapest Research Reactor (BRR) inside a Cd capsule (Section 5.4.3.1.). During the irradiation real time dose monitoring was performed and the irradiation was terminated after 137 min, at which point the required dose 438 milliGrays per minute (mGy min^{-1}) was delivered. The measured surface gamma dose from the seeds was 130*background dose which reduced to <2*background dose upon dispatch (where background dose is ~90 nGy h^{-1}).

5.5.8.1. Testing for faster cycling Noccaea caerulescens mutant lines

From the initial 5500 M_0 seeds irradiated with fast neutrons at 60 Gy, resulting M_1 plants were grown under glasshouse conditions using standard procedures for N. caerulescens, including a 10-week period of vernalisation (Section 5.4.1.). Approximately 80% M_1 seeds germinated, with 2% showing signs of leaf colour variegation and 79% of plants survived to maturity (Ó Lochlainn et al., 2011). Morphologically, plants appeared more heterozygous with 5% producing abnormally few but large leaves or conversely, producing many small tightly bunched leaves. Faciation was visible in approximately 1% of flowering stems. Approximately 80,000 M_2 seeds were maintained in a single pool at an average of 25 seeds per M_1

plant. M_2 seeds were grown initially in modules (Fig. 5.12) and transplanted to 0.32 l pots (height 7.9 cm; diameter 9 cm) under glasshouse conditions (Fig. 5.12). The M_2 plants were screened for early flowering phenotypes with no vernalisation requirement (Fig. 5.13). A total of 0.49% M_2 seedlings demonstrated lethal albinism (Fig. 5.12). Floral initials were observed in 35 individuals in the absence of vernalisation. Of these, nine individual plants flowered within 12 weeks, producing an average of 100 M_3 seeds per selfed plant (Fig. 5.13B). One selfed M_2 individual, 'A2', produced ~800 M_3 seeds ($A2M_3$). Two of these nine plants ($A2M_3$ and $A7M_3$) were selfed, again without vernalisation, to produce $A2M_4$ and $A7M_4$ seeds respectively. These two lines were compared for flowering and mineral uptake traits with an S_2 wild-type (WT) line from the original N. caerulescens Saint Laurent Le Minier population.

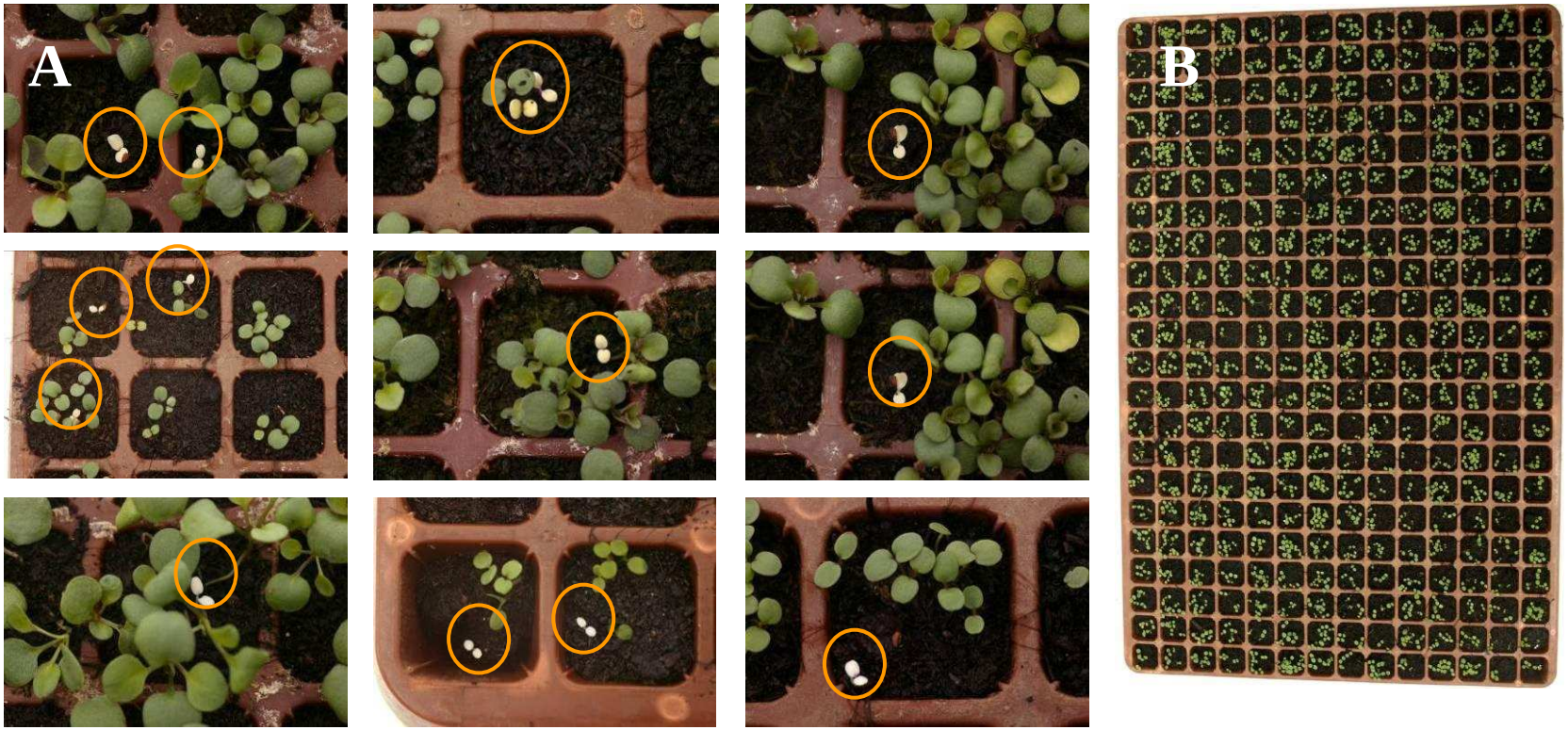


Figure 5.12 **A;** A sample of M_2 fast neutron mutagenised N. caerulescens seedlings which demonstrated lethal albinism (Orange circles) and **B**; M_2 seedlings growing in a 308 cell plug tray. Individual cells measured $1cm^2$.

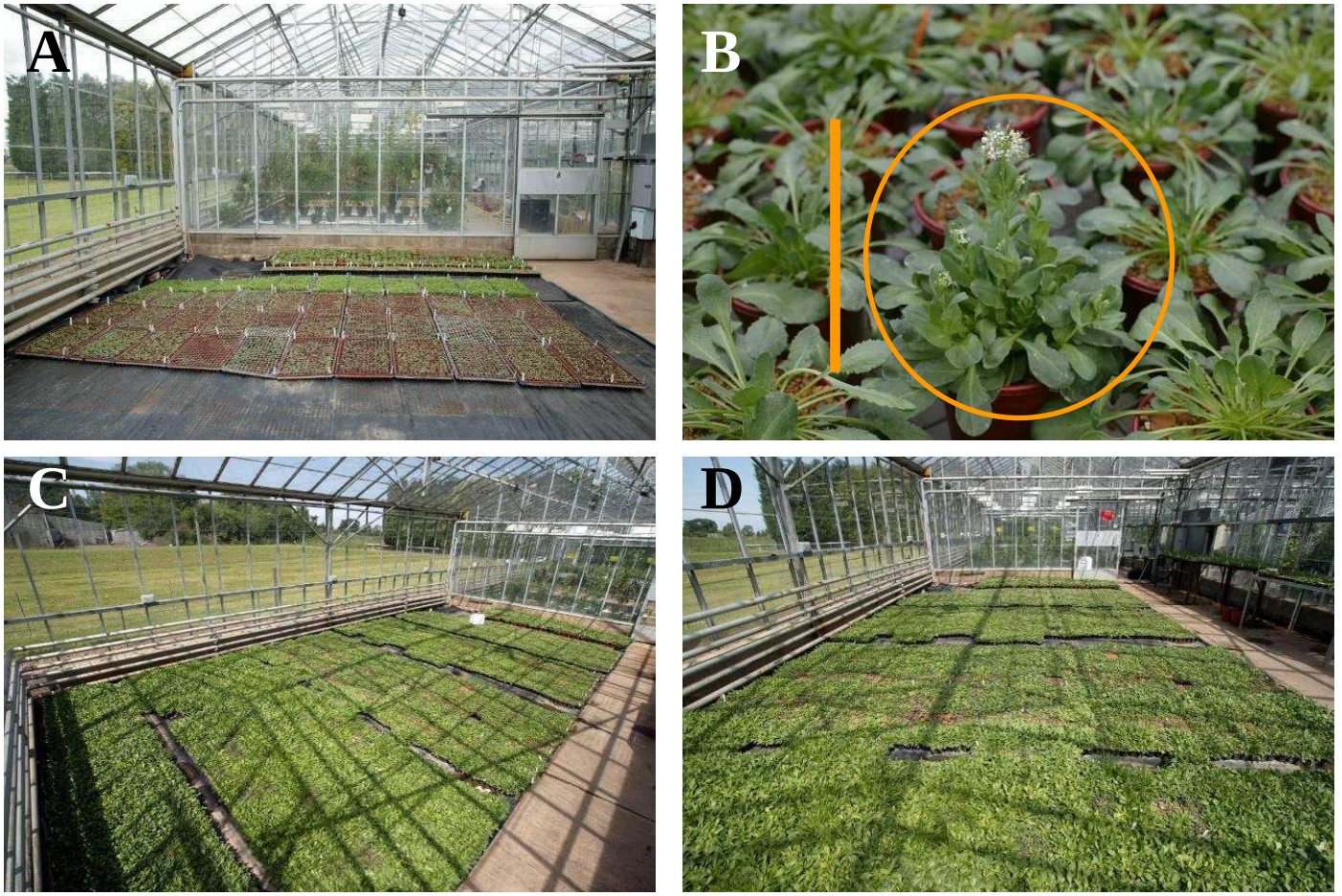


Figure 5.13 **A;** M_2 fast neutron mutagenised N. caerulescens seedlings growing in 308 cell trays under glasshouse conditions. **B**; 12 week old, pre-vernal early flowering M_2 plants in the glasshouse. Scale bar represents 15 cm. **C & D**; A subset of the 80,000 transplanted M_2 plants growing under glasshouse conditions

5.5.8.2. Characterising M_4 fast neutron mutagenised Noccaea caerulescens lines

Lines of $A2M_4$, $A7M_4$ and the S2 WT were sown in modules and transplanted to 0.32 l pots under CE conditions (Section 5.4.3.3) at 7 days after sowing (DAS). Germination in module trays was >98% for all lines by 7 DAS. By 66 DAS, all $A2M_4$ plants had developed floral initials, by 71 DAS all $A2M_4$ plants had unopened flower buds, and by 79 DAS all $A2M_4$ plants had fully opened flowers (Fig. 5.14). The $A7M_4$ flowering was approximately 3 weeks slower than $A2M_4$. Thus, by 87 DAS all $A7M_4$ plants developed floral initials, by 97 DAS all had unopened flower buds, and by 104 DAS all had fully open flowers. Silicle development was well-established for all $A2M_4$ and the majority of $A7M_4$ individuals by 92 and 123 DAS, respectively. No floral or silicle development was observed in any of the S_2 WT plants at these dates (Ó Lochlainn et al., 2011).

Wild type Noccaea caerulescens produced 32% and 23% more leaf biomass than the $A2M_4$ and $A7M_4$ lines respectively however, there was no significant difference in leaf tissue dry weights between the two faster cycling lines (Fig. 5.15). Mineral analysis of dried leaf tissue demonstrated that both $A2M_4$ and $A7M_4$ rapid cycling mutant lines contained similar leaf Zn concentrations to the WT, which were in the range >0.3% Zn on a DW basis (Fig. 5.16 A). This indicates that the hyperaccumulation phenotype (Reeves & Brooks, 1983; Broadley et al., 2007), was retained. Thus, when grown with the addition of 455 mg Zn kg^{-1} to the compost, WT, $A2M_4$ and $A7M_4$ plants accumulated 0.34, 0.33 and 0.35 % Zn on a DW basis respectively (Fig. 5.16 A). However, WT, $A2M_4$ and $A7M_4$ plants differed in leaf concentrations of other minerals, including the macronutrients Mg, Ca and, the micronutrient Cd, which were all typically higher in WT (Ó Lochlainn et al., 2011). Leaf Mg concentrations were 0.42, 0.31 and 0.37%, leaf Ca concentrations were 0.80, 0.61 and 0.69% on a DW basis, while leaf Cd concentrations were 1.1, 0.5 and 0.8 mg kg^{-1} DW; in WT, $A2M_4$ and $A7M_4$ plants respectively (Fig. 5.16 B-D). These variations might be due to phenological differences between lines (Nord and Lynch, 2009; Ó Lochlainn et al., 2011).

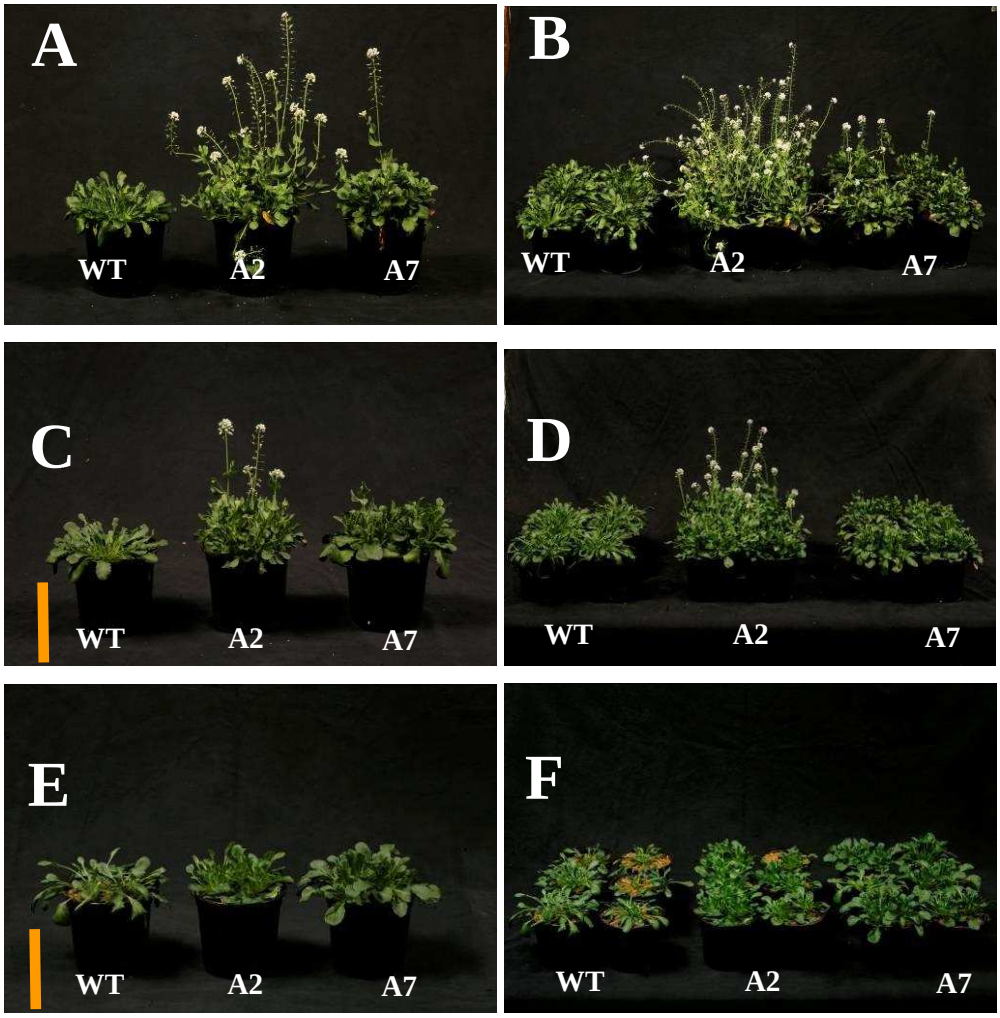


Figure 5.14 N. caerulescens lines M_4A2 (A2) and M_4A7 (A7) and wild type (WT) growing under controlled environmental conditions after 36 (E-F) 62 (C-D) and 78 (A-B) days. Left column; individual plants per line, Right column; groups of six plants per line. Scale bar represents 15 cm.

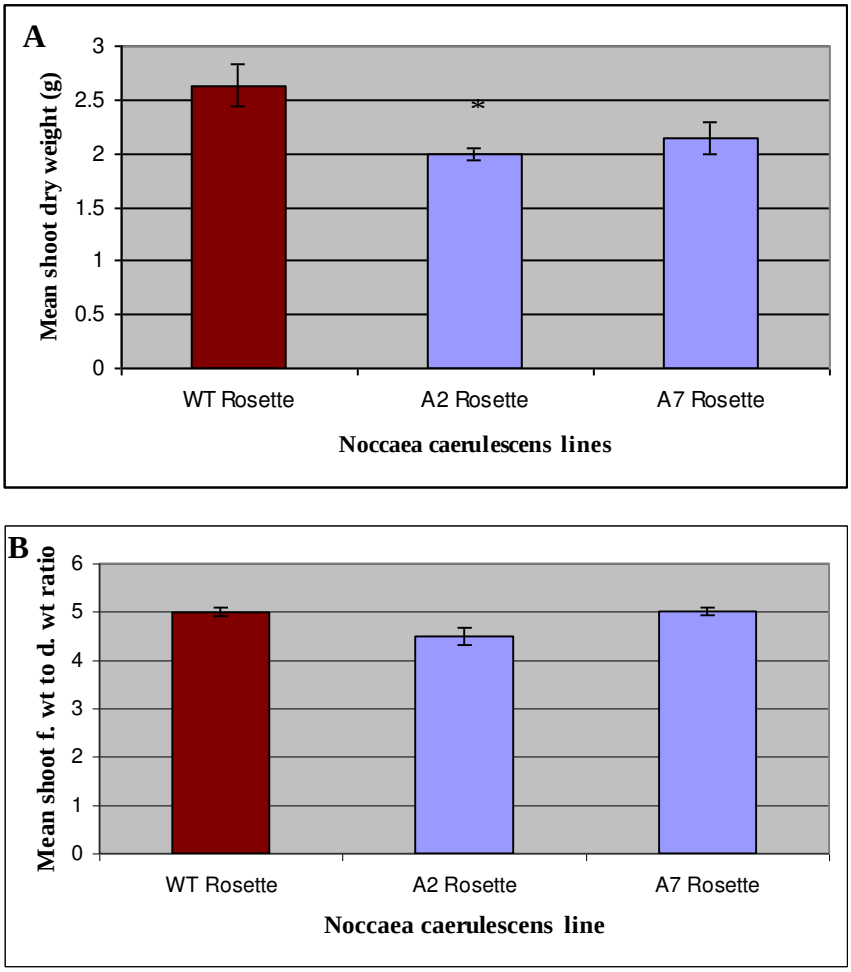


Figure 5.15 **A;** Mean shoot d. wt and **B;** Mean fresh weight to dry weight ratio of Noccaea caerulescens lines M_4A2 (A2) and M_4A7 (A7) and wild type (WT) growing under controlled environmental conditions for 123 days. Error bars represent standard error of the mean d. wt per plant. Asterisk indicates significant difference: *; $P < 0.01$, according to the unpaired Student's t test (n = 36 per line).

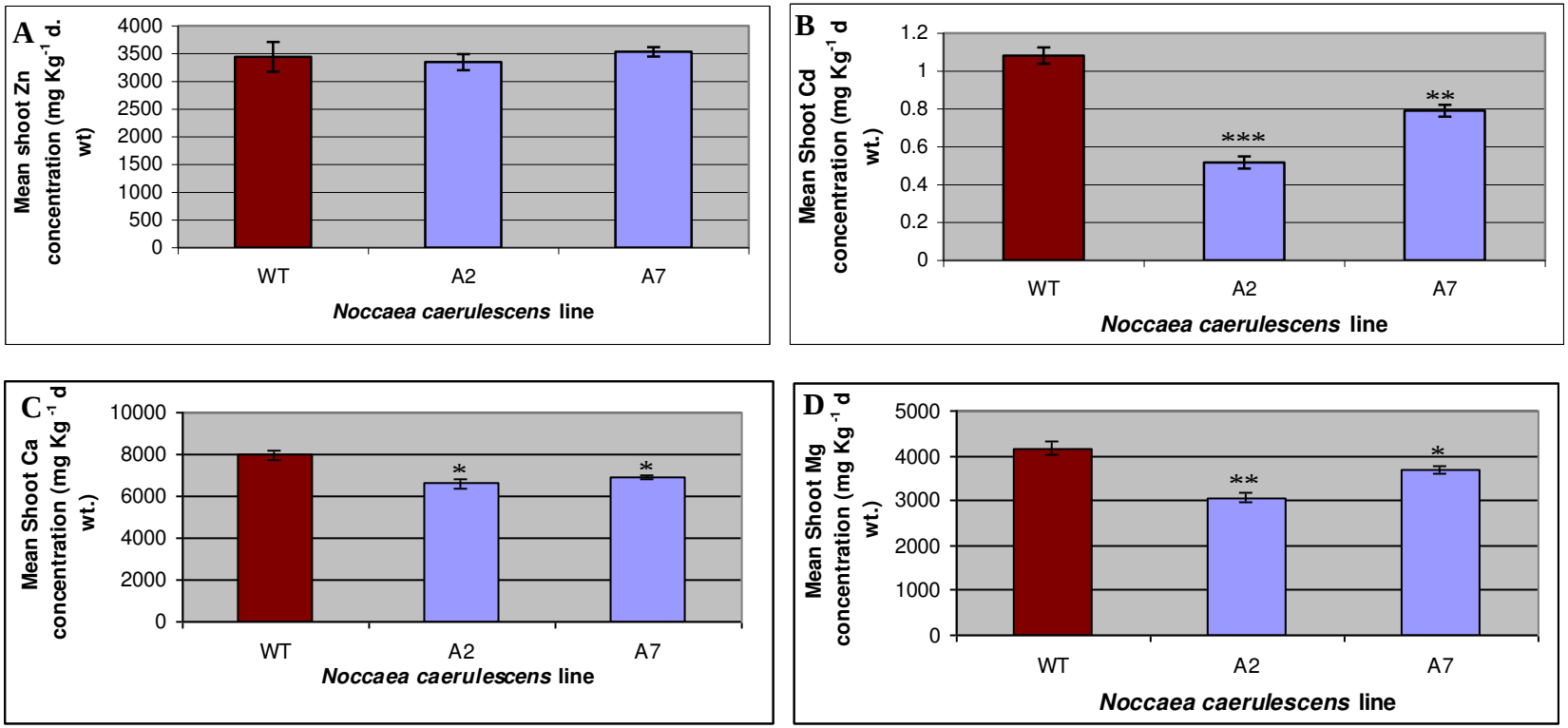


Figure 5.16 Mean leaf concentrations (mg mineral Kg^{-1} dried shoots) of Zn (Panel A), Mg (B), Ca (C) and K (D) in rosette leaves of Noccaea caerulescens lines M_4A2 (A2), M_4A7 (A7) and S_2 wild type (WT) growing under controlled environmental conditions for 123 days. Error bars represent standard error of the mean concentration per plant. Error bars represent standard error of the mean dry weights.
Asterisk indicates significant differences: *; $P < 0.05$, **; $P < 0.005$, ***; $P < 0.001$, according to unpaired Student's t tests (n = 36 per line).

5.6. DISCUSSION

This chapter has reported the establishment of genetically-stable faster cycling lines of Noccaea caerulescens Saint Laurent Le Minier using fast neutron mutagenesis, to support future forward and reverse molecular genetic studies. These lines should offer a rapid platform to perform direct functional analysis of Noccaea candidate genes involved in Zn hyperaccumulation which have been identified through sequencing (Chapter 3) and heterologous expression (Chapter 4) studies. In vitro growth conditions have been optimised to support such molecular analysis of both N. caerulescens and A. thaliana in response to increased exogenous supplies of Zn. Ex vitro growth conditions have similarly been optimised to ensure maximum seed yield for N. caerulescens, thus supporting techniques requiring maximal fecundity including transformation based approaches which are currently being optimised. Together these results provide tremendous support and an important platform to further elucidate the genetics surrounding Zn hyperaccumulation in N. caerulescens. To perform such molecular genetic analyses it is important to establish a robust growth system with minimal environmental variables. It was therefore decided to test for Zn-dependent growth effects on N. caerulescens under in vitro conditions.

5.6.1. Agar growth media may contain background levels of Zn.

Plant growth responses in response to exogenous Zn supplies were tested on an in vitro agar-based medium in order to establish an efficient system for future molecular genetic analyses. The standard concentration of Zn in growth media used by the majority of researchers in Arabidopsis is 15 μM Zn supplied in its sulphate form ($ZnSO_4$) and usually combined with other essential minerals as described by Murashige and Skoog (1962). Nutrient solutions were made according to the Murashige and Skoog (1962) (MS) protocol and their growth effect on plants in vitro were compared to plants supplemented with equivalent concentrations of a standard commercial MS nutrient salt (Sigma M5524, Sigma, Germany). Mean dry weight measurements revealed no significant differences between conditions on the growth of Arabidopsis ($P < 0.05$). Next, growth effects of a Zn deficient media was examined by adding no $ZnSO_4$ and supplementing with the remaining essential elements at half MS strength. Plants showed no significant difference in dry weight ($P < 0.05$). It was hypothesised that background Zn contamination existed in the agar. HCL was applied to agar and mixed overnight to displace bound Zn^{2+}. To test whether this

regime had any phenotypic or morphologic effect on plants, samples were grown on agar which was mixed overnight in either a 10% (v/v) solution of HCL (HCL wash) or in deionised H_2O (H_2O wash). One half of each agar treatment was supplemented with all essential elements (Zn replete MS) while the other was supplemented with all except $ZnSO_4$ (Zn deficient MS). Following 14 d growth, no significant difference was detected between the morphology or dry weights of plants growing on Zn replete MS for both HCL and H_2O wash treatments. Plants grown on agar from the H_2O wash treatment and supplemented with Zn deficient MS were significantly similar to both Zn replete treatments. Conversely, plants grown on HCL wash treated Zn deficient MS demonstrated stunted growth, with unusually small chlorotic leaves and small weak roots. These symptoms were indicative of Zn deficiency symptoms in plants (Marschner, 1995). Thus for all subsequent Zn deficiency growth response experiments, agar was treated with a 10% HCL wash prior plant culture.

5.6.2. Growth response for Arabidopsis thaliana in response to exogenous Zn

An exogenous Zn $[Zn]_{ext}$ dependent growth response curve was established for A. thaliana under in vitro conditions to complement downstream heterologous molecular genetic analyses. Previous studies have shown that A. thaliana did not hyperaccumulate Zn or display any Zn hypertolerance traits, however none have demonstrated a comprehensive $[Zn]_{ext}$ growth response across a wide range of Zn levels. To test for this, plants were grown in vitro for 21 d on increasing levels of Zn supplied to agar medium containing replete concentrations of all remaining essential elements. Mean shoot d wt. was significantly lower for plants grown on agar supplemented with 0 μM $ZnSO_4$ (0 μM $[Zn]_{agar}$) than those growing on $[Zn]_{agar}$ ranging from 0.1 to 100 μM, and contained less than half the shoot d wt. of those grown on 30 μM $[Zn]_{agar}$. There were no significant differences between mean shoot d. wt. among plants growing on $[Zn]_{agar}$ ranging from 0.1 to 30 μM. Mean shoot d wt. showed a significant negative correlation upon agar supplementations of 100 and 300 μM $ZnSO_4$, with shoot d. wt. reducing by 33% and 66% respectively from those grown on 30 μM. Mean root d wt demonstrated a positive correlation with $[Zn]_{agar}$ supplementations ranging from 0 – 30 μM, before reducing by 9% under 100 μM $ZnSO_4$ and sharply by 99% under 300 μM $ZnSO_4$. Phenotypically, shoot tissue appeared chlorotic and stunted when grown on $[Zn]_{agar}$ ranging from 0 – 3 μM. Plants growing on 100 μM demonstrated a reduction in root growth and plant size

although there was no visual difference in shoot colour. Upon supplementation with 300 μM $ZnSO_4$, plant leaf tissue became chlorotic and along with root tissue suffered severe stunting. It was therefore concluded that in terms of both root and shoot physiology for Arabidopsis thaliana under in vitro conditions, Zn deficiency and toxicity symptoms could be robustly recorded at $[Zn]_{agar}$ of 0 and 300 μM respectively while Zn repletion was evident at 30 μM.

5.6.3. Growth response for Noccaea caerulescens in response to exogenous Zn

Noccaea caerulescens growing for 45 d on agar demonstrated a clear positive correlation between $[Zn]_{agar}$ and plant shoot and root d. wt. between 0 μM to 300 μM $[Zn]_{agar}$. Shoot d. wt. was significantly higher than all other treatments at 300 μM $[Zn]_{agar}$. Increasing $[Zn]_{agar}$ above 300 μM resulted in a negative correlation between plant root and shoot d. wt. and $[Zn]_{agar}$. On $[Zn]_{agar}$ from 3000 to 6000 μM Noccaea displayed severe morphological toxicity symptoms including chlorosis, blanching and stunting and eventual death. The severity of necrosis, leaf darkening, chlorosis and root and shoot stunting worsened as $[Zn]_{agar}$ decreased from 300 to 0 μM, thus correlating with decreasing shoot and root d. wt. for these treatments. From these growth observations it was demonstrated that Zn deficiency and toxicity in roots and shoots of Noccaea were found at 0 and 6000 μM. For N. caerulescens, Zn repletion occurred at 300 μM under these conditions. Downstream analyses may require plants under less extreme conditions depending on the mode of assay since many individuals die and tissue samples are extremely small under these conditions.

Noccaea caerulescens demonstrated remarkably higher tolerance to increasing concentrations of $[Zn]_{agar}$ when compared to the Zn non-hyperaccumulator Arabidopsis thaliana. Zn replete levels were 10 fold higher than that of A. thaliana. Conversely Noccaea showed symptoms of Zn deficiency under conditions which were Zn sufficient for A. thaliana, thus illustrating a high Zn requirement for optimal growth. These results are consistent with previously published observations, where N. caerulescens was shown to increase shoot fresh weight (f. wt.) at $[Zn]_{agar}$ up to 300 μM before a reduction in f. wt. was recorded (Hammond et al., 2006). In this same study a more closely related non-hyperaccumulator, Thlaspi arvense, demonstrated a reduction in shoot f. wt. as $[Zn]_{agar}$ increased above 30 μM, resembling A. thaliana in the current study. Haydon and Cobbett (2007) observed a reduction of over 40% in f. wt of 17 d old shoots of A. thaliana grown in 60 μM

$[Zn]_{agar}$ compared with those in 1 μM $[Zn]_{agar}$. Results here confirmed previous reports of N. caerulescens hypertolerance to extremely high external Zn concentrations (Broadley et al., 2007; Verbruggen et al., 2008; Krämer et al., 2010) and provided a basis for future molecular genetic assays. Similarly, it was necessary to maximise seed production in N. caerulescens to compliment such assays. An ex vitro growth trial was thus attempted to determine optimal conditions for maximum fecundity in this species.

5.6.4. Maximising the fecundity of Noccaea caerulescens

Zinc supplementation in peat based Levington M3 growth media resulted in improved fecundity in Noccaea caerulescens Saint Laurent Le Minier. It had been reported previously that metallicolous populations of N. caerulescens which hyperaccumulate Zn also require Zn to maintain optimal growth (Broadley et al., 2007; Verbruggen et al., 2009). Moreover ecotypes from southern France including Ganges, which encompassed the Saint Laurent Le Minier accession, required additional Cd for optimal growth outside of their natural habitat (Roosens et al., 2003; Liu et al 2008). As for all model plant species, exact cultural conditions for optimal growth must be elucidated to complement downstream molecular and physiological trials under both in vitro and ex vitro conditions.

Results observed in the current study demonstrated that the addition of up to 2,000 ppm $ZnSO_4$ (455 mg Zn kg^{-1} dry compost mix) resulted in the maximum mean number of seeds (3,000) produced per plant. For all conditions supplemented with $ZnSO_4$ the addition of increasing concentrations of $CdSO_4$ to composts ($[Cd]_{soil}$) negatively correlated with seed yield, suggesting that this ecotype did not require additional Cd in compost mixes under these conditions to yield optimal seeds. However a relatively small but noticeable positive correlation between seed yield and $[Cd]_{soil}$ was observed for treatments containing no $ZnSO_4$ supplementation. Morphologically plants displayed leaf yellowing under treatments supplemented with 2,000 ppm, 4,000 ppm and 10,000 ppm $ZnSO_4$ + 500 and 1,000 ppm $CdSO_4$ matrix. Leaf yellowing could be a symptom of iron (Fe) deficiency (Marschner, 1995). One explanation for this apparent Zn dependent improved fecundity could be that metalliferous N. caerulescens evolved a requirement for elevated $[Zn]_{soil}$ in order to grow normally (Escarré et al., 2000). This response could be a result of the metalliferous population's ability to reduce Zn uptake as compared to non-

metalliferous populations (Ernst et al., 1990) and to their Zn sequestration mechanism (Lasat et al., 1996) which may result in a Zn deficiency for metabolic processes in normal $[Zn]_{soil}$ levels (Shen et al., 1997). Zn is known to be an essential element in plant reproduction with increases in the number of inflorescences and seed yield in subterranean clover observed following Zn applications (Alloway, 2004). Equally, Brown et al. (1993) observed that Zn deficient wheat developed small anthers and aberrant pollen grains suggesting that Zn played a vital role in sexual development in plants. One of the primary advantages to an efficient ex vitro reproduction system is its application to numerous detailed physiological and molecular genetic approaches including transformation techniques. We therefore tested the efficacy of a commonly used 'floral dip' technique to transform this species.

5.6.5. T-DNA insertion transformation of Noccaea caerulescens

Although still in progress, currently no single T_0 T-DNA transformant was isolated following attempted transformation of 4000 N. caerulescens Saint Laurent Le Minier plants. Two unique T-DNA constructs, NcHMA4RNAi and NPTII::GUS were employed in a modified Noccaea-specific floral dip approach (Peer et al., 2003). As positive controls, both constructs were transformed into A. thaliana Col-0 by floral dip (Clough and Bent, 1998) and successfully selected in vitro on agar containing 50 $\mu g\ ml^{-1}$ kanamycin sulphate and selfed to T_3. Presence of the transgenes from both constructs was confirmed by PCR. The Saint Laurent Le Minier ecotype has never previously been transformed and may be recalcitrant to such techniques. A similar observation was observed for other Noccaea ecotypes such as N. caerulescens Le Bleymard (a Zn and Cd hyperaccumulating population also from southern France) which also appeared recalcitrant to these techniques (Peer et al., 2003). Moreover, within the model plant species Arabidopsis thaliana, transformation efficiencies using the 'floral dip' approach vary considerably between ecotypes (Clough and Bent, 1998). Subsequent transformation events within the Noccaea genus, using the 'floral dip' approaches have never been published. Relatively recently, a tissue culture based approach reported some transformation success in N. caerulescens (Guan et al., 2008), however this technique is lengthy and unsuitable to rapid molecular genetic methods. It has been demonstrated that Agrobacterium tumefaciens mediated 'floral dip' transformation in Arabidopsis was dependent on

the presence of unique apertures in the developing gynoecium (Desfeux et al., 2000). Such morphology may not be present in N. caerulescens making this system unsuitable as a mutagenic approach in this genus. Other methods available to mutate genomes include insertional mutagenesis delivered by Fast Neutron bombardment. Thus, such an approach was attempted in Noccaea caerulescens to remove its 10 week vernalisation requirement to initiate flowering and thus reduce its 7 – 9 month life cycle (Peer et al., 2003; 2006).

5.6.6. Rapid cycling Noccaea caerulescens

This section discusses the generation of genetically stable faster cycling lines of N. caerulescens which were evidently non vernal-obligate (Ó Lochlainn et al., 2011). A total of 5500 M_0 seeds were irradiated at 60 Gy, and subsequent M_1 plants were selfed to M_2 in a single pool. 80,000 M_2 plants were screened for early flowering phenotypes with no vernalisation requirements. A total of 0.49% M_2 seedlings demonstrated lethal albinism. This level of albinism was considerably less than the 2% reported in Arabidopsis which represented approximately 10 induced deletions per line (Koornneef et al., 1982) and 2.57% in Medicago (Rodgers et al., 2009). Extrapolating these comparisons to other species is complex and confounded by species-specific effects of the dose (Gy) as well as variations in the genetic basis of this phenotype (Rodgers et al., 2009). Similarly it has been shown that a 50% M_1 survival rate was a reasonable balance between mutagenesis and fertility however in the present study M_1 Noccaea showed approximately 63.2% survival (Rodgers et al., 2009). Although not attempted here, a more comprehensive approach might be through analysis of tiling arrays to give a more robust estimation of size distribution and number of induced deletions per line (Rodgers et al., 2009). Floral initials were observed in 35 individuals in the absence of vernalisation. Of these, nine individual plants flowered within 12 weeks, producing an average of 100 M_3 seeds per selfed plant. One selfed M_2 individual, 'A2', produced ~800 M_3 seeds ($A2M_3$). Two of these nine plants ($A2M_3$ and $A7M_3$) were selfed, again without vernalisation, to produce $A2M_4$ and $A7M_4$ seeds respectively. These two lines were compared for flowering and mineral uptake traits with an S_2 wild-type (WT) line from the original population. There was no significant difference in germination rates between $A2M_4$, $A7M_4$ and the S2 WT under CE conditions. By 66 days after sowing (DAS), all $A2M_4$ plants had developed floral initials, and by 79 DAS all $A2M_4$ plants had fully

opened flowers. The $A7M_4$ flowering was approximately 3 weeks slower than $A2M_4$. Together this meant that by 14 weeks after sowing (WAS) all M_4 fast neutron mutagenised lines were fully flowering and by 17 WAS silicle development was well established. For both lines this represented a reduction in cycle time by ~53% and a complete removal of any vernalisation requirement to initiate flowering (Ó Lochlainn et al., 2011). For $A2M_4$ in particular, all plants developed silicles by 13 WAS and so demonstrated a reduction in cycle time of ~64%. No floral or silicle development was observed in any of the S_2 WT plants at these dates. Wild type Noccaea caerulescens produced 32% and 23% more leaf biomass than the $A2M_4$ and $A7M_4$ lines respectively. However, there was no significant difference in leaf tissue dry weights between the two faster cycling lines. Mineral analysis of dried leaf tissue demonstrated that both $A2M_4$ and $A7M_4$ rapid cycling mutant lines contained similar leaf Zn concentrations to the WT, which were in the range >0.3% Zn on a DW basis. This indicates that the hyperaccumulation phenotype was retained (Broadley et al., 2007; Ó Lochlainn et al., 2011). Thus, when grown with the addition of 455 mg Zn kg^{-1} to the compost, WT, A2M4 and A7M4 plants accumulated 0.34, 0.33 and 0.35 % Zn on a DW basis, respectively. However, for the majority of the remaining essential plant elements, WT plants accumulated higher concentrations in their leaf tissue than both A2M4 and A7M4 plants e.g. Ca and Mg levels in leaves of $A2M_4$ and $A7M_4$ plants were between 75 – 85% those of WT plants (Ó Lochlainn et al., 2011). Particularly noticeable was the reduction in Cadmium (Cd) leaf concentrations in $A2M_4$ and $A7M_4$ plants which contained < 50% and < 75% respectively of the concentration found in WT leaves. This observation would appear consistent with previous population studies which demonstrated that hyperaccumulation of Zn and Cd in Noccaea was under independent genetic control (Zha et al., 2004; Deniau et al., 2006). Variations in mineral concentrations might be due to induced phenological differences between the lines (Nord and Lynch, 2009). They concluded that phenology affected the acquisition of minerals by plants from the soil, and that a compression in phenology, such as early maturation, would likely reduce this acquisition. Similar observations were reviewed in White and Broadley (2009). Phenotypically, both mutant lines were stable and there was no evidence of significant intra-line variability in flowering, growth or Zn accumulation. Neither A2 nor A7 displayed seed dormancy. The A2 and A7 lines had clearly lost the requirement for vernalisation to initiate flowering whilst remaining self fertile. From

a preliminary back cross experiment to WT Noccaea lines, there was no evidence to date that this non-vernalisation trait was dominant. Both mutant lines exhibited much more rapid flowering and seed maturation phenotypes than any WT grown under our conditions. These lines significantly reduced the period currently required to cultivate vernal-obligate WT N. caerulescens (Peer et al., 2003, 2006), enabling production of up to four generations of seed in a single year (Ó Lochlainn et al., 2011). Crucially, both lines have retained the Zn hyperaccumulator phenotype, and so have considerable potential for establishing further molecular genetic insights, especially when efficient transformation systems and full genome sequence become available. If faster cycling lines of other N. caerulescens ecotypes can be similarly developed, there is scope for establishing additional mapping populations and introgression lines to facilitate locus resolution of interesting traits including Zn hyperaccumulation. As found with rapid-cycling Brassicas, the elimination of a vernalisation requirement greatly accelerates the ability to resolve traits introgressed from a wide range of germplasm, including subsequent selection for reduced time to flowering and seed maturation (Williams & Hill, 1986; Iniguez-Luy et al., 2009; Broadley et al., 2010). It may also be possible to further mutate lines A2 and A7 to produce even faster cycling lines in the future. The effects of fast neutron bombardment on these faster cycling N. caerulescens lines at a molecular level have not yet been elucidated. Previous observations in a range of plant genomes have associated loss of function in FLOWERING LOCUS C (FLC) and FRIGIDA (FRI) orthologues with the removal of vernalisation and induction of early flowering phenotypes in A. thaliana (Michaels and Amasino, 1999; Sung and Amasino, 2004), with conserved roles for FLC homologues observed in Beta vulgaris (Reeves et al., 2007), Arabis alpina (Wang et al., 2009); Brassica rapa (Zhao et al., 2010). Thus it may be that fast neutron bombardment of these lines had resulted in a loss of function or downregulation of these orthologues in N. caerulescens. It may be possible to test this hypothesis using high throughput transcriptome sequencing or DNA hybridisations to tiling or exon arrays designed for Arabidopsis (Mockler et al., 2005) or Brassica (Love et al., 2010) using heterologous- (cross-) species based approaches (Broadley et al., 2008). However, further selfing, backcrossing and complementation will most likely still be required since the mutational load has not yet been elucidated. The experiments described here have resulted in the creation of previously unreported genetically-stable, fast cycling, vernalisation-independent

lines of N. caerulescens through fast neutron mutagenesis. This commonly applied approach to mutation breeding has been shown to create random deletions ranging from one base to > 100 kb (Kodym & Afza, 2003; Salt et al., 2008; Bruce et al., 2009). It has proved to be a relatively inexpensive method of producing large mutant populations and could be applied to species whose genomes are not amenable to T-DNA transformation, thus generating genome-wide saturation in relatively small populations. Here it has been demonstrated that fast neutron mutagenesis is a viable approach to develop mutant populations within this exceptional hyperaccumulating genus and, it is anticipated that these lines will become a valuable community resource for future molecular genetic investigations into metal tolerance and hyperaccumulation. Agrobacterium tumefaciens mediated transformation of wild type N. caerulescens is still in progress and due to time constraints such work has not been attempted on these fast cycling lines. It is anticipated that the research community will continue these efforts to create a fully transformable system within these lines to gain a more sophisticated understanding of the current data surrounding the molecular genetic controls of Zn, and other metal, hyperaccumulation in this species (Chapters 3 and 4). It is hoped that such information can be translated into closely related crop systems such as the commercially important Brassica spp. There is an immense wealth of mineral accumulation as well as quantitative trait loci (QTL) data from large mapping populations within this genus, however elucidating the genetic mechanisms has proved challenging (Chapter 6). However, progress is gaining momentum with several sequencing projects in progress including the Brassica rapa Genome Sequencing Project (http://www.brassica.info/resource/sequencing.php) which should have a fully assembled genomic sequence by the end of 2010. It is hoped therefore that molecular data generated from N. caerulescens (Chapters 3 and 4) as well as this novel fast cycling resource will hasten progress in understanding natural genetic variation in Zn accumulation in Brassicaceae.

5.7. SUMMARY

- Complete plant growth responses to exogenous Zn concentrations in growth media were determined for both Noccaea caerulescens and Arabidopsis thaliana under in vitro conditions.
- Noccaea carulescens showed significantly higher tolerance to increased growth media Zn ($[Zn]_{media}$) concentrations than Arabidopsis thaliana.
- Noccaea demonstrated Zn deficiency symptoms under $[Zn]_{media}$ conditions which were >100 fold greater than that of Arabidopsis.
- A robust protocol was established to reduce Zn contamination from stocks of agar-based growth media.
- Optimal fecundity for a range of Zn and Cd ex vitro supplementations to compost-based growth medium was established for Noccaea caerulescens.
- Floral dip-based transformation approaches are still in progress in Noccaea.
- A robust protocol for developing a fast neutron mutant population of Noccaea caerulescens was established.
- Fast cycling M_4 fast neutron mutant Noccaea lines form fruit within 92 days after sowing compared with 7 – 9 months for wild type (WT) N. caerulescens lines.
- It is now possible to reproduce approximately 4 generations of fast cycling Zn hyperaccumulating Noccaea caerulescens mutant lines in one year.
- Fast cycling M_4 fast neutron mutant Noccaea lines require no vernalisation to initiate flowering compared with a 10 week vernal dependency in WT plants.

CHAPTER 6 REVERSE GENETIC APPROACHES TO UNDERSTANDING ZINC ACCUMULATION IN BRASSICA SPP.

6.1. INTRODUCTION

Innovations in functional molecular analyses have been instrumental in elucidating natural genetic variation in Zn accumulation in Brassicaceae plant aerial tissues ($[Zn]_{leaf}$). Such advances included transgenic approaches in Arabidopsis thaliana which associated HMA4 with $[Zn]_{leaf}$ accumulation (Mills et al., 2003, 2005; Hussain et al., 2004; Verret et al., 2004, 2005). Supporting this, transcriptome profiling demonstrated that HMA4 orthologues were more highly expressed in the Zn hyperaccumulators, A. halleri and Noccaea caerulescens than in related non-hyperaccumulators (Hammond et al., 2006; Talke et al., 2006). Novel sequencing and transformation events in A. halleri revealed that increased $[Zn]_{leaf}$, through enhanced Zn flux from the root symplasm to xylem vessels, was a result of high HMA4 expression, which was specified in cis and amplified by gene triplication (Hanikenne et al., 2008). This was consistent with the present study, where novel sequencing approaches identified a quadruplication of HMA4 in N. caerulescens (Chapter 3) and promoter expression profiles (Chapter 4) were similar to observations for A. halleri. However, HMA4 has not yet been isolated as a potential Zn transporter in the closely related and commercially important crop Brassica spp. Nevertheless, analysis of 376 Brassica oleracea accessions (representing 99% of the common allelic polymorphisms in this species), showed $[Zn]_{leaf}$ varied 26 fold (Broadley et al., 2010). Through quantitative trait loci (QTL) forward genetic analyses, significant genetic variation in Zn accumulation was observed in Brassica rapa (Wu et al., 2007, 2008) and B. oleracea (Broadley et al., 2010). Heritability for $[Zn]_{shoot}$ was significant for both mapping populations with 18.5% for Brassica oleracea (Broadley et al., 2010) and 69% for B. rapa ssp. pekinensis (Wu et al., 2008). Frustratingly, such forward genetic approaches are unsuitable for systematic genome wide gene analysis due to the enormous scale and efforts required for their establishment (Rogers et al., 2009). Additionally, these approaches can be lengthy and often require more than a year to isolate locus specific genes. Furthermore, high sequence identity between paralogous sequences can hinder identification despite increased availability of whole-genome sequences, increasing mapped

polymorphisms and improvements in genotyping technologies (Bruce et al., 2009). However, reverse genetic approaches, such as Targeting Induced Local Lesions in Genomes (TILLing), could be applied to elucidate gene function through phenotypic analysis following induced gene mutations (Colbert et al., 2001; Till et al., 2003; Stephenson et al., 2010). These approaches are becoming increasingly efficient, since faster and improved sequencing technologies have increased the availability of novel gene sequences (Gilchrist and Haughn, 2010; Stephenson et al., 2010). Coupled with this, candidate gene identification from Zn QTL in both B. oleracea and B. rapa has been facilitated with over 1000 genetically linked Restriction Fragment Length Polymorphism (RFLP) loci in B. napus (an allotetraploid of B. rapa and B. oleracea) (Rana et al. 2004) mapped to homologous positions within A. thaliana through comparative sequence analysis (Parkin et al., 2005). Using a combination of comparative genomics with existing forward genetic datasets, it should therefore be possible to identify candidate locus specific genes involved in Zn accumulation and genotype their allelic variation in planta.

This chapter examined the hypothesis that induced allelic variants of locus specific candidate genes involved in Zn accumulation in Brassica rapa could be efficiently identified and genotyped in planta. To test this, candidate genes were identified using a combination of comparative genomic analysis of existing Brassica datasets and functional evidence from closely related Brassicaceae species including A. thaliana, A. halleri and N. caerulescens. Sequences were elucidated using in silico data-mining of all publicly available Brassica sequences. Allelic variants of these genes were screened from M_2 DNA from a TILLing population of B. rapa mutated with ethyl methanesulfonate (EMS) (Stephenson et al., 2010). Subsequent M_3 plants were genotyped using novel molecular techniques combined with high resolution melt (HRM) analysis.

6.1.1. Brassica rapa as a crop for reverse genetic analysis

Brassica spp. is considered a model crop genus due to its extensive and easily accessible genetic variation, and adaptive capabilities to an immense diversity of agro-climatic conditions (Choi et al., 2007). Its monogenomic, diploid AA (B. rapa), BB (B. nigra) and CC (B. oleracea) and allopolyploid, AABB (B. juncea), AACC (B. napus) and BBCC (B. carinata) genomes contain numerous morphologically

distinct but economically important varietas representing a large proportion of global vegetable crop, and over 12% of global edible vegetable oil production (Choi et al, 2007; Broadley et al., 2008; Broadley et al., 2010).

Brassica spp. have a number of advantages over most crop species since they are the closest crop relative to the model plant A. thaliana, having both diverged from a common Brassicaceae ancestor ca. 20 million years ago (Ma), considerably more recent than other crop species such as tomato (150 Ma, ~950Mb) and rice (200 Ma, ~400Mb) (Wolfe et al., 1989; Yang et al., 1999). Subsequent to genome triplication, Brassica oleracea and B. rapa diverged ca. 4 Ma (Inaba and Nishio, 2002) and despite disparity in chromosome number, (B. rapa L., 2n=2x=20, ~550Mb) and (B. oleracea L., 2n=2x=18, >600Mb), both genomes demonstrate immense collinearity (Lagercrantz and Lydiate, 1996; Parkin et al., 2005). Comparative analyses conducted by Rana et al. (2004) confirmed that the microstructure of B. rapa was vastly similar to that of B. oleracea and both essentially contained a triplicated structure of an A. thaliana-like genome. Both displayed a very similar though not identical pattern of interspersed gene loss, and all conserved genes displayed highly conserved collinearity between the three genomes (Rana et al., 2004). In addition there are over 95 published genetic maps within the Brassica genome made publicly available through the CropStore database (http://www.brassica.info/CropStore/populationslinkedphp?pop=BnaTVSL). Over three million Brassica spp. genomic sequences, expressed sequence tags (EST) and bacterial artificial libraries (BAC) have been deposited in GenBank and are currently publicly available (http://www.ncbi.nlm.nih.gov/genbank/). A Brassica Ensemble database (BrassEnsemble) (http://www.brassica.info/BrassEnsembl/index.html) has been established as a fully searchable database for integrating and interrogating diverse Brassica data. The B. rapa Chiifu-401genome is presently being fully sequenced and assembled as part of the Multinational Brassica Genome Sequencing Project (http://www.brassica.info/info/about-mbgp.php) with currently over 130,000 BAC clone end-sequences publically available in addition to 968 fully sequenced BACs. Several independent sequencing efforts have already deposited completed sequences in fully searchable formats including a 12-fold coverage of the rapid cycling B. rapa R-O-18 yellow Sarsons genotype (http://www.brassica.info/resource/sequencing.php). In addition, this genotype has formed the genetic basis for the reverse genetic 'TILLing' population (Stephenson et al., 2010). It was therefore

decided to elucidate locus specific gene sequences for this genotype for future reverse genetic assays.

6.1.2. Targeting Induced Local Lesions IN Genomes (TILLing) as a method for reverse genetics in Brassica rapa

Targeting Induced Local Lesions IN Genomes (TILLing) was first developed as a reverse genetics approach to identifying gene function in plants (McCallum et al., 2000; Colbert et al., 2001). The basis of this approach relies on detection of mismatches between annealed wild type and mutant DNA strands (Till et al., 2003). Mutations are induced by ethyl methanesulphonate (EMS) exposure which produces random mutations in the nucleotide sequences by alkylation on the O^6 position of guanine, resulting in GC to AT transitional changes (Till et al., 2003). A total of 96 positions can therefore be mutated in the genetic code and of these, 33 do not result in an amino acid change (silent mutations), 58 lead to 26 amino acid substitutions (mis-sense mutations) and 5 result in stop codons (nonsense mutations) (Stephenson et al., 2010). Unlike the majority of other mutagenesic approaches EMS can therefore produce both loss of function and gain of function mutants leading to null, hypomorphic, neomorphic or hypermorphic phenotypes (Gilchrist and Haughn, 2010). Since it is desirable to detect such potential regions, an analysis software, CODDLE, is employed to efficiently identify areas within the gene which have the highest probability of altered gene function when mutated with EMS (Till et al., 2003). Once identified, approximately 1 kb fragments of these regions are PCR amplified from pooled DNA using fluorescently labelled primers and mismatched heteroduplexes are then generated with wild type DNA. These heteroduplexes are incubated with CELI (a plant endonuclease which cleaves at heteroduplex mismatched sites) and the results are viewed on a sequencer such as ABI3730 before subsequent individual plant DNA analysis (Till et al., 2003). Within the Brassica spp., Brassica rapa R-O-18 was selected for this approach since it had a similar development ontology to oil seed rape crops (B. napus) and, as a diploid, had greater potential for functional analysis of redundant genes compared to amphidiploid relatives (B. napus and B. juncea) (Stephenson et al., 2010). Seeds were treated with 0.3% and 0.4% ethyl methanesulphonate (EMS) to generate a population of 6,912 and 2,304 M_2 plants respectively with random point mutations.

DNA from all 9,216 M_2 plants was pooled, with each pool comprising 8 DNA samples, to form the TILLing platform (Stephenson et al., 2010). As an initial trial, TILLing analysis of six genes revealed a mutant density of approximately one per 60 kb or ~10,000 mutations per plant, with mutations in approximately half of the GC base pairs in the genome. It was estimated that screening 3,072 M_2 plants would produce an average of 68 mutations and a 97% probability of obtaining a missense mutation and therefore protein truncation (Stephenson et al., 2010). RevGen UK (http://revgenuk.jic.ac.uk/) provide a TILLing service for B. rapa R-O-18 and following successful identification of mutant M_2 DNA, respective M_3 seeds are dispersed for further genotyping using efficient methods such as high resolution melt (HRM) analysis.

6.1.3. High resolution melt (HRM) as a method for genotyping mutant Brassica rapa R-O-18 lines

Techniques to genotype mutant lines containing single nucleotide polymorphisms (SNP) have rapidly developed over recent years, however many are expensive, inefficient to process or require bespoke apparatus. Sanger sequencing is robust and provides comprehensive detail, however it is relatively slow and expensive for large mutant populations. Similarly Taqman SNP typing requires bespoke software for maximum efficiency and is prohibitively expensive (Martino et al., 2010). Cleaved amplified polymorphic sequence (CAPS) marker assays are limited to regions where SNP induction creates a recognition sequence for specific restriction endonucleases (Parsons et al., 1997). High resolution melt (HRM), however has been shown to be a robust, precise, efficient and inexpensive method to detect SNPs in genomes (Garrirano et al., 2009; Martino et al., 2010). It monitors the real time dissociation of double stranded nucleotide fragments, PCR amplified with primers labelled with intercalating dyes that fluoresce strongly when bound to double stranded DNA. As fragments dissociate (melt apart) in response to increasing temperature, the presence of double stranded DNA is decreased and thus fluorescence is decreased. A light scanner (e.g. LightScanner® instrument, Idaho Technology, Salt Lake City, Utah, USA) records and plots this data graphically on a melt curve illustrating fluorescence levels as a function of temperature. Disparity between melt curves from different plants indicate alternative melting temperatures of DNA sequences as a result of SNP mutations and therefore represent allelic variation for that sequence (Martino et al.,

2010). By identifying such differences between melt curves from putative mutant M_3 Brassica rapa R-O-18 TILLing lines and homozygous wild type lines, this technique should provide a fast and efficient method to select lines for further molecular analysis.

This chapter describes the efficient identification, isolation and genotyping for locus specific allelic variants of candidate Zn hyperaccumulator sequences through comparative genomics, novel in silico approaches and classical in vivo techniques from a Brassica rapa R-O-18 TILLing population to facilitate future functional analyses.

6.2. AIMS

The aim of this chapter was to develop an efficient method to test for, isolate and screen for allelic variants in locus specific Brassica rapa candidate zinc transporter sequences from a population of previously generated EMS mutagenised M_2 B. rapa R-O-18.

6.3. OBJECTIVES

- To test for independent paralogues of Brassica rapa HMA4, CAX1 and ESB1 sequences through in silico data mining and PCR confirmation approaches.
- To develop an efficient DNA extraction and PCR protocol for large population screens.
- To efficiently screen B. rapa M_2 and M_3 lines for accessions containing mutations in a single CAX1 locus using a novel nested PCR approach for High Resolution Melt (HRM) SNP detection.

6.4. MATERIALS AND METHODS

6.4.1. Plant materials

Rapid cycling Brassica rapa lines screened from the TILLing population (Stephenson et al., 2010) and wild type lines (donated by Dr. Graham King, Rothamsted Research, U.K.) were derived from B. rapa ssp. oleifera cv. R-0-18. All Plants were generated by ex vitro seed germination. All seeds were stored under darkness at 18 - 20ºC at 30 – 40% relative humidity (RH).

6.4.1.1. Brassica rapa culture

Brassica rapa R-O-18 were cultured under glasshouse conditions with a 16 h photoperiod at 22.3 ±4°C and 13.3 ±2°C mean day and night temperatures respectively (Section 2.3.2.). Plants were sown directly into 1.05 l pots (height 11.3 cm; diameter 13 cm) containing 1 l of Levington M3 high nutrient compost (Monro Group) (Section 2.3.1.) with moisture levels maintained twice daily by overhead irrigation. Following three weeks of vegetative culture, Vitafeed® 2-1-4 nutrient solutions (N-P-K: 16-8-32 + micronutrients) (Vitax Ltd., Coalville, Leicestershire, UK) were applied weekly to plants at a rate of 3 g l^{-1} via overhead irrigation. Plants initiated flowering after approximately 6 weeks of culture and seeds were collected approximately 4 weeks thereafter. Perforated bread bags (150mm x 700mm, W R Wright & Sons Ltd, Walton, Liverpool, UK) were employed to enclose inflorescences to prevent pollination from neighbouring plants.

6.4.2. Comparative analysis of Brassica spp. quantitative trait loci (QTL) involved in zinc-dependent traits

Comparative analysis of published quantitative trait loci (QTL) involved in altered Zn dependent traits (Zn-QTL) in B. oleracea and B. rapa were employed to identify candidate Zn transporter genes associated with altered shoot Zn levels $[Zn]_{shoot}$.

For B. oleracea QTL were identified from a mapping population generated by anther culture of the F_1 of a cross between a DH rapid-cycling accession (B. oleracea var. alboglabra; ‘A12DHd’) and a DH accession derived from an F_1 hybrid cultivar of calabrese (B. oleracea var. italica cv. ‘Green Duke’; ‘GDDH33’ (Sebastian et al., 2000) referred to as the AGDH population (Broadley et al., 2010). A 906 cM linkage

map with a mean distance between marker loci of 1.92 ± 3.49 cM was employed, such that ca 90% of the genome was <5 cM from a marker (Sebastian et al., 2000; Broadley et al., 2008). Nine replicates of 90 AGDH lines and both parents were grown under glasshouse conditions (GE2) in 1 l pots containing compost (Shamrock medium grade sphagnum peat; Scotts, Ipswich, UK) supplemented with 5.25 mg l^{-1} external phosphate (low $[P]_{ext}$) or 15.75 mg l^{-1} (high $[P]_{ext}$) (Broadley et al., 2010). Plant shoots were sampled for 50, 50 and 34 d after sowing (DAS) from three temporally-sequential blocks for fresh and dry weights and mineral analysis using inductively coupled plasma emission spectrometry ICP-AES (Broadley et al., 2010). Twelve replicates of 62 AGDH lines with parents growing under field conditions (FE2) supplemented with four $[P]_{ext}$ treatments equivalent to 0, 298, 1,125, and 2,713 Kg of triple superphosphate ($Ca(H_2PO_4)_2$) ha^{-1} were similarly sampled 105 DAS (Broadley et al., 2010). Data from all experiments were analysed using REML procedures (GenStat 9.1.0.147; VSN International, Oxford, UK). Accession means were estimated and sources of variation allocated to environmental and genetic components for individual experiments, using variance components models defined in Broadley et al. (2008). For GE2 and FE2, QTL mapping was performed using composite interval mapping (CIM) in QTL Cartographer 2.0 (Broadley et al., 2010).

For B. rapa, QTL were identified from a mapping population derived from a cross between two DH Chinese cabbage lines. Y177 originated from a winter type Japanese cultivar ‘Jianchun’ and Y195 derived from a summer type Chinese cultivar ‘Xiayang’ referred to as the BrIVFhn population. A 1,090 cM linkage map was employed, which represented a mean of one marker per 3.8 cM. Two replicates of 142 BrIVFhn lines and parents were grown under field conditions in a randomized two-block design. Plant shoots were sampled 30 d after transplantation (DAT) from two blocks for fresh and dry weights and mineral analysis using inductively coupled plasma-atomic emission spectrometry ICP-AES (Wu et al., 2008). Three replicates of 140 BrIVFhn and parental lines were grown hydroponically under glasshouse conditions in a half strength Hoagland’s nutrition solution (pH 5.5), buffered with 2 mM MES (2-morpholinoethanesulphonic acid) supplemented with deficient, normal or toxic Zn concentrations corresponding to 0, 2 and 100 µM $ZnSO_4$ respectively (Wu et al., 2008). Shoots were sampled 21 DAT and oven dried for 3 d at 65°C for shoot dry biomass (SDB). JoinMap version 3.0 was employed for marker

segregation and linkage analysis while linkage groups (LGs) were determined using a LOD thresholds ≥4.0 and a maximum recombination fraction of 0.4. (Wu et al., 2008). QTL were identified and located on linkage maps using using interval mapping (IM) and multiple-QTL model (MQM) mapping methods from MAPQTL version 5.0 by Wu et al., (2008).

QTL involved in Zn-dependent traits (Zn-QTL) in both B. oleracea and B. rapa were compared to the published B. napus (an allotetraploid of B. rapa and B. oleracea) genetic map containing over 1000 genetically linked loci (Parkin et al., 2005). Data were analysed to identify QTL in similar regions within syntenic chromosomes of B. oleracea and B. rapa for all Zn-QTL. These were located to homologous positions in the Arabidopsis thaliana genome using the published mapped positions from this B. napus map (Parkin et al., 2005). Using a combination of literature reviewing and gene ontology (GO) searches for zinc transport in NCBI (http://www.ncbi.nlm.nih.gov/), candidate A. thaliana genes were identified with associations in Zn transport to shoot tissues in these loci.

6.4.3. Isolation of locus specific fragments of Brassica rapa gene sequences

The Brassica rapa genome was not yet fully sequenced or assembled, despite the imminent arrival of a completely assembled genomic sequence from the multinational Brassica Genome Sequencing Project (BrGSP). An in silico comparative sequence analysis approach was thus adopted to identify locus specific regions for PCR amplification.

6.4.3.1. In silico data mining to identify gene sequences

Basic Local Alignment Search Tool (BLAST) software was employed to identify sequences from all Brassica spp. which were orthologous to Arabidopsis thaliana genes HMA4, CAX1 and ESB1 (Section 2.8.2.1.). All A. thaliana sequences of interest were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/). For the BLASTn algorithm (http://www.ncbi.nlm.nih.gov/blast/producttable.shtml#tab31) the entire GenBank database was interrogated to identify all Brassica sequences sharing > 80% sequence identity at an e-value of $<1e^{-30}$ with the A. thaliana query sequence. For the WuBLAST algorithm (http://brassica.bbsrc.ac.uk/), the entire Brassica Genome Sequencing Project (BrGSP) database was interrogated for all

Brassica BAC-end sequences, fully sequenced BACs, ESTs and GSS which shared > 80% sequence identity with this query sequence. A. thaliana genes sharing high sequence similarities to the query gene i.e. members of the same gene family were similarly used to interrogate both databases. After in silico comparative sequence analyses (Section 2.8.2.2.), Brassica sequences were assigned to the member of the A. thaliana gene family with which they shared the highest sequence identity, thus removing false positives. All selected sequences were formatted into FASTA and entered into a contig assembly program (ContigExpress) at default settings to form longer contiguous Brassica sequences and identify possible locus specific paralogues (Section 2.8.2.3.). These nascent Brassica gene sequences were aligned to the Arabidopsis thaliana gene of interest, flanked at either end with 1 Mbp of sequence, using AlignX software at default settings (Section 2.8.2.2.). The genomic structure of these contigs was elucidated and annotated in the Vector map software (VectorNTI 11, Invitrogen, Paisley, UK). Primers were designed to amplify ~ 1.2 Kbp fragments beginning upstream from the deduced transcriptional start site to a region within the gene using Primer 3 Version 0.4.0 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) with the parameters "Max Self Complementarity" and "Max 3' Self Complementarity" adjusted to 2, in order that hairpin loops and dimerisations were avoided.

6.4.3.2. Amplification of locus specific sequences from Brassica rapa R-O-18

DNA was extracted from leaves of wild type B. rapa R-O-18 using the GenElute® Plant Genomic DNA Miniprep kit (Section 2.6.1.) and subjected to PCR amplification using primers designed to Brassica contig sequences produced from in silico comparative analyses. Phusion™ high fidelity DNA polymerase was employed to ensure the integrity of amplified sequences (Section 2.7.3.). Amplified fragments were electrophoresed and gel purified (Section 2.6.3.), A-tailed (Section 2.5.6.1.) and cloned into Escherichia coli DH5α bacteria (Section 2.5.6.2.). Colony PCRs (Section 2.5.6.4.) were performed on colonies of bacteria selected on solid LB supplemented with spectinomycin (100 µg ml^{-1}) (Section 2.5.6.4.). All clones demonstrating amplicons of the correct size were grown for 16 h in liquid LB supplemented with spectinomycin (100 µg ml^{-1}) at 37°C before plasmid extraction (Section 2.5.7.) for Sanger-based sequencing (Section 2.8.). Sequences from all clones were examined to ensure all were identical and therefore confirm in vivo locus specificity of primers.

6.4.4. Targeting Induced Local Lesions IN Genomes (TILLING) as a method to screen mutant Brassica rapa R-O-18 population

A population of 9,216 M_2 ethyl methanesulphonate (EMS) induced mutant lines were created and deposited as a publicly available genomic resource managed by RevGen UK (http://revgenuk.jic.ac.uk/) (Stephenson et al., 2010). DNA from all lines was pooled, with each pool comprising 8 DNA samples, to form the TILLing platform (Stephenson et al., 2010).

6.4.4.1. Using Codons Optimised to Discover Deleterious Lesions (CODDLE) to define optimal regions for mutation analysis

Sequences of B. rapa R-O-18 genes were sent to RevGen UK for analysis using the Codons Optimised to Discover Deleterious Lesions (CODDLE) software (http://www.proweb.org/coddle/) at default settings to identify regions where GC to AT base pair transitions were more likely to occur. It also deduced whether transitions would result in intronic splice variants, introduced nonsense, missense or silent amino acid changes and therefore possible hypomorphic, neomorphic or hypermorphic phenotypes. Primers were subsequently designed (with T_m of between 60 – 70°C) to amplify these regions and their locus specificity was confirmed using the methods described (Section 6.4.3.2.).

6.4.4.2. TILLing the mutant population

The entire TILLing procedure for the B. rapa R-O-18 mutant population was performed by RevGen UK. Aliquots of primers were separated to fluorescently label the forward primer strand with FAM and the reverse with HEX fluorescent dyes. These primers were mixed in a molar ratio of 3:2 (labelled : unlabelled) to produce final concentrations of 0.2 µM. PCR cycling was performed using a G-Storm (Gene Technology Limited, Essex, UK) 96-well thermalcycler. Reaction volumes of 12.5 µl with ExTaq polymerase (Takara Bio Inc, Shiga, Japan) and 5 µl of genomic DNA (0.5 ng μl^{-1}) from pools of 8 M_2 plants were employed. An aliquot of 5 µl of PCR was cleaved with CEL1 by incubating at 45 °C for 15 min. Supplementation of 5 µl of 150 mM EDTA stopped the reaction and the DNA was precipitated upon addition of isopropanol before resuspending in 9.9 µl of Hi Di Formamide and 0.1 µl MRK1000 (Gel Company, San Francisco, California, USA) Rox size standard. The

sample was subsequently heat denatured for 5 min and loaded on to an ABI 3730 (Applied Biosystems, Foster City, California, USA) using a 96-capillary 50 cm array with POP7. Cleavage products resulting from CEL1 digestion of heteroduplex mismatches recorded by the ABI sequencer output were analysed with GENEMAPPER 4.0. PCR amplifications of candidates from the detected pools of eight were sequenced and analysed using Mutation Surveyor software (SoftGenetics, State College Pennsylvania, USA) to identify individuals with mutations and possible protein changes.

6.4.5. Genotyping segregating CAX1 mutant Brassica rapa lines

6.4.5.1. High-throughput DNA extraction protocol

A 0.25 cm^2 diameter leaf disc from a small unexpanded Brassica leaf was placed in a 96-well plate (Qiagen Collection Microtubes, Qiagen, Crawley, UK). Samples were stored on ice during the collection and placed at -20°C for longer term storage. Once all samples were collected, a single 3 mm stainless steel ball and 600 µl of room temperature DNA extraction buffer (1 M Tris-HCL (pH 7.5), 4 M NaCl, 0.5 M EDTA and 10% w/v SDS) were added to each well. Suspensions were shaken for 3 min in a ball mill until the majority of tissue was disrupted. Plates were subsequently centrifuged in a plate centrifuge (Centrifuge 5804, Eppendorf UK, Cambridge, UK) at 5,600 g for 20 min. Using 1,250 µl long narrow pipette tips (Star Labs, Milton Keynes, UK), 300 µl of supernatant was transferred to a fresh plate (Thermofisher storage plate with rubber sealing mat, Fisher, Loughborough, UK), containing 300 µl of 100% isopropanol (Propan-2-ol 99.5+% HPLC, Fisher, Loughborough, UK), mixed and incubated for 3 min at room temperature (RT). Plates were then centrifuged in a plate centrifuge at 5600 g for 20 min before discarding the supernatant through aspiration. 100 µl of 70% (v/v) ethanol was added and mixed by shaking vigorously for 5 min, centrifuged at 5,600 g for 20 min. The supernatant was discarded by aspiration and the plates were dried in a vacuum manifold for 10 min. DNA was resuspended in sterile distilled H_2O (SDW) (20 µl) for 2 hrs at 4°C.

6.4.5.2. High Resolution Melt (HRM) analysis

Primers were designed for High Resolution Melt (HRM) analysis to amplify 100 to 250 bp regions within the BraA.CAX1.a region which was screened by TILLing (Section 6.4.4.2.). Eight sets of HRM primer pairs were designed using Primer 3

Version 0.4.0 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) with the parameters "Max Self Complementarity" and "Max 3' Self Complementarity" adjusted to 2, and a $T_m \geq 60°C$. in order that hairpin loops and dimerisations were avoided. Primer sequences were examined further by primer thermodynamic software from Vector NTI 11 (Invitrogen, Paisley, UK) to confirm that no secondary structures formed. All eight primer pairs amplified overlapping regions containing SNPs in the BraA.CAX1.a TILLing locus. HRM primer amplification efficiencies and specificities, were determined in vivo by PCR amplification of wild type B. rapa R-O-18 DNA in reaction volumes containing 10.µl of GoTaq buffer (5X), 5 µl of $MgCl_2$ (25mM), 1 µl of dNTPs (10mM), 5 µl of both forward and reverse primers (10 pmol μl^{-1}), DNA (10 ng), 0.2 µl of Taq polymerase (GoTaq) and SDW to a total volume of 50 µl. PCR conditions were 95°C for 5 min and 30 cycles of 95 °C for 10 sec, 55°C for 10 sec and 72°C for 30 sec. The Eppendorf 96-well MasterCycler 5331 (Eppendorf, Cambridge, UK) was employed for all HRM PCR reactions.

6.4.5.3. Nested PCR approaches to amplify locus specific fragments for HRM analysis

All sample DNA extracted from TILLing lines to be genotyped (Section 6.4.5.1.) were amplified with BraA.CAX1.a TILLing primers (Section 6.4.6.) using Phusion™ high fidelity DNA polymerase to ensure sequence integrity of amplified sequences (Section 2.7.3.). Reaction volumes comprised of 4 µl of Phusion HF (5X), 0.4 µl of dNTPs (10mM), 1.4 µl of forward and reverse BraA.CAX1.a TILLing primers (10 pmol μl^{-1}), 1 µl DNA (Section 6.4.5.1.), 0.6 µl of DMSO, 0.2 µl of Phusion DNA polymerase and SDW to a total volume of 20 µl. PCR conditions required an initial denaturation at 98°C (30 sec) followed by 30 cycles of denaturation at 98°C (5 – 10 sec), annealing at 58°C (15 sec) and extension at 72°C (30 sec). HRM was performed using components and directions from the Type-it® HRM™ PCR kit (Qiagen, Crawley, UK). A 1 µl aliquot of this reaction was added as a template to the HRM PCR reaction mix comprising 5 µl of 2 x HRM PCR Master Mix (including EvaGreen fluorescent dye), 1.75 µl of forward and reverse HRM primers (10 pmol μl^{-1}) and SDW to 10 µl. In the Eppendorf 96-well MasterCycler 5331, reaction volumes were subjected to 95°C (5 min), followed by 30 cycles of 95°C (10 sec), 55°C annealing (10 sec) and 72°C extension (30 sec).

6.4.5.4. Lightscanner® analysis of HRM products

PCR plates containing HRM amplicons were placed into a Lightscanner® (Idaho Technology, Salt Lake City, USA) and subjected to a 60°C temperature for 5 mins to ensure all samples were at an equal temperature. The melt temperature range was set at 63.3°C – 95.4°C and a ramp setting at 0.1°C with a second hold at each step. Exposure was set at 126 (Auto), background correction was set to exponential, curve shift set to 0.020, standards set to 'Auto Group', and sensitivity at normal +2.8. HRM analysis was then performed on the dissociation of double-stranded DNA (dsDNA) PCR products, which had been saturated with the low PCR-toxic dye, EvaGreen from the initial HRM PCR (Section 6.4.5.3.). As double stranded DNA dissociated in response to heat, the fluorescence of the intercalating EvaGreen dye diminished, thus characterising the PCR products during this transition. The LightScanner® Data Analysis (Version 2.0) detected fluorescence alterations from EvaGreen intercalating dye during these temperature treatments and plotted the data graphically as melt curves for further analysis.

6.4.6. Primers employed

Sequences of primers employed to identify and isolate locus specific sequences for both TILLing and plant genotyping were from 5' to 3':

BrCAX1_Forward	CTAGTTGCCTGCATGGGGATGCAC
BrCAX1_Reverse	GTAGAAATGGCGGGGATCGTGCAG
BrCAX1 TILLing_Forward	AGAGATTTCCTAGCCATGTG
BrCAX1 TILLing_Reverse	CGACCCCTAATTGTTTTATGTG
BrHMA4_Forward	TGATATAGAAATGGCGTTACA
BrHMA4_Reverse	ATGTTGATGTCGACCCTCTTC
BrESB1_Forward	GGACCTACACCACCACCGCCA
BrESB1_Reverse	CTTATCAAATCAATCT
HRM1_Forward	TCCTCGAAGTTGCCTCTGAT
HRM1_Reverse	GCTGCTGACCATTGTTCCTG
HRM2_Forward	CACCTGTGACGCCATTTTAT
HRM2_Reverse	CCCGCCATTTCTACACTGAT
HRM3_Forward	GCAGGAACAATGGTCAGCAG

HRM3_Reverse	CGAGAATGACTTCTTGGAGATT
HRM4_Forward	CGTGCAGGAACAATGGTCAG
HRM4_Reverse	CGAGAATGACTTCTTGGAGATT
HRM5_Forward	AGAAGTCATTCTCGGCACAAA
HRM5_Reverse	CGCTGTTCTGTTTAGTAATGTGTTG
HRM6_Forward	TCCAGAAGGTTCCATACAAAGG
HRM6_Reverse	CGCTGTTCTGTTTAGTAATGTGTTG
HRM7_Forward	GGAAGTTATTGCGGCTTCAG
HRM7_Reverse	CGACCCCTAATTGTTTTATGTG
HRM8_Forward	AGTCATTCTCGGCACAAAGC
HRM8_Reverse	CGACCCCTAATTGTTTTATGTG

6.5. RESULTS

6.5.1. Identifying the heritability of Zinc accumulation in Brassica oleracea based on allelic recombination from two homozygous lines

In order to determine the heritability of accumulated zinc concentrations in shoots ($[Zn]_{shoot}$) in Brassica oleracea and to identify quantitative trait loci (QTL) associated with $[Zn]_{shoot}$, 90 of the most informative doubled-haploid lines from a population of 206 Brassica oleracea lines (Broadley et al., 2008) were selected and referred to as the AGDH population (Sebastian et al., 2000) (Section 6.4.2.). Data was taken from plants grown under both glasshouse conditions (GE2) and field conditions (FE2) at the (University of Warwick HRI, Wellesbourne, UK) (Section 6.4.2.). For all conditions, shoot fresh weights and dry weights of all above ground material were recorded before being processed for mineral extraction to elucidate Zn concentrations (Section 6.4.2.). A significant heritability of 18.5% for $[Zn]_{shoot}$ was observed (Fig. 6.1) however variance components analysis also revealed significant effects of $[P]_{ext}$ (7.2%), and $[P]_{ext}$/accession interaction (2.5%). This indicated that breeding strategies to increase $[Zn]_{shoot}$ in Brassica oleracea could be demanding because of their dependence on environmental conditions.

AGDH accessions	[a]Variance component (%)	Shoot DW (g)			Leaf-Zn			Leaf-P		
	occasion	14.0			2.3			18.1		
	occasion/replicate	4.6			0.4			3.4		
	occasion/replicate/block	7.4			1.1			5.2		
	occasion/replicate/block/plot	0.0			1.6			0.0		
	accession (V_A)	14.2			18.5			7.4		
	$[P]_{ext}$	24.2			7.2			43.1		
	$[P]_{ext}$/accession	1.0			2.5			0.5		
	residual	34.7			66.2			22.4		
	Fixed term	[b]W	[c]d.f.	[d]χ^2	W	d.f.	χ^2	W	d.f.	χ^2
	accession	784	99	<0.001	594	99	<0.001	648	99	<0.001
	$[P]_{ext}$	633	1	<0.001	98	1	<0.001	1712	1	<0.001
	$[P]_{ext}$/accession	123	99	0.049	132	99	0.015	118	99	0.09
	Standard errors of differences of mean									
	Average:	0.09			8.43			0.02		
	Maximum:	0.10			8.94			0.02		
	Minimum:	0.09			8.18			0.02		

Figure 6.1 Variance components analyses of shoot dry weight (DW) (g), shoot Zn concentration (Leaf-Zn), and shoot P concentration (Leaf-P) of the AGDH mapping population (Broadley et al., 2010). [a]Variance components defined in Broadley et al. (2008); [b]Wald test statistic; [c]Degrees of freedom; [d]Chi squared function. Red circle highlights the calculated $[Zn]_{shoot}$ heritability within this population. Data was analysed using REML procedures in GenStat (Release 9.1.0.147; VSN International, Oxford, UK) (Section 6.4.2.).

6.5.2. Quantitative trait loci (QTL) associated with $[Zn]_{shoot}$ in Brassica oleracea and B. rapa

Under glasshouse conditions three quantitative trait loci (QTL) that associated with $[Zn]_{shoot}$ (Zn-QTL) were detected in the AGDH population (Fig. 6.2). For both high and low $[P]_{ext}$ combined, significant Zn-QTL were detected on C2 (82.2 cM, LOD 3.2) and C9 (69.2 cM, LOD 2.5) and for high $[P]_{ext}$ a unique QTL on C5 (54.7 cM, LOD 2.6) was detected (Fig. 6.2). Under field conditions Zn-QTL were detected in C1, C3 and C7. Several loci were identified in C3 including its upper region (21.1 cM, LOD 4.1). Although there were no significantly consistent QTL between glasshouse and field environments, C3 did also demonstrate Zn-QTL in its upper arm (18.3 cM, LOD 2.3) under glasshouse conditions (Broadley et al., 2010). In B. rapa, two QTL affecting shoot dry biomass (SDB) under increasing Zn nutrition conditions ($[Zn]_{ext}$) were detected on chromosome 3 (R3) (8 cM, LOD 3.9) and (11 cM, LOD 3) (Wu et al., 2008) (Section 6.4.2.). B. rapa chromosome 3 (R3) was syntenic with B. oleracea chromosome 3 (C3) (Rana et al., 2004). One QTL affecting SDB under high $[Zn]_{ext}$ was also detected on R6 (59 cM, LOD 3.1). On the same chromosome a QTL affecting $[Zn]_{shoot}$ was detected (79 cM, LOD 3.9). Finally

one QTL affecting $[Zn]_{shoot}$ was detected on R4 (28 cM, LOD 5.4) (Wu et al., 2008). For all B. oleracea QTL-fitting, a non-conservative likelihood ratio significance threshold of 11.5 (LOD 2.5) was used; the actual likelihood ratio for significant QTL associated with $[Zn]_{shoot}$ (Zn-QTL) was >12.9, based on 1000 random permutations across all Zn-traits (Broadley et al., 2010) (Section 6.4.2.). For B. rapa, significance thresholds accepting the presence of QTLs were determined by 1000 permutation tests for each trait and LOD score ≥ 3 were considered to be significant (Wu et al., 2008) (Section 6.4.2.).

		Chromosome	cM	LOD	Additive effect (A12DHd allele)	1-LOD support (left boundary)	1-LOD support (right boundary)	2-LOD support (left boundary)	2-LOD support (right boundary)
GE2	**Zn (both P-levels)**	2	82.2	3.2	-5.9186	70.2	89.8	70.1	92.6
		9	69.2	2.5	-7.1554	60.8	70	55	73.6
	Zn (high $[P]_{ext}$)	5	54.7	2.6	-4.1992	40.4	61.3	33.7	75.2
FE2	**Zn (all P-levels)**	1	91.2	3.3	-1.3963	74.7	97.2	63.9	97.2
		3	21.1	4.1	-1.4761	16.8	37	8.4	43.7
		7	38.5	3.3	1.636	33.8	50.5	30.5	51.4
	Zn (298 kg TSP ha^{-1})	1	89.2	2.6	-1.7592	69	97.2	60.6	97.2
		3	21.1	3.9	-2.4128	17.5	28.1	12.2	29.4
		3	43.4	4.2	-3.6428	35	43.7	29.4	44
		7	52.4	3	1.7883	39.8	62	38.3	68
	Zn (1125 kg TSP ha^{-1})	7	58.4	2.8	1.6836	41.8	61.5	38.3	64.9

Figure 6.2 Quantitative trait loci QTL (LOD > 2.5) associated with shoot Zn concentration in Brassica oleracea from 90 doubled-haploid (DH) accessions of the AGDH mapping population (Broadley et al., 2008), grown at low $[P]_{ext}$ and high $[P]_{ext}$ in glasshouse conditions (GE2) and for 70 DH accessions grown at four rates of P-fertiliser addition under field conditions (FE2) (Section 6.4.2.)

6.5.3. Identifying candidate genes involved in Zinc accumulation in Brassica spp.

Quantitative trait loci (QTL) associating with altered $[Zn]_{shoot}$ in Brassica (Zn-QTL) represented several thousand genes and so these phenotypes could not be efficiently resolved to specific genes. Comparative mapping analysis was therefore performed on two mapping experiments of B. rapa (Wu et al., 2008) and B. oleracea (Broadley et al., 2010) and a published B. napus map containing over 1000 genetically linked RFLP loci mapped to homologous positions in A. thaliana (Parkin et al., 2005) (Section 6.4.2.). A. thaliana genes with known involvement in Zn accumulation and which mapped to Brassica loci within Zn-QTL from these two mapping experiments were selected as candidate genes. One QTL affecting $[Zn]_{shoot}$ in B. rapa was located on chromosome 4 (R4) (28 cM, LOD 5.4) (Wu et al., 2008) and localised with an RFLP probe pN151 (Parkin et al., 2005) which mapped to a region in A. thaliana containing Enhanced Suberin 1 (ESB1) gene (At2g28670), which is involved in suberin deposition in the root endodermis and altered $[Zn]_{shoot}$ in A. thaliana (Baxter et al., 2009). This probe also hybridised to a QTL on R3 affecting shoot dry biomass in response to external Zn supply (8 – 11 cM, lowest LOD 3) (Wu et al., 2008). QTL affecting $[Zn]_{shoot}$ in B. oleracea grown under glasshouse conditions (GE2) (18.3 cM, LOD 2.3) and under field conditions (FE2) (21.1 cM, LOD 4.1) were located on chromosome 3 (C3) and localised with an RFLP probe pO98 (Parkin et al., 2005) which mapped to a region in A. thaliana containing the vacuolar localized high affinity calcium:hydrogen antiporter CAtion/H^+ eXchanger 1 (CAX1) (At2g38170) (Pittman et al., 2005) which has been associated with altered $[Zn]_{shoot}$ in A. thaliana (Cheng et al., 2005) and tomato (Park et al., 2005). On the same chromosome (C3) another QTL affecting $[Zn]_{shoot}$ in B. oleracea (43.4 cM, LOD 4.2) localised to a region flanked by two RFLP probes, pO79 and F20108 (Parkin et al., 2005), which mapped to a region in A. thaliana containing HMA4 (At2g19110).

6.5.3.1. In silico data mining

Basic Local Alignment Search Tool (BLAST) was initially employed to obtain loci specific Brassica gene orthologues of AtHMA4, AtCAX1 and AtESB1. GenBank/NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and the BrGSP BLAST Server (http://brassica.bbsrc.ac.uk/BrassicaDB/BrGSP_blast.html) were the two

primary databases utilised to retrieve Brassica sequence data. Briefly, 2Mb fragments of A. thaliana sequence obtained from GenBank/NCBI containing the gene of interested were entered into the BLAST search programme for both GenBank/NCBI, (using the algorithms BLASTn, megablast and discontiguous megablast at default settings (http://www.ncbi.nlm.nih.gov/blast/product_table.shtml#tab31)), and the BrGSP BLAST Server (operating a WuBLAST 2.0 algorithm at default settings (http://blast.wustl.edu/)) (Section 6.4.3.1.). Results from both databases identified no full length Brassica orthologous sequences for HMA4 genes. For AtCAX1 one full length sequenced BAC (KBrH004D08, GenBank: AC189543.2) from the closely related B. rapa Chiifu genome contained one CAX1 homologous sequence. Similarly for AtESB1 one full length sequenced BAC (KBrH104J01, GenBank: FP340660.1) from the B. rapa Chiifu genome showed >85% genomic sequence identity to AtESB1 and contained flanking regions demonstrating relative synteny with the Arabidopsis genomic region. Genes EST, GSS and Bacterial Artificial Chromosomes (BAC)-end sequences which demonstrated greater than 80% sequence identity to the A. thaliana region were selected as putative Brassica orthologous sequences and compiled for further refinement.

6.5.3.2. Elucidating Brassica spp. orthologous sequences

An average of 300 clones, displaying an initial 80% identity to each of the A. thaliana fragments, were further analysed to remove false positives which would provide misleading consensus sequences from subsequent sequence alignments. In respect of AtHMA4, false positives were identified as sequences which showed greater homology to closely related members of the HMA gene family such as AtHMA2 (At4G30110) and AtHMA3 (At4G30120). For CAX1, similar false positives were identified among sequences showing greater homology to AtCAX2 (At3G13320) and AtCAX3 (At3G51860). All Brassica spp. sequences homologous for each gene were then assembled into contiguous sequences (contigs) using assembling software ContigExpress, VectorNTI 11 (Invitrogen, Paisley, UK) at default settings; overlap percentage cut off; 70% and gap extension penalty set to 20 (Section 6.4.3.1) (Fig. 6.3). Contig sequences were aligned to their corresponding 2Mb AtHMA region using the Pairwise Alignment tool using AlignX software, (VectorNTI 11), at default settings; gap opening penalty, 15 and gap extension

penalty 6.66. Together these aligned sequences provided fragmented preliminary views of Brassica orthologous sequences (Fig. 6.3). Consensus sequences obtained from these alignments were employed as a template to design primers to amplify locus specific regions of these genes for downstream molecular analyses

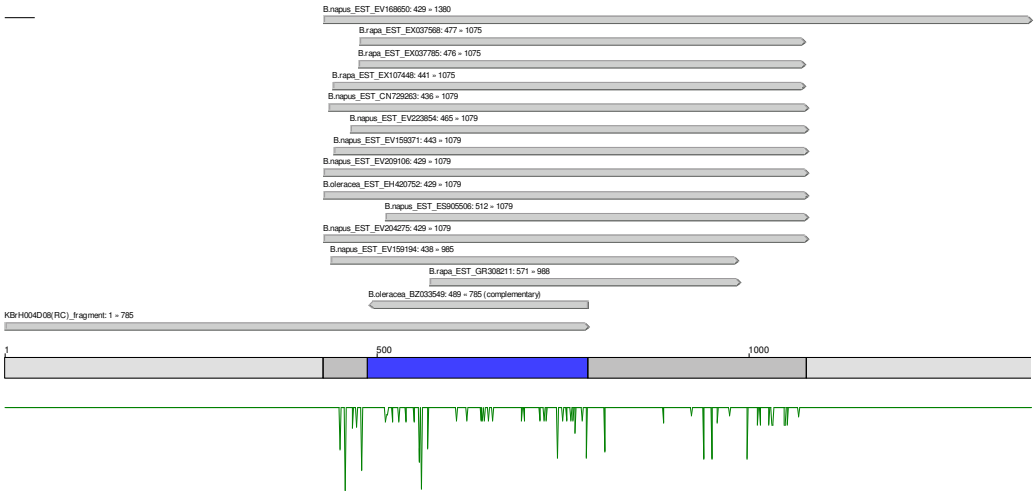


Figure 6.3 Contig assembly of publicly available EST, GSS, and BAC sequences from Brassica spp. showing high (>80%) identity to Arabidopsis thaliana CAX1 (AT2G38170) identified from BLAST (using the algorithms BLASTn, megablast, discontiguous megablast and WuBLAST 2.0 algorithms at default settings). Gray arrows represent directionally assembled Brassica sequences. Bar represents contig coverage with multiple sequences in both directions (blue fill), in one direction (purple hatch), and single sequence coverage (green slant). Green scale represents an overview of assembly quality using default settings of the software (ContigExpress, Vector NTI 11, Invitrogen).

6.5.3.3. Isolating locus specific sequences from Brassica rapa R-O-18

Primers were designed to amplify approximately 1.2 Kb of sequence for each Brassica orthologous fragment generated through in silico data mining and sequence assemblies (Sections 6.4.3.1.). Since it was desirable to obtain truncated mutations in the 5' region of these genes, primers were designed to amplify the first exon of HMA4 and CAX1 sequences. For CAX1 this involved designing primers 600 bp upstream from the transcriptional start site and 600 bp downstream in the first intron (Section 6.4.6.). For HMA4, primers were designed to include the transcriptional start site and 1641bp downstream in the second exon (Section 6.4.6.). For ESB1 however, due to the low sequence identity between promoter regions from B. rapa Chiifu and R-O-18, primers were designed 571 bp downstream from the transcriptional start site in the first exon and 259 bp downstream in the 3' UTR (Section 6.4.6.). Sequences were amplified from genomic DNA extracted from B.

rapa R-O-18 (donated by G. King, Rothamsted Research) (Section 6.4.1.) by PCR using proofreading polymerase (Phusion®, New England Biolabs, Hitchin, UK). To test for locus specificity of amplified sequences, a diagnostic restriction endonuclease digest was performed. Here PCR amplicons were purified by gel electrophoresis (Section 2.6.3.) and 6 µl aliquots were digested individually using the restriction endonucleases EcoRI, HindIII, NotI and XhoI as described in Section 2.6.4. As a control, original amplicons were run alongside digested samples in an electrophoresis gel to compare product band sizes (Fig. 6.4). Only HMA4 displayed multiple banding when digested with NotI. The summation of these band sizes was approximately 1641 bp indicative of a locus specific sequence (Fig. 6.4).

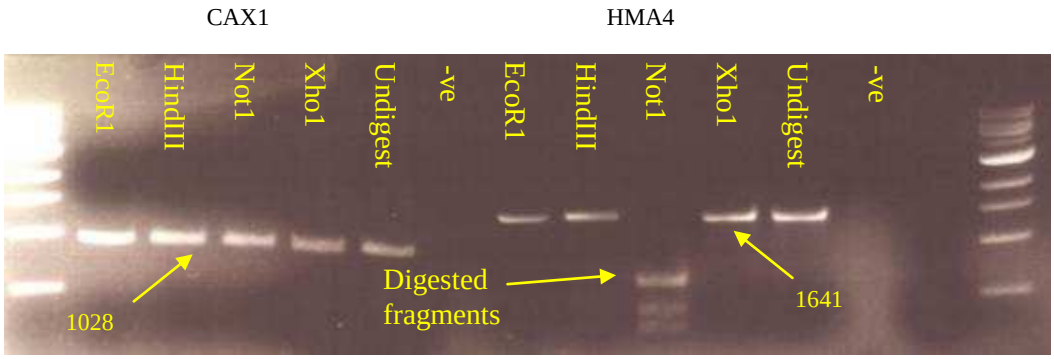


Figure 6.4 Restriction digest of PCR amplified TILLing sequences of B. rapa R-O-18 CAX1 and HMA4 regions using EcoRI, HindIII, NotI and XhoI incubated for 4 hours. Lanes labelled 'Undigested' contain undigested amplified DNA. Arrows represent bp according to the 1 Kb DNA ladder (New England Biolabs). Gel contained 1% (w/v) agarose.

Both the 1028 bp B. rapa R-O-18 CAX1 (BraA.CAX1.a) and 1641 bp BraA.HMA4.a sequences were purified, isolated by gel electrophoresis (Section 2.6.3.) and A-tail adapted for downstream cloning into the pCR8®/GW/TOPO® TA Cloning® system (Invitrogen, Paisley, UK) (Section 2.5.6.). Aliquots of 2 µl of the pCR8® compatible BraA.CAX1.a and BraA.HMA4.a fragments were individually ligated into a pCR8® entry vector as described in Section 2.5.6.2. Ligated fragments were transformed into competent E. coli DH5α cells (50 µl) via chemical and heat shock treatments (Section 2.5.6.3.). Transformed E. coli colonies were selected on solid LB media (Section 2.5.6.4.) and subjected to colony PCR with either BraA.CAX1.a or BraA.HMA4.a specific primers to test for their presence in the bacterial plasmid (Section 2.5.6.4.) (Fig. 6.5). The positive "Genomic" control contained B. rapa R-O-

18 genomic DNA as a template (Fig. 6.5). Three colonies containing BraA.HMA4.a and six containing BraA.CAX1.a were grown overnight in 50 ml of LB, supplemented with 50 μg ml^{-1} spectinomycin (Section 2.5.6.4.). Following plasmid isolation (Section 2.5.7.), inserts were sequenced using the Value Read, Single Read Service by Sanger sequencing (Eurofins MWG GmbH, Ebersberg, Germany) (Section 2.8). All clones returned single locus specific sequences for both BraA.HMA4.a and BraA.CAX1.a. Glycerol stocks were made of all confirmed E. coli DH5α cells for both sequences (Section 2.5.5.).

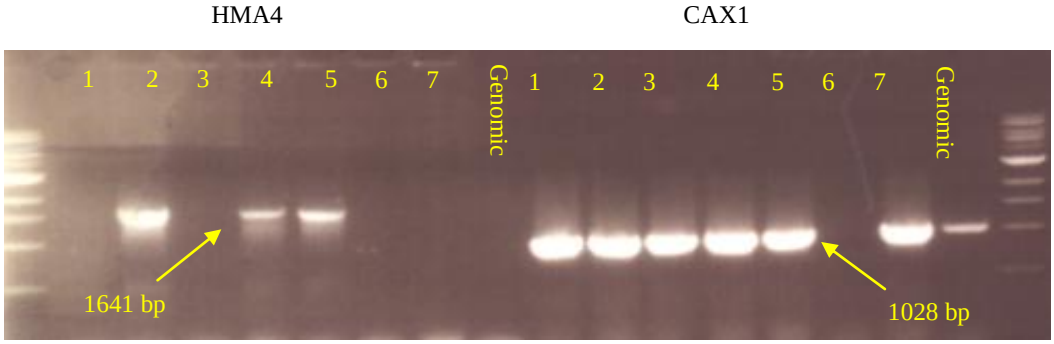


Figure 6.5 Single colony PCR on Escherichia coli DH5α cells transformed with PCR8® entry vector containing either B. rapa HMA4 or CAX1 TILLing sequences. Lanes labelled 'Genomic' contained PCR amplified B. rapa genomic DNA. Arrows represented bp according to the 1 Kb DNA ladder (New England Biolabs). Gel contained 1% (w/v) agarose.

The 1178 bp BraA.ESB1 PCR amplicon was purified, isolated by gel electrophoresis (Section 2.6.3.) and A-tail adapted for downstream cloning into the pCR8®/GW/TOPO® TA Cloning® system (Section 2.5.6.1.). Competent E. coli DH5α cells transformed with these products were selected on solid LB media (Section 2.5.6.4.) and subjected to colony PCR with BraA.ESB1.a specific primers to test for their presence in the bacterial plasmid (Fig. 6.6). The positive "Genomic" control contained B. rapa R-O-18 genomic DNA as a template (Fig. 6.6). Plasmids extracted from 14 colonies containing BraA.ESB1 (Section 2.5.7.), were sequenced through the Value Read, Single Read Service by Sanger (Section 2.8.). which returned three paralogous sequences branded BraA.ESB1.a, BraA.ESB1.b and BraA.ESB1.c. Glycerol stocks were made for all three confirmed paralogues in E. coli DH5α cells (Section 2.5.5.). BraA.ESB1.a and BraA.ESB1.c contained two independent single nucleotide polymorphisms (SNP) from the homologous sequence published in B. rapa Chiffu. BraA.ESB1.b demonstrated 100% homology with the published B. rapa Chiffu sequence. Phylogram models of unrooted DNA parsimony (Section 2.8.2.2.) demonstrated that for both ESB1 and CAX1 orthologues, sequences

from B. rapa Chiifu and B. rapa R-O-18 maintained closer sequence identity than with their homologous in the closely related A. thaliana (Fig. 6.7).

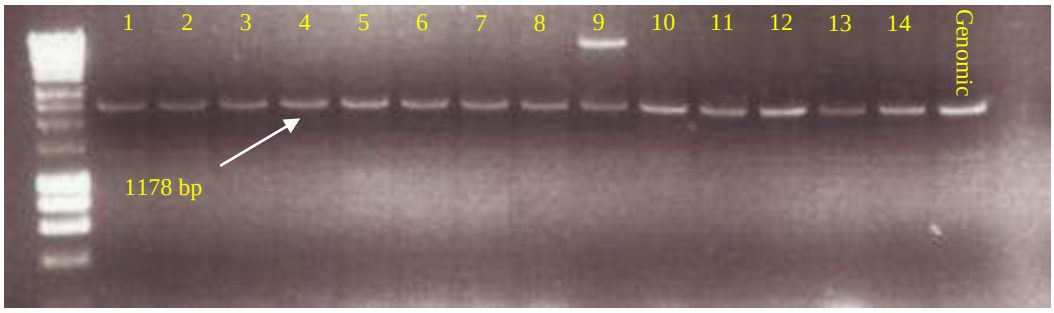


Figure 6.6 Single colony PCR on Escherichia coli DH5α cells transformed with PCR8® entry vector containing B. rapa ESB1 TILLing sequences. Lane labelled 'Genomic' contained PCR amplified B. rapa genomic DNA. Arrows represented bp according to the 1 Kb DNA ladder (New England Biolabs). Gel contained 1% (w/v) agarose.

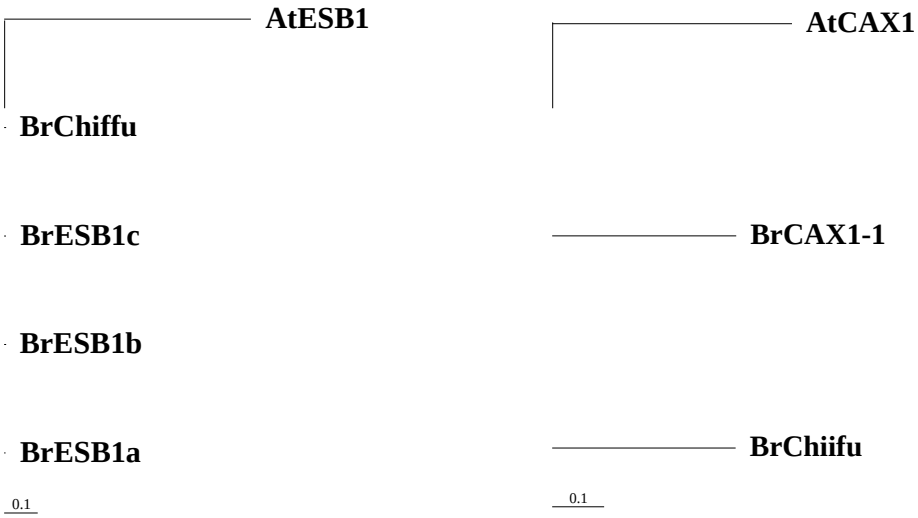


Figure 6.7 Phylogram of unrooted DNA parsimony of PCR amplified sequences from A. thaliana (At), B. rapa Chiifu (BrChiifu) and B. rapa R-O-18 (Br) of CAX1 and ESB1 orthologs. Diagrams were produced from DNAPars software at default settings, Phylip version 3.68.

6.5.3.4. Identifying CAX1 allelic variants within the Brassica rapa R-O-18 genome through Targeting Induced Local Lesions IN Genomes (TILLING)

The cloned locus specific B. rapa R-O-18 BraA.CAX1 sequence was sent to RevGen UK (http://revgenuk.jic.ac.uk/) to be employed as a template to screen and select for R-O-18 plants containing mutations within this sequence. B. rapa R-O-18 seeds were previously subjected to ethyl methanesulfonate (EMS) based mutagenesis and resulting M_1 plants were selfed to produce 9216 M_2 plants whose DNA was

extracted and stored in a library of 384-well plates (Stephenson et al., 2010). EMS based mutagenesis induced base pair transitions i.e. from AT – GC which resulted in intronic splice variants and introduced nonsense, missense and silent amino acid changes in the coding sequences (Colbert et al., 2001). Prior to screening, RevGen processed the BraA.CAX1 sequence through a CODDLE program (Codons Optimised to Discover Deleterious Lesions) (http://www.proweb.org/coddle/) at default settings (Section 6.4.4.1.) to identify regions where anticipated point mutations would result in deleterious effects on gene function. Several primer pairs were produced by the program to amplify the BraA.CAX1 locus specific region, and were functionally examined, using wild type (WT) B. rapa R-O-18 DNA, to test for BraA.CAX1 locus specificity. Two locus specific BraA.CAX1 primers (CAX1 Tilling Primer) were selected to amplify a 760 bp fragment, with the forward primer located 347 bp upstream from the transcriptional start site and the reverse located in the first intron, 141 bp downstream from the first exon (Section 6.4.6.) (Fig. 6.10). Primers were then fluorescently labelled with FAM (Fluorescein amidite, a blue labeling fluorescent tag fused to the forward primer) and HEX (Hexachlorofluorescein phosphoramidite, a green labeling fluorescent tag fused to the reverse primer) and employed to amplify pools of DNA (8 individual samples) before inducing heteroduplexes between WT and mutant DNA (Section 6.4.4.2.). Cleavage of these heteroduplexes, induced by incubation with the endonuclease Cel1 revealed fluorescently labelled traces of mutated genes through ABI3730 sequencer analyses (Section 6.4.4.2.). A total of 26 M_2 lines were identified with independent SNP mutations in BraA.CAX1. Of these mutations, 13 lines were missense, one was nonsense, five were silent and seven were located in non coding regions. Three missense lines and one silent line were homozygous for the mutation. Mutations in the BraA.CAX1 region were asymmetrically distributed along the 760 bp amplicon (Fig. 6.8). This may have been a result of the position of the 5' primer in an AT-rich non-coding region, which was necessary to maintain locus specificity. Conversely the 3' primer was located close to the first exon in an area more densely populated with mutable GC sites and so contained more detected point mutations.

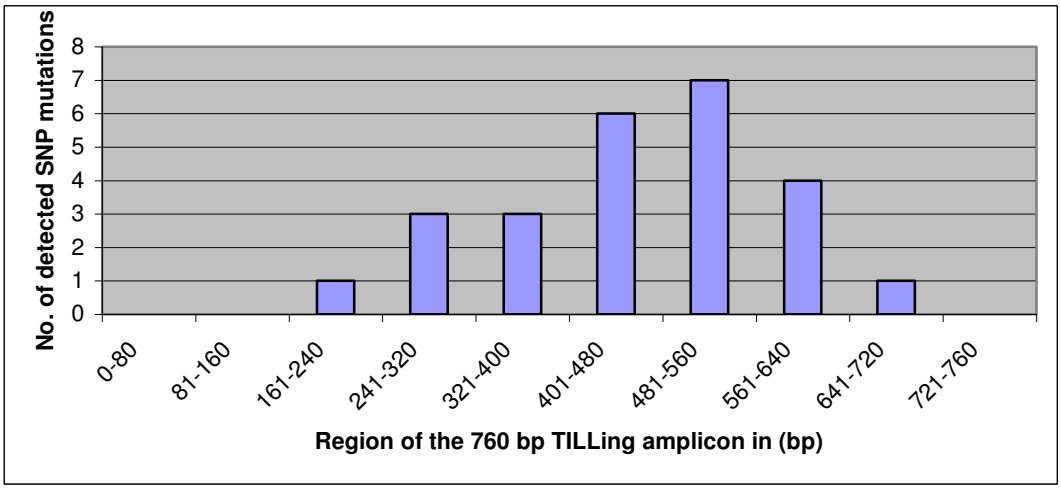


Figure 6.8 Distribution of SNP mutations in the BraA.CAX1 760 bp TILLing amplicon for all 26 lines. The detected mutations were unequally distributed with fewer located near to the flanking primers, and most appeared between 401 – 560 bp in exon 1 of the CAX1 ortholog.

6.5.3.5. Efficient sampling of M_2 and M_3 Brassica rapa R-O-18 mutant DNA

M_3 mutant seeds derived from the 26 M_2 lines identified by TILLing in the BraA.CAX1 region were sourced from RevGen UK. Due to poor fecundity, 7 M_2 lines did not produce seeds, therefore seeds for a total of 19 M_3 lines were received. In order to recover the specific mutations lost in the 7 sterile lines, M_2 sibling lines, which were not screened by TILLing, were adopted for all including these 7 lines. This resulted in a population of seeds for 19 M_3 and 26 M_2 lines. Depending on seed availability a maximum of 5 replicates for M_3 and 2 replicates for M_2 lines (n=135) were sown in 1 l pots containing Levington M3 compost, and grown on benches under glasshouse conditions (22.3°C and 13.3°C mean day and night temperatures respectively) at a 16 hr photoperiod, with overhead irrigation. An efficient and inexpensive DNA extraction protocol was thus adopted to initiate the genotyping of all 135 samples "Speed DNA extraction protocol" (Mervin Poole, pers. comm.) (Section 6.4.5.1.). Once the first true leaves had begun to expand, 0.25 cm^2 leaf discs were taken from each plant and placed in individual wells of a 96 deep-well plate (Qiagen Collection Microtubes). When all samples were taken, a 3 mm stainless steel ball was added to each well, and the material dissolved in 600 µl of DNA extraction buffer solution (Section 6.4.5.1.). Whole plates were vibrated vigorously in a ball mill for 3 minutes before being centrifuged in a plate centrifuge at 5600 g

for 20 min. DNA contained in a 300 µl aliquot of the supernatant was then purified and resuspended as described in Section 6.4.5.1.

6.5.3.6. Efficient screening for BraA.CAX1 mutations in M_2 and M_3 Brassica rapa R-O-18 TILLing lines

To efficiently identify mutated BraA.CAX1 sequences in M_2 and M_3 lines growing under glasshouse conditions, a High Resolution Melt (HRM) analysis was adopted (Section 6.4.5.1.). This involved designing primers to amplify between 100 and 250 bp fragments within the 760 bp Tilling region, since larger fragments reduced HRM sensitivity due to increased multiple melt domains and complicated melt curves (Section 6.4.5.1.) (Fig. 6.9). Prior to selection, primer pairs were assayed for secondary structures including primer dimer duplexes, hairpin loops, palindromes and repeats of greater than 4 nucleotides using the secondary structure profiling software Oligo Analysis at default settings (Vector NTI 11, Invitrogen, Paisley, UK) (Section 2.8.2.3). Primers were then PCR tested on WT B. rapa R-O-18 genomic DNA to affirm efficiency, functionality and single band amplification with no resulting primer dimer formations (Fig. 6.9). Once assayed, four optimal primer pairs were selected, whose amplified fragments together represented the entire region containing SNPs (Fig. 6.9). A genomic illustration of this process was represented for primer set six, which amplified a region containing six independent SNPs, representing TILLing lines 1, 2, 4, 10, 11 and 25 (Fig. 6.10).

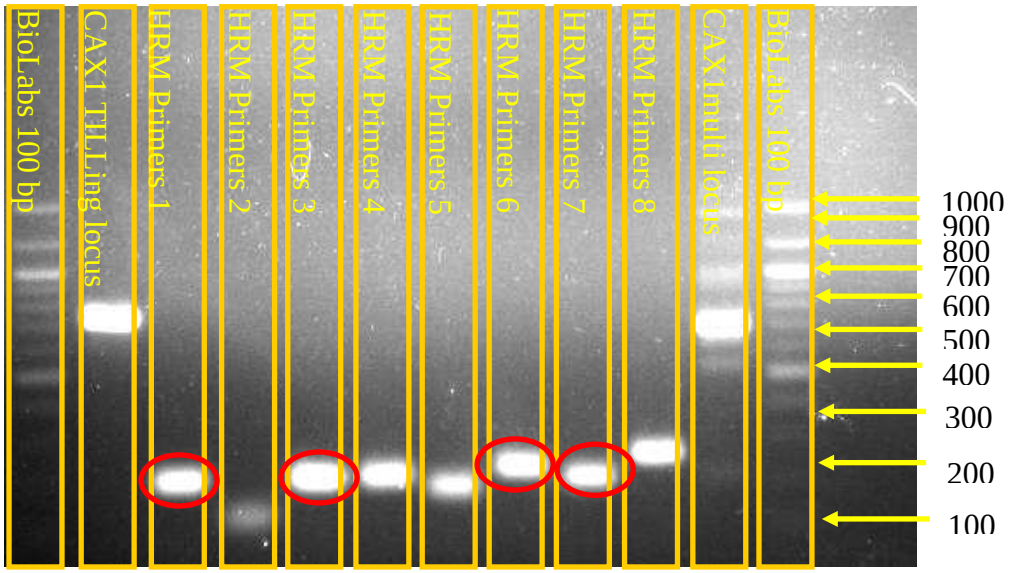


Figure 6.9 PCR amplicons from Brassica rapa R-O-18 DNA using 8 HRM primers sets (Annealing temp = 58 °C). The locus specific "CAX1 TILLing locus" and non-specific "CAX1 multi locus" primers were used as positive controls. Red circles indicate HRM primer pairs continued for mutant TILLing line genotyping. Arrows represent bp according to a 100 bp DNA ladder (New England Biolabs). Gel contained 1% (w/v) agarose.

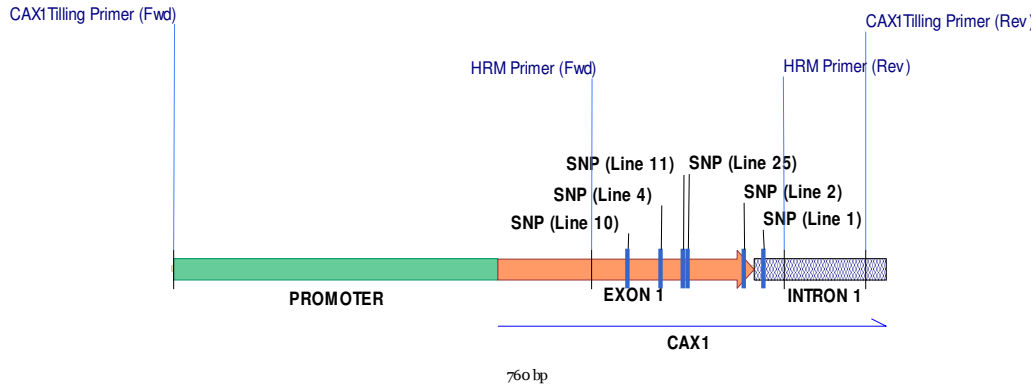


Figure 6.10 Representation of the locus specific 760 bp B. rapa CAX1 fragment used for TILLing. Primers used for TILLing (CAX1TILLing Primer) and for HRM (HRM Primer) are highlighted in blue. Six EMS induced base pair translocation locations are annotated as SNP lines. Blue arrow represents translational direction of CAX1 from its translational start site. The bar represents the CAX1 promoter (green), CAX1 exon (orange) and intron (hatched) regions. Image generated using VectorNTI 11 software (Invitrogen, Paisley, UK).

In order to ensure locus specificity for HRM primers a nested PCR approach was employed. This involved amplifying the 760 bp TILLing fragment with locus specific BraA.CAX1 TILLing primers for all M_2 and M_3 mutant lines using proofreading polymerase (Section 6.4.5.3.). A 1 µl aliquot from each reaction was then applied to each corresponding HRM PCR mix and amplified for 30 cycles to maintain optimum amplification with minimum artifacts (Section 6.4.5.3). Following

HRM PCR, samples were viewed on a gel to verify that all had amplified to a plateau phase (Fig. 6.11). At this phase all reactions had amplified to relatively similar quantities despite no prior quantification of starting DNA material, as recommended by Type-it®-HRM PCR Kit (Qiagen, Crawley, UK). All concentrations of DNA were estimated to be equal and therefore the T_m of samples amplified by each primer set were similar.

For HRM analysis, lines which were amplified by individual HRM primer sets were grouped on 96-well PCR plates. All samples were triplicated on each plate, with three independent triplicated water (negative samples) and WT B. rapa R-O-18 (positive samples). Plates were placed into a Lightscanner® (Idaho Technology, Salt Lake City, USA) and subjected to a 60°C temperature for 5 mins to ensure all samples were at an equal temperature. The melt temperature range was set at 63.3°C – 95.4°C and a ramp setting at 0.1°C with a second hold at each step. Exposure was set at 126 (Auto), background correction was set to exponential, curve shift set to 0.020, standards set to 'Auto Group', and sensitivity at normal +2.8 (Section 6.4.5.4.).

Figure 6.11 PCR products from HRM primer set 6 amplifications from 13 Brassica rapa R-O-18 M_3 TILLing samples (representing lines 1, 2, 4, 10, 11 and 25) which were identified as SNP mutants using HRM analysis. Two wild type B. rapa R-O-18 genomic samples were amplified as positive controls. No template was used as a negative control (-ve). Arrows represented bp according to a 100 bp DNA ladder (New England Biolabs). Gel contained 1% (w/v) agarose.

6.5.3.7. Genotyping Brassica rapa R-O-18 TILLing lines for mutations in BraA.CAX1

From a total of 26 genotyped lines, 19 M_3 and 26 M_2 (n=135), for CAX1 SNP mutations, 59 individual plants (> 37%) were identified, representing 24 independent CAX1 lines. Of the seven M_2 lines employed as sibling replacements for sterile M_3 lines, five exhibited CAX1 SNPs mutations. A subset of 6 lines which were positively genotyped for CAX1 SNPs by HRM and two wild type plant controls (n=15) were sequenced by Sanger (Section 2.8.1.). All 13 mutant plants were shown to contain a CAX1 SNP mutation and of these, four were homozygous (exemplified by samples 1C and 1E, Figs. 6.12 & 6.13) and nine were heterozygous (exemplified by samples 11A and 11C, Figs. 6.14 & 6.15) for the mutation. Within this pool of samples, HRM was capable of detecting all samples containing CAX1 SNPs and returned no false positives.

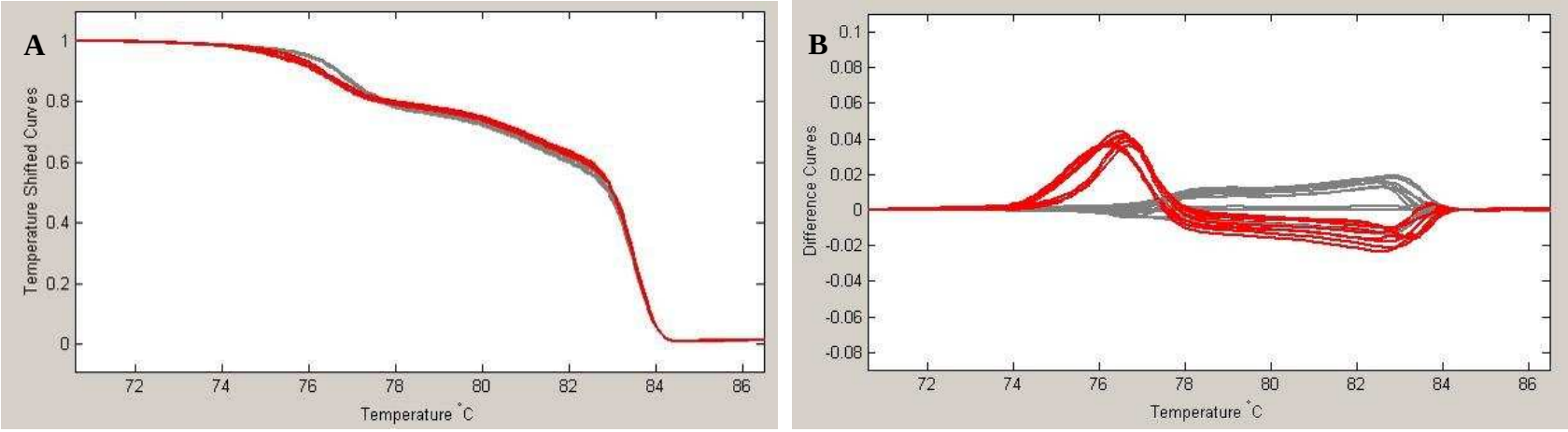


Figure 6.12 LightScanner® Data Analysis (Version 2.0) for 260 bp sequences from Brassica rapa R-O-18 wild type and M_3 TILLing mutants 01C and 01E. **A:** Normalised melt curves for EvaGreen dye (Qiagen, Crawley, UK) fluorescent signals from DNA strand dissociation of triplicated technical replicates of B. rapa wild type (grey lines) and M_3 TILLing mutants, 01C and 01E (red lines). All genotypes were auto-called by the software packages and were not altered by the operator. **B:** Difference plot of genotypes' fluorescence normalised to wild type samples (grey lines). Run Parameters: Melt Temperature Range; 63.3 – 95.5 °C, Exposure; 126 (Auto). Scanning Analysis Parameters: Temperature Range; lower 70.6 – 72.6 °C, upper; 85.5 – 86.5 °C, Background Correction; exponential, Curve Shift; 0.020, Standards; Auto Group, Sensitivity; normal +2.8.

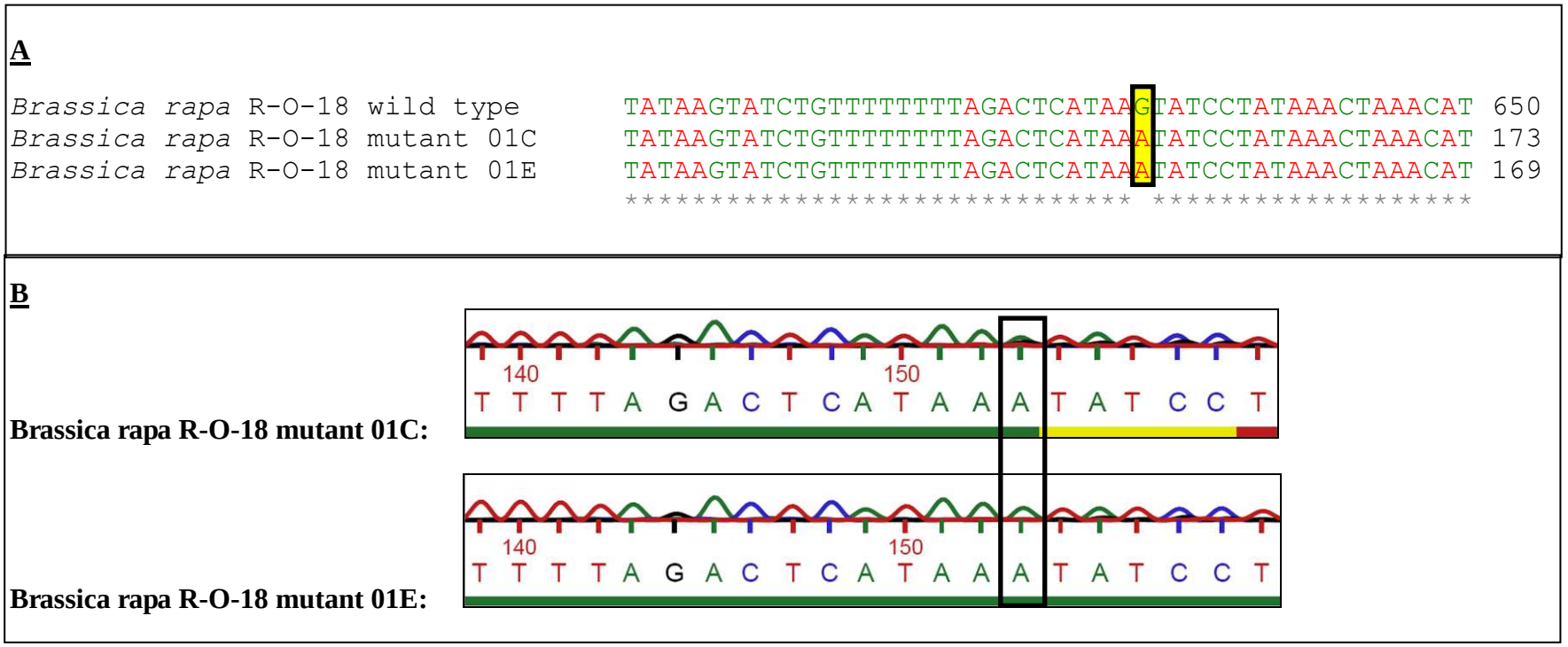


Figure 6.13 **A:** Sequencing text data for Brassica rapa R-O-18 wild type and M_3 TILLing mutants, 01C and 01E. Rectangular box indicates the site of EMS induced transitional mutation from G to A.

B: Sequencing chromatograms of Brassica rapa R-O-18 mutants 01C and 01E. Rectangular box indicates homozygosity for the site of EMS induced transitional mutation. Average spacing: 83.0. Line under base letters represents the quality of each base call: Red line, 10 – 19; Yellow line, 20 – 29; Green line, ≥ 30.

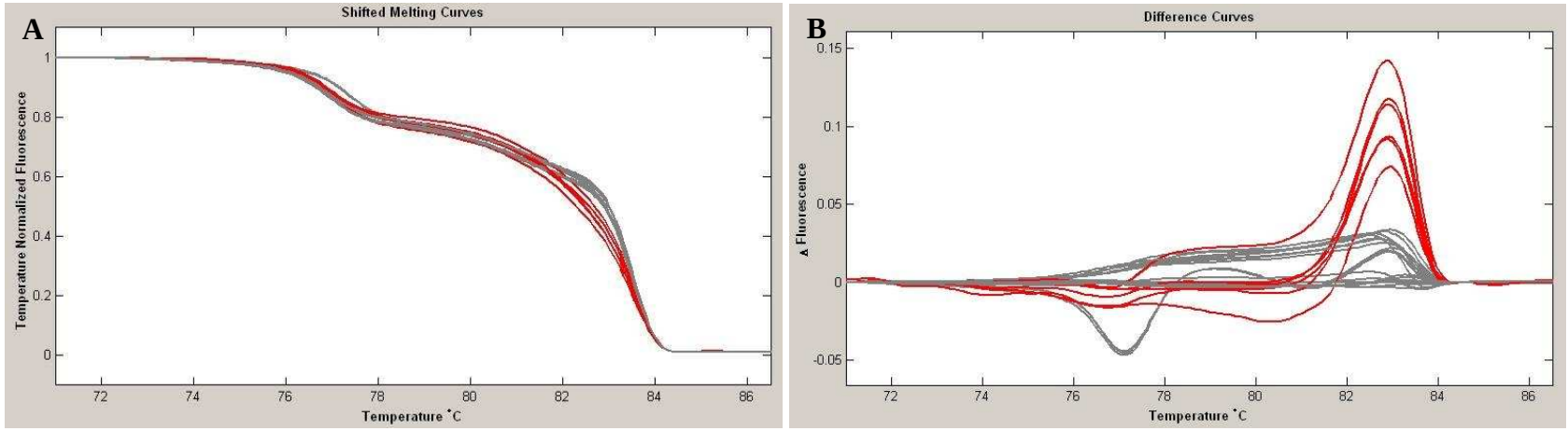


Figure 6.14 LightScanner® Data Analysis (Version 2.0) for 260 bp sequences from Brassica rapa R-O-18 wild type and M_3 TILLing mutants 11C and 11A. **A:** Normalised melt curves for EvaGreen dye (Qiagen, Crawley, UK) fluorescent signals from DNA strand dissociation of triplicated technical replicates of B. rapa wild type (grey lines) and M_3 TILLing mutants, 11C and 11A (red lines). All genotypes were auto-called by the software packages and were not altered by the operator. **B:** Difference plot of genotypes' fluorescence normalised to wild type samples (grey lines). Run Parameters: Melt Temperature Range; 63.3 – 95.5 °C, Exposure; 126 (Auto). Scanning Analysis Parameters: Temperature Range; lower 70.6 – 72.6 °C, upper; 85.5 – 86.5 °C, Background Correction; exponential, Curve Shift; 0.020, Standards; Auto Group, Sensitivity; normal +2.8.

A

```
Brassica rapa R-O-18 wild type      TCATTCTCGGCACAAAGCTAGCCATTCTTTTTCCGGCCATACCTGCCGCC 550
Brassica rapa R-O-18 mutant 11C     TCATTCTCGGCACAAAGCTAGCCATTCTTTTTCCGGCCATACCTACCGCC 82
Brassica rapa R-O-18 mutant 11A     TCATTCTCGGCACAAAGCTAGCCATTCTTTTTCCGGCCATACCTACCGCC 73
                                    ******************************************* *****
```

B

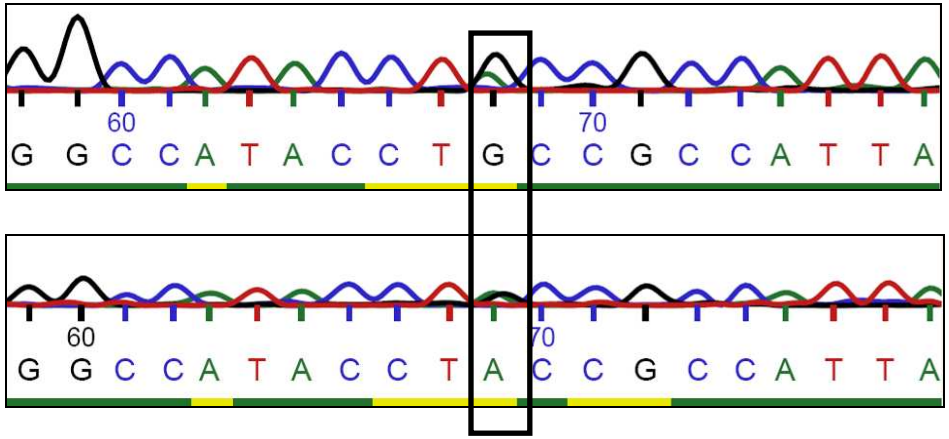


Figure 6.15 **A:** Sequencing text data for Brassica rapa R-O-18 wild type and M_3 TILLing mutants, 11A and 11C. Rectangular box indicates site of EMS induced transitional mutation from G to A.
B: Sequencing chromatograms of Brassica rapa R-O-18 mutants 11A and 11C. Rectangular box indicates heterozygosity for the site of EMS induced transitional mutation. Average spacing: 83.0. Line under base letters represents the quality of each base call: Red line, 10 – 19; Yellow line, 20 – 29; Green line, ≥ 30.

6.6. DISCUSSION

This chapter described the efficient identification and isolation of locus specific candidate Zn transporters for roles in Zn accumulation in the commercially significant crop Brassica rapa L. Until recently, classical approaches to identify genes of interest chiefly involved lengthy fine mapping of quantitative trait loci (QTL) and were often hindered by the presence of gene paralogues (Bruce et al., 2009). However, by exploiting the colinearity between the closely related Brassica spp. and fully sequenced A. thaliana (Parkin et al., 2005), comparative genomics should permit earlier allele identification. These approaches were employed to select three orthologues associated with Zn transport throughout the Brassicaceae which mapped to regions within quantitative trait loci (QTL) involved in Zn dependent traits in B. rapa (Wu et al., 2008) and B. oleracea (Broadley et al., 2010). These included HMA4, whose selection was supported by novel quadruplication and altered cis regulation observed in Zn hyperaccumulators N. caerulescens (Chapters 3 and 4) and A. halleri (Hanikenne et al., 2008) as well as functional and expression evidence throughout the Brassicaceae (reviewed by Broadley et al., 2007; Verbruggen et al., 2008; Kramer et al., 2010). Functional evidence similarly existed for CAX1 (Shigaki et al. 2003, 2005; Cheng et al., 2005) and ESB1 (Lahner et al., 2003; Baxter et al., 2009) orthologues which co-localised to these QTL regions. To identify allelic variants for one of these, CAX1, rapid cycling B. rapa R-O-18 M_3 TILLing mutant lines were successfully genotyped using a robust and extremely efficient high resolution melt (HRM) analysis technique. Thus, with the potential for improved and accelerated functional data from fast cycling Zn hyperaccumulators (Chapter 5), and the imminent release of the entire genomic sequence of B. rapa, such efficient and high throughput techniques should facilitate a more rapid dissection of the molecular mechanisms behind Zn accumulation in Brassica.

6.6.1. HMA4, CAX1 and ESB1 are candidates genes involved in zinc accumulation in Brassica rapa R-O-18

HMA4, CAX1 and ESB1 orthologues were selected as candidate genes for reverse genetic studies in Zn accumulation in Brassica rapa R-O-18 through comparative

genomic techniques. Loci in Brassica napus which were mapped to orthologous regions within the Arabidopsis thaliana genome (Parkin et al., 2005), were compared to recently published QTL associating with shoot Zn accumulation $[Zn]_{shoot}$ in B. oleracea (Broadley et al., 2010) and Zn trait data in B. rapa (Wu et al., 2008). Here it was found that $[Zn]_{shoot}$ associated with QTL on C2, C3, C5, C7 and C9 in B. oleracea (Broadley et al., 2010) and R3, R4 and R6 in B. rapa (Wu et al., 2008). Interestingly both R3 and C3 were previously shown to be syntenic chromosomes (Rana et al., 2004; Parkin et al., 2005) and so may contain conserved QTL involved in $[Zn]_{shoot}$.

Both C3 and R3 QTL intervals were similarly syntenic with N3 and N13 regions of Brassica napus (Parkin et al., 2005). A Cation Exchanger 1 (CAX1) (At2G38170) vacuolar localised high affinity calcium:hydrogen antiporter (Pittman et al., 2005) was present in the orthologous mapped region in Arabidopsis thaliana (Parkin et al., 2005). Interestingly Shigaki et al. (2003, 2005) demonstrated increased Zn accumulation in yeast cells expressing a 36 amino acid N-terminal truncation of AtCAX1 which was also consistent with similar results found in tomato (Park et al., 2005). In A. thaliana CAX1 was expressed primarily in leaf tissue with cax1 mutant lines showing 9% reduction in $[Zn]_{shoot}$ when grown in standard compost. A similar trend was observed with a cax1cax3 double mutant which demonstrated a 38% increase in $[Zn]_{shoot}$ grown on similar compost (Cheng et al., 2005). The authors speculated that these fluctuations in Zn could have been the result of either alterations in H^+ pumps or due to a lack of association between CAX transporters (Cheng et al., 2005).

QTL involved in $[Zn]_{shoot}$ in B. oleracea associated with a second region on C3 (Broadley et al., 2010) which was syntenic with both linkage groups N3 and N13 in B. napus that have been mapped to a region in A. thaliana containing P_{1B}-type ATPase 4 Heavy Metal Hyperaccumulator (HMA4) (Parkin et al., 2005). HMA4 (At2G19110) is associated with Zn accumulation throughout the Brassicaeae and has a role in maintaining Zn homeostasis through Zn transport into the root xylem (Hussain et al., 2004; Papoyan and Kochian, 2004; Verret et al., 2004; Hanikenne et al., 2008; Barabasz et al. 2010). Transcriptome expression studies have revealed constitutively high expression of HMA4 orthologues within the Zn

hyperaccumulators Noccaea caerulescens and A. halleri (Hammond et al., 2006; Talke et al., 2006), while expression of A. halleri HMA4cDNA under its endogenous promoter in A. thaliana resulted in increased Zn concentrations in xylem parenchyma, resembling typical Zn distribution in A. halleri roots (Hanikenne et al., 2008).

Both R4 and R3 Zn-QTL intervals in B. rapa (Wu et al., 2008) were found to map to Enhanced Suberin1 ESB1 (At2G28670) in the A. thaliana genome (Parkin et al., 2005). ESB1 has been strongly associated with altered $[Zn]_{shoot}$ in A. thaliana (Lahner et al., 2003; Baxter et al., 2009). It was demonstrated that expression of ESB1 was primarily located in the root endodermis including lower expression levels in the quiescent centre, stele and cortex (Baxter et al., 2009). Two independent loss of function mutants (esb1-1 and esb1-2) were shown to contain depositions of suberin in the endodermal root layer and in the root tip that were twice that of wild type (WT) plants. Further, peak transpiration rates were 15% that of WT, while $[Zn]_{shoot}$ was significantly reduced when compared to WT. Interestingly it was concluded that since suberin deposition reduced radial apoplastic transport in tandem with a significant reduction in $[Zn]_{shoot}$ then in agreement with White et al. (2002), a significant root apoplastic bypass may exist for Zn to reach the xylem (Baxter et al., 2009).

6.6.2. In silico comparative DNA analysis is a robust method to identify locus specific candidate gene sequences

Through a combination of complex in silico comparative genomics and traditional in vivo laboratory techniques, locus specific fragments for the R-O-18 inbred line of Brassica rapa A genome HMA4 locus a (BraA.HMA4.a), BraA.CAX1.a, BraA.ESB1.a, BraA.ESB1.b and BraA.ESB1.c were successfully isolated. A Basic Local Alignment Search Tool (BLAST) was initially employed to obtain locus specific Brassica gene orthologues of AtHMA4, AtCAX1 and AtESB1 GenBank/NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and the BrGSP BLAST Server (http://brassica.bbsrc.ac.uk/BrassicaDB/BrGSP_blast.html) databases were both interrogated to retrieve relevant Brassica sequence data. For all three genes, 2Mb fragments of A. thaliana sequence obtained from GenBank/NCBI containing the corresponding Arabidopsis gene of interest were

entered into two BLAST search programme, GenBank/NCBI, (using the algorithms BLASTn, megablast and discontiguous megablast at default settings (http://www.ncbi.nlm.nih.gov/blast/ producttable.shtml#tab31), and the BrGSP BLAST Server (operating a WuBLAST 2.0 algorithm at default settings (http://blast.wustl.edu/)).

For AtHMA4, no full length Brassica orthologous sequence was initially identified. Short sequences (<1 Kbp) from Expressed Sequence Tags (EST), Genomic Sequence Surveys (GSS) and Bacterial Artificial Chromosomes (BAC)-end sequences showing >80% sequence identity to the AtHMA4 region were selected and assembled into contigs using ContigExpress software (default parameters, VectorNTI, Invitrogen, UK) and aligned to the AtHMA4 locus using AlignX software (default parameters, VectorNTI, Invitrogen, UK) primer design. One 1641 bp R-O-18 locus specific HMA4 fragment (BraA.HMA4.a) was amplified from genomic DNA and confirmed through restriction endonuclease and sequencing analyses.

For AtCAX1 one full length sequenced BAC (KBrH004D08, GenBank: AC189543.2) from the closely related B. rapa Chiifu genome and short Brassica sequences (<1 Kbp) showing >80% sequence identity to the AtCAX1 region were assembled and formed contigs using ContigExpress software. Contigs were aligned to the AtCAX1 region, and primers amplified a locus specific 1028 bp CAX1 fragment (BraA.CAX1.a) from B. rapa R-O-18, which was cloned and confirmed through sequencing analyses.

For AtESB1 one full length sequenced BAC (KBrH104J01, GenBank: FP340660.1) from the B. rapa Chiifu genome showing >80% sequence identity to the AtESB1 region was aligned to the AtESB1 from which primers were designed. Three R-O-18 locus specific 1178 bp ESB1 fragments (BraR.ESB1.a, BraR.ESB1.b and BraR.ESB1.c) were amplified from genomic DNA, cloned and confirmed through sequencing analyses.

6.6.3. Targeting Induced Local Lesions IN Genomes (TILLing) is a robust method to identify locus specific mutated lines

Targeting Induced Local Lesions IN Genomes (TILLing) within the Brassica rapa R-O-18 population was an efficient approach to identifying allelic variants of BraA.CAX1.a. from a mutant population of 9216 plants. A total of 26 lines were selected which contained independent mutations in CAX1. Following identification and isolation of the locus specific CAX1 sequence, information was forwarded to RevGen UK (http://revgenuk.jic.ac.uk/). Here the sequence was initially processed to optimise primers to amplify regions more suitable to efficient TILLing using the Codons Optimized to Discover Deleterious LEsions (CODDLE) software (http://www.proweb.org/coddle/). From this, primers were thus designed to amplify a sequence within the first exon of BraA.CAX1.a. to identify early truncations from novel STOP codons and therefore possible gene knock-outs or splice variants and missense or silent amino acid changes within the gene sequence. Primers were re-examined by RevGen UK to confirm locus specificity before being employed for TILLing. Prior to seed dispatch the zygosity of all lines was confirmed, along with mutation coordinates and transitional changes i.e. from G-A or C-T. Previous observations from Arabidopsis (Greene et al., 2003) and the current B. rapa R-O-18 (Stephenson et al., 2010) TILLing populations reported difficulties detecting mutations in regions within 100 bp of each primer. In terms of BraA.CAX1.a., this resulted in designing CODDLE optimised primers 347 bp upstream of the translational start site in the promoter region and 141 bp downstream of the first exon (273 bp), thus producing a TILLing region of 760 bp. Similar to Stephenson et al. (2010) mutations in BraA.CAX1.a. were asymmetrically distributed along the amplicons, probably since the 5' primer was located 347 bp upstream in an AT-rich non-coding region. The 3' end primer was designed near to the first exon, in an area with more GC content and therefore higher mutation probability. A total of 26 M_2 lines were selected which contained CAX1 nucleic acid mutations. 50% produced missense mutations, one line or ~4% resulted in nonsense mutations, 20% initiated no amino acid change and 26% were located in non-coding regions. These figures compared favourably with those observed for six genes screened previously in this population (Stephenson et al., 2010). The mutation density within this population was calculated to be ~1 per 56 kb after normalisation to the

average 35% GC level for the B. rapa genome (Trick et al., 2009) representing 10,000 mutations per plant, which is the highest ever reported for any diploid plant or animal species (Stephenson et al 2010). Within the BraA.CAX1.a locus the mutation density was considerably higher at 1 per 29 bp. Gene redundancy from paralogues originating from paleopolyploidic events may have reduced the lethality of such a high mutation load (Blanc and Wolfe, 2004). In addition, a high degree of heterozygosity ~13:1 (heterozygote:homozygote), caused by pollen transfer between racemes derived from progenitor cells containing different sets of mutations from the same plant, may also contribute to such low mortality rates (Stephenson et al., 2010). Contrary to reports in some TILLing populations with high mutation loads, the B. rapa R-O-18 population has presented phenotypic evidence of gene manipulation which were comparable to those witnessed in orthologous knockout regions in A. thaliana (Stephenson et al., 2010). Transformation of Brassica remains an expensive and inefficient process and thus hampers T-DNA insertion mutagenesis. Furthermore, T-DNA and physical mutagenesis such as fast neutron bombardment may increase lethality by knocking out essential functions (McCallum et al., 2000; Colbert et al., 2001). Conversely, TILLing EMS mutagenised populations can provide an allelic series of phenotypic severity and permit a relatively sophisticated understanding of gene function (Till et al., 2003). To this end, seeds were ordered for all 26 M_3 mutant lines to initiate a back crossing population as a basis for future functional experiments.

6.6.4. High Resolution Melt curve analysis as an approach to genotyping

An efficient protocol to genotype Brassica rapa R-O-18 M_2 and M_3 TILLing mutant lines was established here. This chapter illustrated that a combination of novel nested PCR approaches and a High Resolution Melt (HRM) analysis was sufficiently sensitive to detect heterozygous single nucleotide polymorphisms (SNP) in locus specific genes of interest, despite the presence of multiple paralogous sequences. The method was applied to all CAX1 B. rapa R-O-18 mutants identified through TILLing the EMS mutated M_2 population. HRM involved the analysis of the dissociation of double-stranded DNA (dsDNA) PCR products, which had been saturated with the low PCR-toxic dye, EvaGreen. The method characterised the PCR products as they transitioned from dsDNA to

single–stranded DNA (ssDNA) (Dong et al., 2009). Highly concentrated dyes such as these, coupled with high DNA saturation, result in less redistribution of dye to non-denatured regions of the nucleic strands during melting, thus increasing the fidelity of measured fluorescent signals (Monis et al., 2005). In addition, EvaGreen has been shown to demonstrate an equal binding affinity for GC-rich and AT-rich regions, with no sequence preferences. These features improved DNA melt sensitivity and thus returned higher resolution melt profiles such that precise SNP detection in PCR products was feasible (Dong et al., 2009). In order to increase efficiency, two lengthy steps were removed. First, DNA isolation was performed using a rapid rudimentary extraction protocol in order to facilitate large population analyses. DNA was not further purified and a nested-PCR technique was employed as an efficient alternative to designing and testing locus specific HRM primers. The Lightscanner® version 2.0 (Idaho Technology, USA) software is commercially available for HRM analysis and was effectively employed to test and validate CAX1 mutations in these EMS mutated B. rapa R-O-18 M_2 and M_3 TILLing lines.

All M_3 and M_2 TILLing lines were genotyped for CAX1 SNP mutations by HRM analysis. A total of 26 lines, 19 M_3 and 26 M_2 (n=135) were grown and genotyped, and 59 individuals (> 37%) were identified as CAX1 mutants representing 24 independent cax1 lines. Seven M_3 lines however, including those containing a truncated CAX1, were not supplied by RevGen due to fecundity limitations. M_2 siblings were genotyped in their place and were shown to contain five out of the seven mutations of interest. Thus, it was shown that M_2 sibling lines, which have not been screened by TILLing, can be employed for genotyping in cases where seed supply is an issue. Results from a subset of lines genotyped by HRM (8 lines, 6 identified as SNP mutants and two as wild type, n=15) were corroborated by sequencing. No false positives were recorded from HRM within this subset, which is an improvement to that reported by Dong et al. (2009), where it was estimated that HRM returned a false positive rate between 2.8 and 7.1%. Zygosity was not determinable through HRM, however sequencing identified 69% heterozygous and 31% homozygous for the mutations compared with ~ 90% heterozygotes observed by Stephenson et al. (2010).

Here, a resourceful and effective approach to the identification and genotyping of locus specific allelic variants in Brassica rapa has been described. These methods, which encompassed comparative analysis, in silico data-mining and classical in vivo techniques, have efficiently isolated five locus specific B. rapa candidate Zn transporters, including a HMA4 homolog. Moreover, for one of these transporters, CAX1, 24 lines of independent allelic variants have been genotyped using HRM analysis with novel nested PCR approaches. These high throughput techniques can therefore be applied to routinely isolate locus specific candidate Zn transporters from important crop systems such as B. rapa. With the advent of a complete B. rapa genomic sequence release, and the availability of multi allelic TILLing populations, such reverse genetic approaches now have the potential to associate gene function to pre-existing Brassica forward screen datasets. Further, functional genetic data generated from across the Brassicaceae, including Zn hyperaccumulators such as N. caerulescens (Chapter 3 and 4) should enhance understanding of the mechanisms behind increased Zn accumulation in B. rapa. The creation of a fast cycling N. caerulescens genotype (Chapter 5) should more rapidly deliver such functional information and thus accelerate genetic improvement of Zn accumulation in Brassica crops. Together these approaches have the potential to compliment crop improvement trials, such as Zn biofortification strategies, and thus contribute to a sustainable reduction of human Zn malnutrition, which now affects over one third of humans globally.

6.7. SUMMARY

- One locus specific HMA4 gene BraA.HMA4.a was identified and cloned from Brassica rapa R-O-18.
- One locus specific CAX1 gene BraA.CAX1.a was identified and cloned from Brassica rapa R-O-18.
- Three ESB1 paralogues (BraA.ESB1.a, BraA.ESB1.b and BraA.CAX1.c) have been identified and cloned from Brassica rapa R-O-18.
- 26 lines have been successfully identified as containing EMS induced allelic variations in the CAX1 gene through TILLing assays on a population of 9,216 M_2 Brassica rapa R-O-18 mutants.
- A rapid DNA extraction protocol and nested PCR approach has removed the lengthy locus specific primer design for HRM genotyping.
- HRM genotyping has efficiently isolated 24 independent CAX1 genotypes from segregating M_2 and M_3 mutant populations for future downstream phenotyping and molecular genetic approaches.

CHAPTER 7 GENERAL DISCUSSION

7.1 GLOBAL ZN DEFICIENCIES AND MALNUTRITION

Zinc (Zn) is one of the most important metals in biological systems with roles as a cofactor in more than 300 enzymes and numerous transcription factors as well as maintaining the structural architecture of many vital proteins (Broadley et al., 2007; Palmgren et al., 2008). A report by the World Health Organisation (WHO) demonstrated that next to indoor smoke inhalation, Zn deficiency was ranked the fifth most serious risk affecting human health in the developing world and eleventh globally (WHO, 2002). The reason behind such deficiency lies in the fact that staple crops are cultivated on soils with low Zn availability or that their edible fractions contain inherently low Zn concentrations (Welch and Graham, 2002, 2004; White and Broadley, 2005, 2009). Solutions proposed to alleviate this problem include dietary diversification, mineral supplementation and food fortification however, due to a reliance on economic independence and political stability, these strategies have often failed to realise their potential (Bouis et al., 2003; White and Broadley 2005, 2009; Shrimpton et al., 2005; Stein et al. 2007). More immediate biofortification strategies have thus been proposed.

7.2 AGRONOMIC BIOFORTIFICATION STRATEGIES

Biofortification of crops, has been widely regarded as an immediate, sustainable and effective approach to delivering adequate Zn levels to human diet. Agronomic biofortification depends on the application of mineral Zn or the improvement of the solubilisation and mobilization of Zn in the soil (White and Broadley, 2009). Applications of Zn to soil or leaves has increased Zn concentrations in grains, leaves, tubers and fruits on many Zn unavailable soils (Rengel, 2001; Cakmak, 2002, 2004, 2008; Broadley et al., 2007). Evidence from Zn biofortification of cereals has demonstrated improved Zn status of vulnerable populations in Turkey (Cakmak, 2004). One major drawback however, is that Zn fertilisers require regular applications and are relatively costly to manufacture, distribute and apply. The sustainability of such an approach is also limited since current Zn reserves (480 million tonnes) have been estimated to reach exhaustion within 60 yrs at current consumption rates (Cohen,

2007; Kesler, 2007). Increasing the accumulation potential of crops through genetic biofortification could therefore benefit this approach.

7.3 GENETIC BIOFORTIFICATION IN THE BRASSICACEAE

The majority of economic analysts concur that genetic biofortification is the most sustainable and commercially viable approach to increasing dietary Zn (Stein et al., 2007; White and Broadley, 2009). However, within cereals, zinc appears to be preferentially stored as Zn-phytate complexes (phytate is a strong chelator of divalent cations) and thus significantly impedes Zn bioavailability (Cakmak, 2008; Palmgren et al., 2008). Additionally, recent studies have demonstrated a 'dilution effect' where concentrations of Zn have declined as higher yielding genotypes were developed (reviewed by White and Broadley, 2009). However this was not a constitutive trait and some studies observed no significant relationships between plant yield and micronutrient concentrations such as between Zn concentrations and bean yield in Phaseolus vugaris (Graham et al., 2001). In addition leafy vegetables such as Brassica generally accumulate relatively high concentrations of Zn (Welch and Graham, 2002; White and Broadley, 2005, 2009, Broadley et al., 2010). Moreover, a recent study by Broadley et al. (2008) demonstrated no significant negative relationship between mineral concentrations and biomass production within the majority of subtaxa of Brassica oleracea. In addition, the Brassicaceae harbour an enormous variation in $[Zn]_{shoot}$ among species, with some independently evolving hyperaccumulator species such as Arabidopsis halleri, Arabis alpina and Noccaea caerulescens that accumulate >10,000 mg Kg^{-1} DW which represents >100 fold more Zn than the majority of other species (Broadley et al., 2007; Verbruggen et al., 2009; White and Broadley, 2009; Krämer, 2010). This natural genetic variation in Zn accumulation in Brassicaceae thus represents a vital resource to assist in elucidating the molecular mechanisms behind Zn accumulation in crops.

7.4 ELUCIDATING GENETIC VARIATION IN ZN ACCUMULATION IN BRASSICACEAE

It was therefore decided to explore this genetic diversity in Brassicaceae to further dissect the molecular aspects associated with Zn accumulation. The first objective was to clone and sequence a candidate Zn transporter (HMA4) from the model Zn hyperaccumulator plant Noccaea caerulescens Saint Laurent Le Minier. This involved the creation of a genomic fosmid library of 36,864 clones containing approximately 40 Kbp genomic inserts, representing ~5-fold coverage of the ~250 Mbp genome. Sequencing was performed using Next Generation FLX 454 technology (Chapter 3). The second objective was to perform expression analysis of HMA4 promoters from N. caerulescens, A. halleri and A. thaliana to identify differences in regulation of these orthologues between hyperaccumulators and non-hyperaccumulators (Chapter 4). This was performed heterologously due to poor transformation efficiencies in N. caerulescens, thus the third objective was to test for an efficient transformation approach in this genotype (Chapter 5). A slow life cycle (9 months), as well as an obligatory 10 week vernalisation requirement has hampered research in N. caerulescens, and so, the fourth objective was to create a rapid cycling population using fast neutron mutagenic techniques (Chapter 5). In conjunction with this, a robust in vitro growth system was established as a fifth objective, to support efficient future assays of molecular and physiological responses to exogenous Zn supplies (Chapter 5). To efficiently transfer this molecular knowledge into commercially important Brassica rapa crops, a robust platform to identify, isolate and genotype locus specific allelic variants of Zn transporters was developed (sixth objective, Chapter 6). In tandem with this, comparative genomic approaches were employed to demonstrate that known Zn transporters (including HMA4) co-localised with QTL associated with Zn dependent traits from pre-existing datasets in Brassica rapa (Wu et al., 2008) and B. oleracea. (Broadley et al., 2010) (Chapter 6). The following sections will discuss in detail the results reported in this thesis.

7.5 NOCCAEA CAERULESCENS SAINT LAURENT LE MINIER AS A MOLECULAR MODEL FOR ZN ACCUMULATION

Noccaea caerulescens (J. & C. Presl.) F.K.Mey. Saint Laurent Le Minier was selected as a model to investigate the molecular genetics of Zn accumulation since it is a well characterised Zn hyperaccumulator, with the highest Zn concentration ever recorded in any plant species of 53,450 µg g^{-1} dry biomass (Reeves et al., 2001). It is functionally nonmycorrhizal (Regvar et al., 2003) and all members of this genus constitutively hyperaccumulate metals (Broadley et al., 2007; Verbruggen et al., 2009; Krämer, 2010). It is considered to be a model hyperaccumulator species due to its relatively small diploid genome (2n=2x=14) and its inherent ability to hyperaccumulate Zn, Cd and Ni. Unlike the related Zn/Cd hyperaccumulator, Arabidopsis halleri, N. caerulescens is fully self compatible and is more closely related to the Brassica spp. than Arabidopsis thaliana, from which it diverged some 20 million years ago (Mya) and shares 88% sequence identity (Peer et al., 2003; Broadley et al., 2007; Krämer, 2010). In addition, N. caerulescens has been shown to demonstrate significantly more inter- and intra- population variation in Zn accumulation and Zn tolerance than A. halleri (Verbruggen et al., 2009; Krämer, 2010). Populations from soils polluted with Zn, Cd and lead (Pb) and from non-contaminated soils have been shown to all hyperaccumulate Zn, while Cd hyperaccumulation has also been observed in the majority of populations on contaminated soils (Baker and Brooks, 1989; Reeves et al., 2001). In contrast to A. halleri, some accessions from ultramafic and non-contaminated soils demonstrated no Cd hypertolerance and for some, higher exogenous supplies of Zn and Cd have been associated with low $[Zn]_{shoot}$ or $[Cd]_{shoot}$ respectively (Meerts et al., 2003; Broadley et al., 2007; Verbruggen et al., 2009; Kramer et al., 2010). Also distinct from many species, N. caerulescens demonstrated higher Cd concentrations in leaves than in roots (Assunção et al., 2003) and hypertolerance was confined to populations from calamine soils (Roosens et al., 2003). A number of quantitative trait loci have been identified which account for altered Zn and Cd accumulation in N. caerulescens, including one common locus for Cd and Zn in roots, one for Zn in roots and Zn in shoots and one for Cd in roots and Cd in shoots (Assunção et al.,

2006; Deniau et al., 2006). Unfortunately candidate genes associating with these QTL have not yet been identified.

7.6 *HMA4* AS A CANDIDATE FOR MOLECULAR ANALYSIS OF ZN ACCUMULATION IN PLANTS

Zinc is usually acquired by plants by transporting Zn^{2+} or Zn-phytosiderophores complexes from the soil solution, across the root cell membranes, and both symplastically and apoplastically towards the xylem (White et al., 2002; Broadley et al., 2007). Functional and expression based evidence has indicated that the majority of Zn^{2+} influx to the cytoplasm was mediated by ZIP genes (Pence et al., 2000; Assunção et al., 2001; Colangelo and Guerinot, 2006; Talke et al., 2006 Broadley et al., 2007) and members of the cation diffusion facilitator (CDF) family, such as metal tolerance protein MTP1 and MTP3, which transport Zn into the vacuole via a Zn^{2+}/H^{+} antiport mechanism (Arrivault et al., 2006; Hammond et al., 2006; Krämer, 2010). For Zn to accumulate into shoot tissue however, it must be actively transported into the xylem stream. Strong evidence has associated P_{1B}-type ATPase Heavy Metal Associated (HMA4) orthologues with this role on the basis of expression and functional data from yeast (Papoyan and Kochian, 2004), from A. thaliana (Verret et al., 2004; Hussain et al., 2004; Mills et al., 2004; Siemianowski et al., 2010) and from the Zn hyperaccumulators A. halleri (Talke et al., 2006; Hanikenne et al., 2008; Barabasz et al., 2010) and N. caerulescens (Hammond et al., 2006; van de Mortel et al., 2006). Similarly, HMA4 expression was shown to be constitutively more highly expressed in both N. caerulescens (Hammond et al., 2006) and A. halleri (Talke et al., 2006) than Zn non-hyperaccumulators for all Zn concentrations tested and was induced further under Zn deficiency (A. halleri) and under very high $[Zn]_{ext}$ (150 – 300 μM) in N. caerulescens. Despite such overwhelming evidence associating HMA4 with Zn accumulation, the Noccaea caerulescens genomic ortholog has remained unsequenced.

7.7 SUCCESSFUL SEQUENCING OF THE *HMA4* IN NOCCAEA CAERULESCENS

A Genome Sequencer FLX 454 sequencing approach employed to sequence all clones obtained from a novel Noccaea caerulescens Saint Laurent Le Minier

fosmid library (Chapter 3). N. caerulescens HMA4 (NcHMA4) orthologues were identified by hybridising a radiolabelled NcHMA4 probe, generated by PCR using HMA4 cDNA data published from the Prayon ecotype (Papoyan and Kochian, 2004) to the N. caerulescens fosmid library. All orthologues were confirmed by PCR and Southern blotting (Section 3.5). Pyrosequencing of five fosmids and in silico assembly of their contiguous sequences using Vector NTI 11 software revealed tandem quadruplication of NcHMA4 in a locus syntenic with that of both Arabidopsis halleri and A. thaliana (Section 3.5.11). This tandem multiplication event is remarkably similar to a tandem triplication of HMA4 which evolved independently in the Zn hyperaccumulator A. halleri (Hanikenne et al., 2008) and thus could associate HMA4 multiplication with Zn hyperaccumulation. This is consistent with higher HMA4 expression levels observed between Zn hyperaccumulators and closely related Zn non-hyperaccumualtors i.e. expression of HMA4 expression was higher in N. caerulescens and A. halleri than closely related Zn non-hyperaccumulators Thlaspi arvense (Hammond et al., 2006) and A. lyrata (Talke et al., 2006) respectively. At a protein level, HMA4 amino acid sequences from N. caerulescens, A. thaliana and A. halleri showed relatively high conservation, sharing up to 84% protein sequence identity, suggesting conserved functional roles. This was consistent with observations by Krämer (2010) where it was concluded that there was no evidence for a difference between the protein functions of HMA4 from both A. halleri and A. thaliana. Similarly a high proportion of genes involved in enhanced Zn accumulation have undergone gene copy number expansion including Zn deficiency response genes such as metal transport proteins 1 (MTP1), which is a Zn/H^+ antiporter and involved in sequestering Zn inside leaf cells. An A. halleri allele was shown to co-segregated with Zn hypertolerance in BC_1 of a cross between A. halleri and A. lyrata. This accounted for over 20 fold higher MTP1 expression in A. halleri than A. thaliana (Krämer et al., 2010) and was also shown to be similarly highly expressed in N. caerulescens (Assunção et al., 2001). Duplication of other Zn deficiency response genes including ZIP3 and ZIP9 in A. halleri and ZIP4/ZNT1-2 and IRT1 in N. caerulescens, which have roles in Zn uptake into root cells, xylem unloading and cellular uptake in leaf tissue, could be a contributing factor to their high expression levels in Zn hyperaccumulators (Verbruggen et al., 2009). To

fully elucidate the role of these HMA4 orthologues an investigation of A. halleri mutants with silenced HMA4 expression induced by HMA4RNAi knock-down constructs was performed in A. halleri (Hanikenne et al., 2008). These studies demonstrated that $[Zn]_{leaf}$ in mutant plants was reduced to between 12 – 35% of WT plant levels and more resembled that of A. thaliana (Hanikenne et al., 2008). In contrast the roots of HMA4RNAi mutant A. halleri plants were shown to accumulate 49 – 134 fold higher Zn concentrations than WT plants which again resembled concentrations found in A. thaliana roots (Hanikenne et al., 2008). It was also observed that partitioning and accumulation of metals other than Zn and Cd were not consistently altered in these lines. In A. thaliana, similar studies using gene knock-out lines generated by T-DNA insertions, demonstrated that plants accumulated two fold less Zn than both wild type (WT) and HMA2 knock-out plants (Hussain et al., 2004). Remarkably in double HMA4,HMA2 knock-out A. thaliana mutants, plants contained two fold less Zn in the leaves but two fold more Zn in the roots than WT, thus indicating a role in transporting Zn to shoots but not in Zn uptake (Hussain et al., 2004). Additionally, expression of Zn deficiency response genes including IRT3 and ZIP4 were correlated with HMA4 expression (r = 0.99 and 0.98, respectively) and thus decreased in the roots of HMA4RNAi A. halleri mutant lines (Hannikenne et al., 2008). This indicated that their high expression in hyperaccumulators appeared to be a secondary consequence of increased HMA4 expression. Together these results provide strong evidence for the role of HMA4 in enhanced Zn flux from root to shoots in both A. halleri and A. thaliana. With such substantial evidence supporting its role in enhanced $[Zn]_{shoot}$, it was decided to adopt a similar HMA4RNAi mediated gene knock-down approach to test the function of HMA4 in N. caerulescens.

7.8 AGROBACTERIUM TUMEFACIENS MEDIATED TRANSFORMATION OF NOCCAEA CAERULESCENS

An Agrobacterium tumefaciens mediated transformation approach was thus attempted to integrate a NcHMA4RNAi construct into the Noccaea caerulescens genome to test if reduced HMA4 expression associated with altered $[Zn]_{shoot}$ as described in other genomes (Hussain et al., 2004; Hanikenne et al., 2008). To facilitate high throughput molecular genetic approaches, an efficient floral dip transformation technique was attempted (Clough and Bent, 1998; Peer et al.,

2003). The NcHMA4RNAi construct which (donated by Dr. Victoria Mills) was created by ligating two 400 bp HMA4 cDNA fragments into a pK7GW1WG2 (II) vector using the Gateway® system (Section 2.5.3.2) and was cloned into the A. tumefaciens strain C58. As a transformation control a second construct, the GUS::NPTII construct was employed which contained NOS.NPTII.NOS and CaMV35S.GUS-intron sequences between the left and right borders (Section 2.5.3.1) and was cloned into the A. tumefaciens 1065 strain. Following repeated optimally defined transformation attempts, no T_1 transformant was recovered. The lack of primary transformants was shown not to be due to incompetent vectors or bacteria, since these were fully functional when transformed into A. thaliana. All successfully transformed T_1 A. thaliana plants demonstrated antibiotic resistance and through PCR, T-DNA integration from both vectors was confirmed (Section 5.5.7). Previous transformation studies in Noccaea caerulescens have shown that certain ecotypes including N. caerulescens Le Bleymard were not compatible with floral dip mediated transformation approaches (Peer et al., 2003). They found that the majority of silicles were aborted and that seed fecundity was drastically reduced following floral dip in this ecotype. A number of other factors which could have contributed to transformation failure in this study include possible mildew contamination during the initial days after the dipping treatment, a result of maintaining high humidity by placing pots into plastic sleeves (Section 5.4.3.5). The floral morphology of Noccaea caerulescens may also be incompatible with floral dip transformation. Floral dip transformation experiments in A. thaliana revealed that the female reproductive organs were the primary targets for A. tumefaciens mediated T-DNA integration (Desfeux et al., 2000). Within A. thaliana it was shown that the gynoecium developed as an open, vase-like structure and which, approximately 3 days prior to anthesis, fused to form closed locules at which point transformation could not be achieved (Desfeux et al., 2000). These experiments were assayed using CRABS-CLAW (crc-1) A. thaliana mutants which maintained an open gynoecium due to incomplete carpel fusion at the apex. It could be that the gynoecium of some Noccaea caerulescens ecotypes remains fused during its development and that this results in poor transformation efficiencies. One possible method to increase transformation efficiency could be to create a CRABS-CLAW mutant homolog in Noccaea caerulescens through

physical (Chapter 5) or chemical (Chapter 6) mutagenesis techniques. If successful this would provide a vital genetic resource for downstream reverse genetic and physiological studies. Alternatively a less efficient tissue culture approach could be employed, as demonstrated by Guan et al. (2008). Here shoot nodes from in vitro generated N. caerulescens plantlets were excised and incubated in vitro on media supplemented with a range of plant growth regulators to differentiate over a one week period, then transferred to a second media which was replaced over 3 week intervals to initiate shoot clusters to be submerged in A. tumefaciens solutions, harbouring the construct of interest (Guan et al., 2008). Although successful as a transformation procedure, the process was lengthy, and labour intensive. Therefore this method is unlikely to complement high throughput strategies for in planta molecular analyses and could additionally introduce epigenetic effects and somaclonal variation into plants (Vazquez and Linacero, 2010).

An additional explanation to increased expression of specific genes in Zn hyperaccumulators could be due to deregulation of the root Zn deficiency response (Talke et al., 2006; Verbruggen et al., 2009). This was observed in A. halleri where substantially higher expression levels were found for FRD3, ZIP4 and MTP8 genes which were not associated with more than a single genomic copy (Talke et al., 2006). Induced overexpression levels of Zn deficiency response genes measured in Zn hyperaccumulators is far greater than homologues in Zn non-hyperaccumulators, however how plants gauge their metal status or whether there are metal sensor proteins in plants is not fully understood (Verbruggen et al., 2009). Similarly characterisation of tissue specificity for genes involved in Zn transport in hyperaccumulators is in its infancy. Distribution of metals in leaves between hyperaccumulator species is not conserved and differences could exist between populations in the tissue specific accumulation of the same metal. e.g. for N. caerulescens the highest concentrations of Cd and Zn were found in epidermal vacuoles (Ma et al., 2005) but in A. halleri this was located in mesophyll cells (Kupper at al., 2000). Restricted understanding of tissue specific expression patterns or cellular localisation of gene products greatly limits the prediction of their role in hyperaccumulation (Verbruggen et al., 2009). The complexities involved in

metal hyperaccumulation are currently poorly understood both at the tissue and subcellular level and it is probable that general metal homeostasis processes in all endomembranes are modified in hyperaccumulators (Verbruggen et al., 2009). Poor transformation efficiency currently hampers the understanding of molecular genetics in N. caerulescens, however future possibilities to transform specific cell types should help decipher tissue specific adaptations to extreme environments (Verbruggen et al., 2009). Moreover previous trials elucidating tissue specificity and expression patterns of genes including those from wheat (Tittarelli et al., 2007) and A. halleri (Hanikenne et al., 2008), have demonstrated that gene marker activity was quantitatively and qualitatively similar under both heterologous (A. thaliana) and homologous transformation systems. With limited efficient stable transformation systems available for N. caerulescens, a heterologous approach was thus adopted using A. thaliana.

7.9 ANALYSIS OF CIS REGULATION OF NCHMA4

Upon initial examination of genomic data obtained from pyrosequencing and subsequent locus assembly (Chapter 3) it was shown that 5' promoter regions of HMA4 between all A. thaliana, A. halleri and N. caerulescens exhibited tremendous variation. Promoter regions from all four N. caerulescens shared between just 40 – 51% identity with homologous regions in A. thaliana and A. halleri. To test if sequence variation resulted in altered gene expression, series of promoters fused with GUS constructs were created. Promoter fragments which were PCR amplified from upstream regions of the translational start site of AtHMA4, AhHMA4-3 and NcHMA4-2 were ligated into pGWB3 destination vectors using the Gateway® cloning system (Section 4.4.1). These pGWB3 vectors were employed since the expression system was compatible with a wide range of marker genes for promoter expression analyses including *β*-glucuronidase (GUS), luciferase (LUC), synthetic green fluorescent protein with the S65T mutation (sGFP), enhanced yellow fluorescent protein (EYFP), and enhanced cyan fluorescent protein (ECFP) (Nakagawa et al., 2007). In addition, both functionality and reproducibility of this expression vector was verified through comparative promoter::GUS analysis of qualitative and quantitative GUS activity using a previously characterised phosphate induced promoter (PHT1) promoter::GUS construct (Nakagawa et al., 2007). Observations of transformed

A. thaliana lines growing on agar-based growth media supplemented with 15 μM Zn showed variation in both expression intensities and spatial representation of GUS activity between A. thaliana and both Zn hyperaccumulators (Chapter 4). GUS under the control of HMA4 promoters from both N. caerulescens (NcHMA4-2p) and A. halleri (AhHMA4-3p) was expressed more intensely in both root and shoot tissues compared with AtHMA4p. Spatially all three constructs were expressed in the root vasculature supporting similar previous expression analyses using AtHMA4p::GUS in A. thaliana, where it was shown that GUS activity was localised to endodermal and pericycle cell layers in the roots (Hussain et al., 2004). Subsequent localisation of Zn in these lines using the Zn fluorophore, Zinpyr-1 identified a coincidence of Zn content in these regions (Sinclair et al., 2007). This was substantiated by Hanikenne et al. (2008) who also demonstrated that all three AhHMA4p::GUS were localised in xylem parenchyma and pericycle cell layers in both A. thaliana and A. halleri, which was further resolved using in situ HMA4 mRNA hybridisations. Together these results strongly associated HMA4 from A. thaliana and A. halleri with xylem loading of Zn and the current study indicated that NcHMA4-2 could have a similar role in N. caerulescens (Bernard et al., 2004; Hussain et al., 2004; Papoyan and Kochian, 2004; Verret et al., 2004, 2005; Hanikenne et al., 2008). However, only AhHMA4-3p and NcHMA4-2p exhibited expression in leaf tissue which appeared to be widespread throughout the lamina. This was similarly observed by Hanikenne et al. (2008) for all three AhHMA4 promoters that demonstrated high expression in xylem parenchyma and cambium cell layers, consistent with the putative role for HMA4 in Zn exclusion from specific cell types and distribution within the leaf (Hussain et al., 2004; Talke et al., 2006; Hanikenne et al., 2008). Moreover, such extensive expression patterns throughout the leaf were consistent with highly concentrated Zn sequestration patterns reported in mesophyl vacuoles and cell walls of epidermal cells in N. caerulescens (Frey et al., 2000; Ma et al., 2005). NcHMA4-2 could therefore be associated with increased Zn distribution throughout the lamina of N. caerulescens plants. The present study has demonstrated that heterologous AhHMA4-3p expression in A. thaliana was consistent with observations recorded by Hanikenne et al. (2008) and could therefore by employed as a robust expression system for NcHMA4 analyses until an efficient transformation system

is established for Noccaea caerulescens. Due to time constraints, expression of the remainder of these NcHMA4 promoters could not be assessed, however it is hoped that the scientific community will progress this research using the methods established in this thesis. Further to these initial trials, it may be more informative to assay these lines on a range of Zn concentrations to gain a more comprehensive picture of Zn regulation and HMA4 expression in response to Zn. To facilitate this work an in vitro tolerance screen was performed for both Zn hyperaccumulator (N. caerulescens) and non-hyperaccumulator (A. thaliana) plant systems.

7.10 IN VITRO ZN TOLERANCE

To fully understand the genetic components of a trait the environment must be as homogenous and replicable as possible. Ex vitro environmental conditions can be confounded by variation in light intensities and lack of homogeneity in growth media such as peat based-compost batches, thus impeding on result reproducibility. An in vitro approach was thus tested, since it offered maximal control over environmental conditions; it was reproducible; it required minimal space for experimental assays, and required little input in terms of infrastructure and labour (Chapter 5). Plants were cultured on agar-based growth media supplemented with phytonutrients as described by Murashige and Skoog (1962) (MS media) but with increasing Zn concentrations from 0 – 6000 μM Zn for Noccaea caerulescens and 0 – 300 μM Zn for Arabidopsis thaliana. It was shown that the concentration for optimal A. thaliana growth required 30 μM Zn supplementation, but healthy plants could be obtained on as little as 1 μM Zn over a two week period. Results from this trial also demonstrated that to grow normally, N. caerulescens required Zn supplementation levels which were lethal to A. thaliana (>300 μM Zn) and >10 fold higher than that supplied by standard MS media recipes. The majority of previous observations of Zn dependent growth responses in N. caerulescens have involved either hydroponic (Pollard and Baker, 1996; Roosens et al., 2003) or soil (Escarré et al., 2000; Whiting et al., 2000) based tolerance screens using large populations. However in agreement with these screens, under in vitro conditions N. caerulescens did demonstrate tolerance to high Zn concentrations. Concurring with the current study's findings, a recent investigation demonstrated that under in vitro

conditions, N. caerulescens appeared to grow optimally when supplemented with 300 μM Zn (Hammond et al., 2006). The control Zn non-hyperaccumulator used for their screen was Thlaspi arvense L., which demonstrated optimal growth at 30 μM Zn and lethality at 300 μM Zn, which was consistent with observations for A. thaliana in the current study. In contrast however, the current study examined a greater range of Zn concentrations (9 for A. thaliana and 13 for N. caerulescens) including Zn deficient agar (0 μM Zn) in order to obtain a more thorough Zn dependent growth response for each species. These results reaffirmed N. caerulescens high Zn tolerance under in vitro conditions and demonstrated the reproducibility of in vitro growth systems among independent experiments. The growth system established here thus provided a fully regulated environment to confidently perform molecular genetic and physiological assays in both Zn hyperaccumulator and Zn non-hyperaccumulator plants. This system should similarly complement reverse genetic investigations into Zn dependent traits associated with endogenous genes. To maximise efficiency of such systems a rapid cycling model Zn hyperaccumulator was required, however as yet, none has previously been reported.

7.11 FAST CYCLING ZN HYPERACCUMULATING NOCCAEA CAERULESCENS

Effective plant research demands fast cycling and self compatible model plant systems, such as Arabidopsis thaliana, which are simple to culture and require minimal management (Clough and Bent, 1998; Peer et al., 2003, 2006). However, wild type Noccaea caerulescens is currently unsuitable for such high throughput research since it has an extensive cycle time (9 months) and requires vernalisation (10 weeks) to initiate flowering (Peer et al., 2003, 2006). It was hypothesised that genetically stable, fast cycling N. caerulescens mutant lines could be identified by screening a large population of mutagenised individuals. Since N. caerulescens were recalcitrant to A. tumefaciens mediated transformation, T-DNA mutagenesis was not attempted (Section 5.5.6). Fast neutron (FN) mutagenesis was selected, since it was more likely to produce null alleles through physical deletions of > 100 kb and therefore result in fewer allelic variants of undesirable genes as characterised by chemical mutagenesis such as Ethyl methanesulfonate (Kodym & Afza, 2003; Salt et al., 2008; Bruce et al., 2009). In addition, deletions of large fragments through FN mutagenesis should

knock-out tandem arrays of undesirable genes which would prove extremely challenging, if not impossible to silence by combining single mutants through sexual crosses (Koornneef et al., 1982; Bruce et al., 2009). From an initial 5500 M_1 population of seeds irradiated at 60 Gy, a single pool of 80,000 M_2 N. caerulescens were screened for fast cycling phenotypes in the absence of vernalisation. In contrast to levels of albinism reported in A. thaliana (2%) (Koornneef et al., 1982) and Medicago (2.57%) (Rodgers et al., 2009), a comparatively low 0.49% was observed in this study. M_1 N. caerulescens plants exhibited a much higher survival rate (63.2%) than the 50% believed to indicate optimal saturation among plant genomes (Koornneef et al., 1982; Rogers et al., 2009). A total of 35 M_2 lines exhibited floral initials in the absence of vernalisation, and nine flowered within 12 weeks and produced seed. Two lines were selfed to M_4, $A2M_4$ and $A7M_4$, and analysed for mineral uptake and cycle time. $A7M_4$ plants initiated flowering and seed development approximately 3 weeks later than $A2M_4$ lines and may therefore represent an independent deletion event in the N. caerulescens genome. However, both lines produced fruit within 13 weeks after sowing, which represented a 64% reduction in cycle time, and thus the possibility to produce up to four generations of seed in one year independent of vernalisation. Crucially both lines maintained the Zn hyperaccumulation levels as observed in wild type (WT) plants. Concentrations of the majority of other phytonutrients were lower in both fast cycling lines, however this may be a result of altered phenology, where early maturation could have resulted in reduced mineral acquisition (Nord and Lynch, 2009; White and Broadley, 2009). Such early maturation may also explain the observed reduction in dry weights for both lines in comparison to WT, however fresh weight to dry weight ratios demonstrated no significant differences between lines. Importantly both mutant lines were stable, demonstrating self fertility, no seed dormancy, and no intra-line variability in biomass, flowering time or mineral content. Elucidating the mechanism controlling fast cycling N. caerulescens lines has not yet been attempted, however it is clear that deletion mutagenesis has removed the genetic requirement for this trait. Attempts to introduce faster cycling phenotypes into A. thaliana (Michaels and Amasino, 1999; Sung and Amasino, 2004) demonstrated that loss of function of FLOWERING LOCUS C (FLC) and FRIGIDA (FRI) orthologues resulted in earlier flowering phenotypes. Indeed,

observations in Beta vulgaris (Reeves et al., 2007), Arabis alpina (Wang et al., 2009) and Brassica rapa (Zhao et al., 2010) demonstrated a conserved role for FLC homologues in regulating flowering times. It may be therefore, that fast neutron bombardment has resulted in a deletion of the Noccaea orthologues in both fast cycling lines. Although not attempted here, DNA hybridisations to tiling or exon arrays designed for Arabidopsis (Mockler et al., 2005; Hammond et al., 2006) or novel Brassica arrays (Love et al., 2010) using heterologous-cross-species based approaches as well as high throughput transcriptome sequencing may help elucidate the genetic control of vernalisation in N. caerulescens.

This study has thus described a robust protocol to develop stable fast cycling, vernalisation independent Zn hyperaccumulating Noccaea caerulescens lines from a relatively small initial population of 5500 seeds. It is anticipated that this fast cycling resource will be employed by the scientific community as a model Zn hyperaccumulator species for future molecular genetic and physiological assays. Once adapted to efficient transformation techniques, this resource should accelerate our current understanding of natural genetic variation in Zn accumulation in Brassicaceae. The efficient transfer of such knowledge to important crop species such as Brassica rapa however, requires the establishment of an effective reverse genetic crop system. Currently, the majority of scientific understanding of mineral accumulation in Brassica spp. relies on classical genetic techniques such as quantitative trait loci (QTL) analyses (Broadley et al., 2008, 2010; Wu et al., 2008). To resolve such traits to individual genes, these approaches require inefficient fine mapping strategies (Bruce et al., 2009). However, recent advances in comparative genomics are now permitting a more rapid access to novel genetic regulation. Comparative genomic approaches were therefore tested to access novel candidate genes involved in Zn transport in Brassica rapa.

7.12 REVERSE GENETIC APPROACHES IN BRASSICA RAPA

To effectively exploit reverse genetic analyses, locus specific candidate gene sequences must be easily isolated and characterised from the studied organism. This is hampered in Brassica spp. by the presence of paralogues from palaeopolyploidic events (Blanc and Wolfe, 2004) and a lack of completely assembled genomic sequence data. However, large conserved collinear regions between Brassica napus (an allotetraploid of B. rapa and B. oleracea) (Rana et al. 2004), and the fully sequenced A. thaliana (Parkin et al., 2005) were exploited to identify and isolate locus specific sequences of candidate Brassica rapa Zn transporters. QTL associating with altered $[Zn]_{shoot}$ (Zn-QTL) in B. oleracea on chromosomes C2, C3, C5, C7 and C9 (Broadley et al., 2010) and chromosomes R3, R4 and R6 in B. rapa (Wu et al., 2008) were compared to B. napus loci mapped to A. thaliana regions. Remarkably, chromosome 3 from both B. rapa and B. oleracea (R3 and C3 respectively) which were previously shown to be syntenic (Rana et al., 2004; Parkin et al., 2005), both contained significant Zn-QTL (Wu et al., 2008; Broadley et al., 2010). Further comparative analysis demonstrated that these Zn-QTL mapped to regions in A. thaliana containing known Zn transporters including, HMA4, CAX1 and ESB1 (Chapter 6). Loss of function of the Enhanced Suberin1 (ESB1) (AT2G28670) homolog in A. thaliana has been strongly associated with reduced $[Zn]_{shoot}$ (Lahner et al., 2003; Baxter et al., 2009) (Section 6.6.1). A 36 amino acid N-terminal truncation of the Cation Exchanger 1 (CAX1) (AT2G38170) vacuolar localised high affinity calcium:hydrogen antiporter was shown to increase Zn accumulation in yeast (Shigaki et al., 2003, 2005) and in tomato (Park et al., 2005) however, loss of function in A. thaliana was associated with a 9% reduction in $[Zn]_{shoot}$ (Cheng et al., 2005) (Section 6.6.1). Together these analyses demonstrated that candidate Zn transporter genes co-localised with QTL associated with altered $[Zn]_{shoot}$ in B. rapa and B. oleracea and suggested that chromosome 3 could contain vital genetic regulators of $[Zn]_{shoot}$ in both Brassica species.

To isolate locus specific sequences of these genes for downstream reverse genetic assays a matrix of novel in silico and traditional in vivo tools were employed. Basic Local Alignment Search Tool (BLAST) obtained locus specific Brassica

gene orthologues of AtHMA4, AtCAX1 and AtESB1 from GenBank/NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and the BrGSP BLAST server (http://brassica.bbsrc.ac.uk/BrassicaDB/BrGSP_blast.html) databases. Although the fully assembled B. rapa genome sequence is not publically available, partially or fully sequenced cloned Brassica spp. fragments were assembled using the Vector NTI 11 software to form contiguous gene sequences (Section 6.4.3.1). This data was thus employed to design locus specific gene primers to PCR amplify and sequence B. rapa R-O-18 homologues. Five locus specific alleles, Brassica rapa A genome HMA4 locus a (BraA.HMA4.a), BraA.CAX1.a, BraA.ESB1.a, BraA.ESB1.b and BraA.ESB1.c were successfully isolated and sequenced. In order to execute effective molecular analyses of phenotypes associated with these Zn transporters it was necessary to efficiently mutate, identify and isolate in planta. T-DNA transformation techniques, such as those employed in A. thaliana, are relatively expensive and inefficient in Brassica. In addition other methods of physical mutagenesis such as fast neutron bombardment have been shown to increase lethality by inducing loss of essential functions (McCallum et al., 2000; Colbert et al., 2001). A Targeting Induced Local Lesions IN Genomes (TILLing) approach was thus adopted by screening M_2 DNA from 9216 Brassica rapa R-O-18 plants derived from a population mutated with ethyl methanesulfonate (EMS) (Stephenson et al., 2010). Such mutagenesis was preferred since EMS introduced single nucleotide polymorphisms (SNP) that caused base pair transitions from G-A or C-T, which resulted in both loss of function and gain of function mutants leading to null, hypomorphic, neomorphic or hypermorphic phenotypes (Gilchrist and Haughn, 2010). Such mutations should therefore offer a more sophisticated understanding of the functional roles of genes in Zn accumulation (Till et al., 2003).

This population was thus screened by TILLing M_2 DNA (Stephenson et al., 2010) (Section 6.4.4) for allelic variants of the locus specific BraA.CAX1.a and identified 26 independent M_3 mutant lines. Initial examination illustrated that 50% produced missense mutations, ~4% resulted in nonsense mutations, 20% initiated no amino acid change and 26% were located in non-coding regions. These results revealed a broad range of allelic mutants in BraA.CAX1.a and presented an excellent platform for future reverse genetic analysis of associated

Zn phenotypes in this genome. The mutation density within the 760 bp BraA.CAX1.a locus was ~1 per 29 kb which was considerably higher than that reported previously (1 per 56 bp) by (Stephenson et al., 2010). Low mortality rates of these lines containing such high mutation densities (10,000 per plant) could be due to gene redundancy from paleopolyploidic paralogues and a high degree of heterozygosity ~13:1 (heterozygote:homozygote) (Blanc and Wolfe, 2004; Stephenson et al., 2010). Additionally this population has demonstrated previous phenotypic evidence from allelic variants which were comparable to orthologous gene knockouts in A. thaliana (Stephenson et al., 2010), thus confirming its suitability to reverse genetic molecular assays.

To identify SNP allelic variants in large mutant populations a high throughput and easily adaptable genotyping approach was required. Robust strategies such as Sanger sequencing and Taqman SNP detection are relatively inefficient and too expensive for large population screens (Martino et al., 2010). Others, such as Cleaved Amplified Polymorphic Sequence (CAPS) marker assays, require specific DNA sequences to function (Parsons et al., 1997). In contrast, High Resolution Melt (HRM), represented a robust, precise, efficient but inexpensive genotyping option (Garrirano et al., 2009; Martino et al., 2010). To further increase the efficiency of this approach, novel high throughput DNA extraction (Section 6.4.5.1) and nested PCR approaches (Section 6.4.5.3) were established, thus reducing handling time and lengthy locus specific primer design steps. DNA was thus amplified with locus specific CAX1 primers before a nested HRM PCR reaction that incorporated EvaGreen intercalating dye (Section 6.4.5.2). Using the Lightscanner® version 2.0, dissociation of double stranded PCR fragments was plotted as a function of temperature forming melt curves characterised by EvaGreen fluorescence (Dong et al., 2009). Allelic variation between plants was identified as alterations in melt curves, indicating differences in melting temperatures induced by base pair changes (Martino et al., 2010). Using this method, 96 HRM samples were genotyped in 10 min with only a modest cost incurred through purchasing a HRM PCR kit (Section 6.4.5.2). This novel approach to genotyping B. rapa represented significant economic and labour efficiencies when compared to all other robust genotyping strategies. In terms of SNP identification, from a total of 135 M_2 and M_3 plants, 59 individuals (> 37%)

representing 24 out of 26 CAX1 lines were identified. Fifteen individuals (13 mutant and 2 WT plants) were subsequently sequenced, and confirmed that all 13 lines identified as mutants by HRM, did contain SNPs in the specified locations. No false positives were identified as was reported by Dong et al. (2009) using HRM techniques. Furthermore, 69% were identified as heterozygous and 31% as homozygous for these SNPs, thus HRM was shown to robustly identify heterozygous mutants and could therefore be confidently employed for genotyping future heterologous mutant backcrossed lines.

An efficient and effective protocol to identify, isolate and genotype locus specific candidate Zn transporters for downstream reverse genetic analyses in B. rapa have been described. It is anticipated therefore, that the scientific community will adopt these techniques to efficiently perform future molecular genetic assays within commercially important Brassica spp. crop systems. Moreover, these approaches should become increasingly appropriate to high throughput molecular analyses as we approach the release of a completely assembled and fully sequenced B. rapa genome.

7.13 DIRECTIONS FOR FUTURE RESEARCH

Zn deficiency affects over two billion people worldwide and is one of the top 5 most serious risks affecting human health in the developing world (WHO, 2002). Zn biofortification of crops could resolve this problem (White and Broadley, 2005, 2009; Palmgren et al., 2008). Genetic biofortification has been advocated as the most sustainable and cost effective method to deliver increased Zn to the diets of all vulnerable populations globally (Bouis et al., 2003; White and Broadley, 2005, 2009; Stein et al., 2007; Palmgren et al., 2008), however knowledge of the biological processes of Zn accumulation is at an embryonic stage. The data presented in this thesis has provided further insights into possible molecular mechanisms behind this trait in a Zn hyperaccumulator and defined novel approaches to adapt this knowledge to important crop systems. A number of issues however, remain to be addressed.

In the immediate term, with regard to the Zn hyperaccumulator N. caerulescens, it will be necessary to clone the remainder of NcHMA4 tandem sequences and

perform promoter::GUS fusion studies to identify possible altered cis regulations. Heterologous expression of these sequences in A. thaliana should be examined on agar-based growth media supplemented with increasing levels of Zn ($[Zn]_{ext}$) (Chapter 5). In addition quantitative analysis of GUS activity could be assayed using the MUG technique described by Martin et al. (1992) to give a more comprehensive understanding of Zn-dependent regulation of these genes. Since N. caerulescens is recalcitrant to A. tumefaciens mediated transformation, a Viral Induced Gene Silencing (VIGs) technique (Gilchrist and Haughn, 2010) could be attempted as an initial transient approach to more clearly elucidate the function of HMA4 in planta. Zinc and other mineral concentrations in shoot and root tissues could be assayed in these plants growing on media supplemented with increasing $[Zn]_{ext}$. Truncations of HMA4 genes should be induced and expressed in planta (A. thaliana) to further dissect the roles of various domains contained in each of the tandem repeats, as was previously tested in yeast (Saccharomyces cerevisae) by Papoyan and Kochian (2004). Similarly, promoter::GUS heterologous expression of a series of truncated promoter regions of all tandem HMA4 genes (Chapter 3 and 4) should help identify key regions involved in cis regulation of HMA4. Elucidating these regulatory sequences may facilitate future marker assisted breeding strategies to increase Zn accumulation in crop species. Further modifications to the floral dipping protocol are necessary to develop stable N. caerulescens transformants and complement robust downstream molecular genetic understanding of these sequences in the endogenous genome (Chapter 5). As with transient HMA4 VIGs lines, stably transformed HMA4RNAi lines should provide clear functional evidence for the role of HMA4 in Zn accumulation. These plants could be crossed with lines harbouring RNAi of other Zn transporters and their resulting physiological processes analysed in response to altered $[Zn]_{ext}$. This should provide a clearer understanding of the interactions between HMA4 and other Zn genes such as Zn deficiency response genes (Hammond et al., 2006; Talke et al., 2006; Broadley et al., 2007; Hanikenne et al., 2008). Extending from this, RNA expression analysis of individual HMA4 tandem repeats, using highly specific Taqman probes, could corroborate any prior functional associations and give a clear understanding of associated expression levels for each repeat (Martino et al., 2010). Characterising the expression levels individual repeats present in metallicolous and non-

metallicolous populations in response to $[Zn]_{ext}$ using Taqman probes could help define the molecular control of inter-population variation in Zn accumulation (Pollard and Baker, 1996; Escarré et al., 2000; Roosens et al., 2003; Deniau et al., 2006) or variations in the accumulation of the toxic Zn analogue, cadmium (Cd) (Zha et al., 2004). Understanding the molecular mechanism restricting Cd accumulation will be crucial to safely increasing Zn in the edible portions of crop plants (Palmgren et al., 2008). Expression of 'minigene' constructs, composed of N. caerulescens HMA4 promoters fused to their corresponding HMA4 cDNA, in Zn non-hyperaccumulator plant systems will be necessary as an initial step to appreciating the effects of transferring high Zn accumulation traits to crops (Barabasz et al., 2010; Siemianowski et al., 2010). To expedite these processes, rapid cycling N. caerulescens lines, developed during this study (Chapter 5), should be employed for these molecular genetic approaches. In advance of this, backcrossing to wild type lines will be required to reduce the mutant background remaining from fast neutron mutagenesis, which could create perturbations in molecular genetic and physiological results. DNA could subsequently be hybridised to DNA Tiling arrays to help elucidate the molecular mechanisms controlling vernalisation and cycle time in Noccaea caerulescens which could ultimately have wider implications for research in other slow cycling or vernalisation dependent crop species.

In the longer term it may be possible to create more faster cycling lines by reducing the cycling time from 13 weeks to 6 weeks and thus more resemble the model plant A. thaliana. This could be achieved by crossing existing fast cycling lines such as A2 with A7 and selecting for early flowering mutants. Creating fast neutron mutagenised populations of different ecotypes of N. caerulescens which hyperaccumulate metals at different levels should permit more rapid and effective population studies to elucidate genetic aspects of hyperaccumulation. In addition EMS mutant TILLing populations of these lines, as described for B. rapa (Chapter 6), could offer a rapid and more sophisticated approach to dissecting the effects of allelic variation on Zn phenotypes. In tandem with this, a complete genomic sequence using the rapidly evolving GS FLX 454 technology (Chapter 3) should greatly accelerate our molecular understanding of these processes. Understanding these mechanisms in N. caerulescens should therefore

provide the basis for manipulating similar mechanisms in crops and help enhance their Zn content.

The genus Brassica spp. contains a wealth of genetic diversity that represents numerous globally important crops (Broadley et al., 2008). It has the advantage of being the closest crop relative to the model plant A. thaliana and model Zn hyperaccumulator N. caerulescens, and thus has a wealth of potential comparative genomic and reverse genetic tools at its disposal (Parkin et al., 2005; Stephenson et al., 2010). By applying these tools to pre-existing datasets, the current study identified five previously unreported candidate B. rapa Zn transporters, and genotyped 24 allelic variants of one using novel molecular techniques (Chapter 6). In the immediate term these CAX1 mutant lines should be backcrossed to wild type lines to reduce the mutant background from EMS treatment. Subsequently these lines could be investigated for altered Zn content on a range of growth media supplemented with increasing $[Zn]_{ext}$. This may provide an indication of allele dependent Zn phenotypes. More likely however, further in silico data searches will be required to identify, isolate and genotype lines containing null alleles for all paralogues of these genes, and through reciprocal crossing remove gene redundancy from future molecular genetic and physiological growth trials (Blanc and Wolfe, 2004; Stephenson et al., 2010). Current sequencing efforts predict the release of a fully assembled B. rapa genome sequence by the end of 2010. Identification of paralogous sequences and comparative sequence analyses should therefore become relatively routine, and thus accelerate current interpretations of classical genetic datasets (Wu et al., 2008; Broadley et al., 2010). Coupled with this, advances in expression QTL (eQTL) and more recently, the development of a 135,201 gene model Brassica Exon array (Love et al., 2010) should drastically hasten current capabilities to identify candidate sequences involved in increased $[Zn]_{ext}$. A replicated forward genetic approach is currently being tested with a subset of the entire M_3 TILLing population (3464 lines) to identify plants with altered mineral leaf content. Data from this experiment could provide an alternative approach to identifying novel genetic mechanisms controlling Zn accumulation in this crop and thus complement current reverse genetic strategies.

7.14 CONSIDERATIONS FOR FUTURE RESEARCH

A major factor to consider when enhancing Zn accumulation with altered HMA4 expression will be the dangers associated with potential increases in cadmium (Cd) levels, since Cd has been shown to be a toxic Zn analogue (Palmgren et al., 2008). It is readily acquired by roots and is more bioavailable than any other toxic heavy metal. It has been shown to cause cellular damage by inactivating or denaturing proteins (by readily binding to free sulfhydryl residues), displacing Zn cofactors from a range of proteins resulting in inactivation of transcription factors or enzymes, as well as triggering an increase in oxidation within the cell (Palmgren et al., 2008). In rice, Cd levels have been shown to increase to levels which were not phytotoxic but resulted in human toxicity and induced itai-itai disease (Palmgren et al., 2008). Researchers must therefore examine related populations of plants which differentially accumulate Cd to understand how transporters discriminate between these heavy metals.

Although increasing the Zn concentrations of crops does offer a promising avenue to sustainable Zn sufficiency in the human diet a number of other factors must be considered. Complementary research must also increase the concentration of food Zn promoters such as ascorbate, β-carotene and cysteine-rich polypeptides which stimulate the absorption of essential mineral elements by the gut. Similarly a reduction in 'antinutrient' concentrations, such as oxalate, polyphenolics or phytate, which interfere Zn absorption must be achieved (White and Broadley, 2010). It has been shown that reducing the expression of genes encoding enzymes involved in the synthesis or sequestration of these substances has successfully resulted in reduced phytate in cereal and legume crops. Overexpression of phytases has similarly reduced endogenous phytate concentrations reviewed in White and Broadley (2009).

It must also be remembered however that despite being a sustainable and robust system, breeding for increased Zn and other mineral elements has some limitations. Firstly, it is a long term approach, requiring knowledge and careful balancing of a blend of technologies including comparative genomics, transgenic data and traditional breeding strategies. Outcomes cannot be guaranteed, despite requiring a substantial initial investment. Additionally, genetic variation of Zn

accumulation must be identified between suitable parents and results can only be confirmed after years of crossing and back-crossing activities. The stability of traits such as high grain or leaf Zn concentrations must be established across a range of disparate environments with variations in both soil and climatic conditions. Genotypes producing altered Zn concentrations in edible tissue must also be compatible with agricultural practices in targeted regions (Cakmak, 2008; Palmgren et al., 2008; White and Broadley, 2005, 2009). Biofortified crops should be widely accepted by consumers in both developed and developing countries however it must be ensured that appearance taste texture and cooking quality should not be compromised (Bouis et al., 2003). It is assumed that minute alterations in micronutrient content should not significantly alter food quality, however taste and colour could be affected by these compounds (White and Broadley, 2009). To be accepted, the consumption of such products must be associated with health benefits to humans. If transgenes are to be employed, then the associated risks to wild populations and the environment with must be considered. Moreover, a balance must be struck between these perceived risks, with the economic and sustainable provision of essential nutrients to a growing global population.

7.15 CONCLUSION

This thesis described the identification and isolation of four HMA4 tandem repeats in Noccaea caerulescens Saint Laurent Le Minier. Comparative analysis of their deduced protein sequences showed a high conservation with known HMA4 metal binding motifs in published sequences from A. thaliana and A. halleri. Furthermore, altered cis regulation between both hyperaccumulators and A. thaliana suggested that these regions could play a significant role in altered HMA4 expression and associated Zn accumulation. To accelerate in planta molecular analyses of Zn accumulation, a rapid cycling population of N. caerulescens was established. The development of a robust and efficient Brassica rapa genotyping platform has provided an invaluable tool for downstream molecular genetic approaches. Together these developments should permit more efficient in planta molecular functional analyses and therefore facilitate a more rapid dissection of their role in regulating mineral accumulation in Brassicaceae. Ultimately, an enhanced appreciation of genetic variation in

$[Zn]_{shoot}$ could facilitate downstream genetic biofortification strategies in crops. It is clear that more progress is required to thoroughly exploit the genetic resources currently available and thus complement biofortification approaches. Moreover, with progress in genome sequencing, reverse genetics and comparative and transcriptome analyses gaining momentum, the future outlook appears very positive.

The research presented in this thesis has provided a novel, fast cycling Noccaea caerulescens genomic resource to the scientific community, which should considerably facilitate understanding of genetic variation in Zn accumulation in Brassicaceae. In the longer term, application of novel molecular data identified during this study should assist biofortification strategies to implement sustainable approaches to improving Zn nutrition which currently affects over 2 billion people world wide.

REFERENCES

Abdel-Ghany, S.E., Müller-Moulé, P., Niyogi, K.K., Pilon, M. and Shikanai, T. (2005) Two P-type ATPases are required for copper delivery in Arabidopsis thaliana chloroplasts. The Plant Cell **17:** 1233-1251

Agius, F., Amaya, I., Botella, M.A. and Valpuesta, V. (2005) Functional analysis of homologous and heterologous promoters in strawberry fruits using transient expression. Journal of Experimental Botany **56:** 37-46

Al-Shehbaz, I.A., Beilstein, M.A. and Kellogg, E.A. (2006) Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. Plant Systematics and Evolution **259:** 89-120

Alloway, B.J. (1995) Heavy metals in soils. (second edition). Blackie Academic & Professional, London, UK pp. 368

Alloway, B.J. (2004) Zinc in soils and crop nutrition. International Zinc Association, Brussels, Belgium pp. 123

Al-Shehbaz, I.A., Beilstein, M.A. and Kellogg, E.A. (2006) Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. Plant Systematics and Evolution **259:** 89-120

Aras, N.K., Nazli, A., Zhang, W. and Chatt, A. (2001) Dietary intake of zinc and selenium in Turkey. Journal of Radioanalytical and Nuclear Chemistry **249:** 33-37

Arrivault, S., Senger, T. and Kramer, U. (2006) The Arabidopsis metal tolerance protein AtMTP3 maintains metal homeostasis by mediating Zn exclusion from the shoot under Fe deficiency and Zn oversupply. Plant Journal **46:** 861-879

Assunção, A.G.L., Bookum, W.M., Nelissen, H.J.M., Vooijs, R., Schat, H. and Ernst, W.H.O. (2003) Differential metal-specific tolerance and accumulation patterns among Thlaspi caerulescens populations originating from different soil types. New Phytologist **159:** 411-419

Assunção, A.G.L., Herrero, E., Lin, Y.F., Huettel, B. Talukdar, S., Smaczniak, C., Immink, R.G.H., van Eldik, M., Fiers, M., Schat, H. and Aarts, M.G.M. (2010) Arabidopsis thaliana transcription factors bZIP19 and bZIP23 regulate the adaptation to zinc deficiency. Proceedings of the National Academy of Sciences of the United States of America **107:** 10296-10301

Assunção, A.G.L., Martins, P.D., De Folter, S., Vooijs, R., Schat, H. and Aarts, M.G.M. (2001) Elevated expression of metal transporter genes in three accessions of the metal hyperaccumulator Thlaspi caerulescens. Plant Cell and Environment **24:** 217-226

Assunção, A.G.L., Pieper, B., Vromans, J., Lindhout, P., Aarts, M.G.M. and Schat H. (2006) Construction of a genetic linkage map of Thlaspi caerulescens and quantitative trait loci analysis of zinc accumulation. New Phytologist **170:** 21-32

Auld, D.S. (2001) Zinc coordination sphere in biochemical zinc sites. BioMetals **14:** 271-313

Baker, A.J.M. and Brooks, R.R. (1989) Terrestrial higher plants which hyperaccumulate metallic elements. A review of their distribution, ecology and phytochemistry. Biorecovery **1:** 81-126

Baker, A.J.M., McGrath, S.P., Reeves, R.D. and Smith, J.A.C. (2000) Metal hyperaccumulator plants: a review of the ecology and physiology of a biochemical resource for phytoremediation of metal-polluted soils. In N Terry, G Bañuelos and J Vangronsveld (eds), Phytoremediation of contaminated soil and water. Lewis Publishers, Boca-Raton, FL, USA., pp 85–107

Barabasz, A., Krämer, U., Hanikenne, M., Rudzka, J. and Antosiewicz, D.M. (2010) Metal accumulation in tobacco expressing Arabidopsis halleri metal hyperaccumulation gene depends on external supply. Journal of Experimental Botany **61:** 3057-3067

Barak, P. and Helmke, P.A. (1993) The chemistry of zinc In A.D. Robson, (ed), Zinc in Soil and Plants. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 1-13

Barber, S.A. (1995) Soil Nutrient Bioavailability: A Mechanistic Approach (second edition). John Wiley & Sons, New York, NY, USA.

Barber, S.A. and Claassen, N. (1977) A mathematical model to simulate metal uptake by plants growing in soil. In H. Drucker and R.E. Wildung (eds), Proceedings of the Fifteenth Hanford Life Sciences Symposium, Oak Ridge, TN, USA, pp. 358-364

Barceloux, D.G. (1999) Zinc. Journal of Toxicology-Clinical Toxicology **37:** 279-292

Bar-Yosef, B., Fishman S. and Talpaz, H. (1980) A model of zinc movement to single roots in soils. Soil Science Society of America Journal **44:** 1272-1279

Baxter, I., Hosmani, P.S., Rus, A., Lahner, B., Borevitz, J.O., Muthukumar, B., Mickelbart, M.V., Schreiber, L., Franke, R.B. and Salt, D.E. (2009) Root suberin forms an extracellular barrier that affects water relations and mineral nutrition in Arabidopsis. PLoS Genetics **5:** pp. 12

Baxter, I., Ouzzani, M., Orcun, S., Kennedy, B., Jandhyala, S.S. and Salt, D.E. (2007) Purdue Ionomics Information Management System. An integrated functional genomics platform. Plant Physiology **143:** 600-611

Beilstein, M.A., Al-Shehbaz, I.A. and Kellogg, E.A. (2006) Brassicaceae phylogeny and trichome evolution. American Journal of Botany **93:** 607-619

Bernal, M.P., McGrath, S.P., Miller, A.J. and Baker, A.J.M. (1994) Comparison of the chemical changes in the rhizosphere of the nickel hyperaccumulator Alyssum murale with the non-accumulator Raphanus sativus. Plant and Soil **164:** 251-259

Bernard, C., Roosens, N., Czernic, P., Lebrun, M. and Verbruggen, N. (2004) A novel CPx-ATPase from the cadmium hyperaccumulator Thlaspi caerulescens. Federation of European Biochemical Societies (FEBS) Letters **569:** 140-148.

Bert, V., Bonnin, I., Saumitou-Laprade, P., de Laguérie, P. and Petit, D. (2002) Do Arabidopsis halleri from nonmetallicolous populations accumulate zinc and cadmium more effectively than those from metallicolous populations? New Phytologist **155:** 47-57

Biswas, A.K. (2006) Zinc and related alloys: The pioneering traditions in the ancient and medieval India. Transactions of the Indian Institute of Metals **59:** 911-924

Black, R.E., Allen, L.H., Bhutta, Z.A., Caulfield, L.E., de Onis, M., Ezzati, M., Mathers, C. and Rivera, J. (2008) Maternal and child undernutrition 1 - Maternal and child undernutrition: global and regional exposures and health consequences. LANCET , **371**: 243 – 260.

Blanc, G. and Wolfe, K.H. (2004) Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. The Plant Cell **16:** 1667-1678.

Blaudez, D., Kohler, A., Martin, F., Sanders, D. and Chalot, M. (2003) Poplar Metal Tolerance Protein 1 confers zinc tolerance and is an oligomeric vacuolar zinc transporter with an essential Leucine Zipper motif. The Plant Cell **15:** 2911-2928

Bouis, H.E., Chassy, B.M. and Ochanda, J.O. (2003) Genetically modified food crops and their contribution to human nutrition and food quality. Trends in Food Science & Technology **14:** 191-209

Bowen, J.E. (1973) Kinetics of zinc absorption by excised roots of two sugarcane clones. Plant and Soil **39:** 125-129

BPMP (2007) Adaptation des plantes aux métaux, Institut de Biologie Intégrative des Plantes. www.bpmp.cnrs.fr/Groupes/Metaux.htm (accessed 24 March 2007).

Broadley, M.R., Hammond, J.P., King, G.J., Astley, D., Bowen, H.C., Meacham, M.C., Mead, A., Pink, D.A.C., Teakle, G.R., Hayden, R.M., Spracklen, W.P. and White P.J. (2008) Shoot calcium (Ca) and magnesium (Mg) concentrations differ between subtaxa, are highly heritable, and associate with potentially pleiotropic loci in Brassica oleracea. Plant Physiology, **146**: 1707-1720.

Broadley, M.R., Ó Lochlainn, S., Hammond, J.P., Bowen, H.C., Cakmak, I., Eker, S., Erdem, H., King, G.J. and White, P.J. (2010) Shoot zinc (Zn) concentration varies widely within Brassica oleracea L. and is affected by soil Zn and phosphorus (P) levels. Journal of Horticultural Science and Biotechnology **85:** 375-380

Broadley, M.R., White, P.J., Bryson, R.J., Meacham, M.C., Bowen, H.C., Johnson, S.E., Hawkesford, M.J., McGrath, S.P., Zhao, F.J., Breward, N., Harriman, M. and Tucker, M. (2006) Biofortification of UK food crops with selenium. Proceedings of the Nutrition Society **65:** 169-181

Broadley, M.R., White, P.J., Hammond, J.P., Zelko, I. and Lux, A. (2007) Zinc in plants. New Phytologist **173:** 677-702

Brooks, R.R., Lee, J., Reeves, R.D. and Jaffre, T. (1977) Detection of nickeliferous rocks by analysis of herbarium specimens of indicator plants. Journal of Geochemical Exploration **7:** 49-57

Brown, P.H., Cakmak, I. and Zhang, Q. (1993) Form and function of zinc in plants. In A.D. Robson, (ed), Zinc in Soils and Plants. Kluwer Academic Publishers, Dortrecht, pp 90 - 106

Bruce, M., Hess, A., Bai, J.F. Mauleon, R., Diaz, M.G., Sugiyama, N., Bordeos, A., Wang, G.L., Leung, H. and Leach, J.E. (2009) Detection of genomic deletions in rice using oligonucleotide microarrays. BioMed Central Genomics **10:** pp. 11

Cakmak, I. (2002) Plant nutrition research: Priorities to meet human needs for food in sustainable ways. Plant and Soil **247:** 3-24

Cakmak, I. (2004) Identification and correction of widespread zinc deficiency in Turkey – a success story (a NATO-Science for Stability Project). Proceedings of the International Fertiliser Society **552:** 1-26

Cakmak, I. (2008) Enrichment of cereal grains with zinc: Agronomic or genetic biofortification? Plant and Soil **302:** 1-17

Cakmak, I. and Marschner, H. (1987) Mechanism of phosphorus-induced zinc deficiency in Cotton. III. Changes in physiological availability of zinc in plants. Physiologia Plantarum **70:** 13-20

Cakmak, I. and Marschner, H. (1988) Enhanced superoxide radical production in roots of zinc-deficient plants. Journal of Experimental Botany **39:** 1449-1460

Cakmak, I. and Marschner, H. (1988) Increase in membrane permeability and exudation in roots of zinc-deficient plants. Journal of Plant Physiology **132:** 356-361

Cakmak, I. and Marschner, H. (1988) Zinc-dependent changes in Electron-Spin-Resonance signals, NADPH Oxidase and plasma membrane permeability in cotton roots. Physiologia Plantarum **73:** 182-186

Cambia (2007)
http://www.cambia.org/daisy/cambia/materials/vectors/1204/1201.html (Accessed 24 April 2007).

Cappuyns, V., Swennen, R., Vandamme, A. and Niclaes, M. (2006) Environmental impact of the former Pb-Zn mining and smelting in East Belgium. Journal of Geochemical Exploration **88:** 6-9

Carr, G. (2004) Brassicaceae: Cruciferae, University of Hawaii at Manoa. http://www.botany.hawaii.edu/faculty/carr/brassic.htm (Accessed, 23 March 2006).

Castiglione, S., Franchin, C., Fossati, T., Lingua, G., Torrigiani, P. and Biondi, S. (2007) High zinc concentrations reduce rooting capacity and alter metallothionein gene expression in white poplar (Populus alba L. cv. Villafranca). Chemosphere **67:** 1117-1126

Cavell, A.C., Lydiate, D.J., Parkin, I.A.P., Dean, C. and Trick, M. (1998) Collinearity between a 30-centimorgan segment of Arabidopsis thaliana chromosome 4 and duplicated regions within the Brassica napus genome. Genome **41:** 62-69

Chandler, W.H., Hoagland, D.R. and Hibbard, P.L. (1931) Little leaf rosette in fruit trees. Proceedings of the American Horticultural Society **28:** 556 - 561

Chaney, R.L. (1983) Plant uptake of inorganic waste constituents. In J.F. Parr, P.B. Marsh and J.M. Kla (eds), Land Treatment of Hazardous Wastes. Noyes Data Corporation, Park Ridge, New Jersey pp 50–76

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

Choi, S.R., Teakle, G.R., Plaha, P., Kim, J.H., Allender, C.J., Beynon, E., Piao, Z.Y., Soengas, P., Han, T.H., King, G.J., Barker, G.C., Hand, P., Lydiate, D.J., Batley, J., Edwards, D., Koo, D.H., Bang, J.W., Park, B.-S. and Lim, Y.P. (2007) The reference genetic linkage map for the multinational Brassica rapa genome sequencing project. Theoretical and Applied Genetics **115:** 777 - 792

Clark, R. M., Wagler, T. N., Quijada, P. and Doebley, J. (2006) A distant upstream enhancer at the maize domestication gene tb1 has pleiotropic effects on plant and inflorescent architecture. Nature Genetics **38:** 594-597

Clough, S.J. and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant Journal **16:** 735-743

Colangelo, E.P. and Guerinot, M.L. (2006) Put the metal to the petal: metal uptake and transport throughout plants. Current Opinion in Plant Biology **9:** 322-330

Cohen D. (2007) Earth's natural wealth: an audit. New Scientist **2605:** 34-41

Colbert, T., Till, B.J., Tompa, R., Reynolds, S., Steine, M.N., Yeung, A.T., McCallum, C.M., Comai, L. and Henikoff, S. (2001) High-throughput screening for induced point mutations. Plant Physiology **126:** 480-484

Craddock, P.T. (1987) The early history of zinc. Endeavour **11:** 183-191

Curtis, I.S., Power, J.B., Blackhall, N.W., Delaat, A.M.M. and Davey, M.R. (1994) Genotype-independent transformation of lettuce using Agrobacterium tumefaciens. Journal of Experimental Botany **45:** 1441-1449

Dardenne, M. (2002) Zinc and immune function. European Journal of Clinical Nutrition **56:** S20-S23

Davis, M.A., Murphy, J.F. and Boyd, R.S. (2001) Nickel increases susceptibility of a nickel hyperaccumulator to Turnip mosaic virus. Journal of Environmental Quality **30:** 85-90

Dechamps, C., Lefebvre, C., Noret, N. and Meerts, P. (2007) Reaction norms of life history traits in response to zinc in Thlaspi caerulescens from metalliferous and nonmetalliferous sites. New Phytologist **173:** 191-198

Dechamps, C., Roosens, N.H., Hotte, C. and Meerts, P. (2005) Growth and mineral element composition in two ecotypes of Thlaspi caerulescens on Cd contaminated soil. Plant and Soil **273:** 327-335

Deniau, A.X., Pieper, B., Ten Bookum, W.M., Lindhout, P., Aarts, M.G.M. and Schat, H. (2006) QTL analysis of cadmium and zinc accumulation in the heavy metal hyperaccumulator Thlaspi caerulescens. Theoretical and Applied Genetics **113:** 907-920

Desfeux, C., Clough, S.J. and Bent, A.F. (2000) Female reproductive tissues are the primary target of Agrobacterium-mediated transformation by the Arabidopsis floral-dip method. Plant Physiology **123:** 895-904

Dewey, K.G. and Brown, K.H. (2003) Update on technical issues concerning complementary feeding of young children in developing countries and implications for intervention programs. (Special Issue Based on a World Health Organization Expert Consultation on Complementary Feeding). Food and Nutrition Bulletin **24:** 5 - 28

Dijkhuizen, M.A., Winichagoon, P., Wieringa, F.T., Wasantwisut, E., Utomo, B., Ninh, N.X., Hidayat, A. and Berger, J. (2008) Zinc supplementation improved length growth only in anemic infants in a multi-country trial of iron and zinc supplementation in South-East Asia. Journal of Nutrition **138:** 1969 – 1975

Dong, C., Vincent, K. and Sharp, P. (2009) Simultaneous mutation detection of three homoeologous genes in wheat by High Resolution Melting analysis and Mutation Surveyor®. BioMed Central Plant Biology **9:143** pp. 12

Drake, P.M.W., John, A., Power, J.B. and Davey, M.R. (1997) Expression of the gus A gene in embryogenic cell lines of Sitka spruce following Agrobacterium-mediated transformation. Journal of Experimental Botany **48:** 151-155

Englbrecht, C.C., Schoof, H. and Böhm, S. (2004) Conservation, diversification and expansion of C2H2 zinc finger proteins in the Arabidopsis thaliana genome. BioMed Central Genomics **5:39**

Epicentre Biotechnologies (2010) *CopyControl™ Fosmid Library Production* Kit.. http://www.epibio.com/item.asp?ID=385 (Accessed 19 July 2010)

Eriksson, T. (1965) Studies on the growth requirements and growth measurements of cell cultures of Haplopappus gracilis. Physiologia Plantarium **18:** 976-993

Ernst, W.H.O., Schat, H. and Verkleij, J.A.C. (1990) Evolutionary biology of metal resistance in Silene vulgaris. Evolutionary Trends in Plants **4:** 45-51

Escarré, J., Lefèbvre, C., Gruber, W., Leblanc, M., Lepart, J., Riviere, Y. and Delay, B. (2000) Zinc and cadmium hyperaccumulation by Thlaspi caerulescens from metalliferous and nonmetalliferous sites in the

Mediterranean area: implications for phytoremediation. New Phytologist **145:** 429-437

FAO (2005) Core consumption data: FAOSTAT. Food and Agriculture Organization of the United Nations. http://faostat.fao.org/site/345/default.aspx (Accessed 23 March 2007).

Filatov, V., Dowdle, J., Smirnoff, N., Ford-Lloyd, B., Newbury, H.J. and Macnair, M.R. (2006) Comparison of gene expression in segregating families identifies genes and genomic regions involved in a novel adaptation, zinc hyperaccumulation. Molecular Ecology **15:** 3045-3059

Fior, S., Vianelli, A. and Gerola, P.D. (2009) A novel method for fluorometric continuous measurement of β-glucuronidase (GUS) activity using 4-methyl-umbelliferyl-β-D-glucuronide (MUG) as substrate. Plant Science **176:** 130-135

Fischer, E.H. and Davie, E.W. (1998) Recent excitement regarding metallothionein. Proceedings of the National Academy of Sciences of the United States of America **95:** 3333-3334

Frérot, H., Lefèbvre, C., Petit, C., Collin, C., Dos Santos, A. and Escarré, J. (2005) Zinc tolerance and hyperaccumulation in F_1 and F_2 offspring from intra and interecotype crosses of Thlaspi caerulescens. New Phytologist **165:** 111-119

Frey, B., Keller, C., Zierold, K. and Schulin, R. (2000) Distribution of Zn in functionally different leaf epidermal cells of the hyperaccumulator Thlaspi caerulescens. Plant, Cell and Environment **23:** 675-687

Friedland, A.J. (1990) The movement of metals through soils and ecosystems. In A.J. Shaw (ed), Heavy Metal Tolerance in Plants: Evolutionary Aspects. CRC Press, Boca Raton, FL, USA, pp 7-19

Frossard, E., Bucher, M., Mächler, F., Mozafar, A. and Hurrell, R. (2000) Potential for increasing the content and bioavailability of Fe, Zn and Ca in plants for human nutrition. Journal of the Science of Food and Agriculture **80:** 861-879

Gangopadhyay, G., Roy, S.K. and Mukherjee, K.K. (2009) Plant response to alternative matrices for in vitro root induction. African Journal of Biotechnology **8:** 2923 – 2928

Garritano, S., Gemignani, F., Voegele, C., Nguyen-Dumont, T., Le Calvez-Kelm, F., De Silva, D., Lesueur, F., Landi, S. and Tavtigian, S.V. (2009) Determining the effectiveness of High Resolution Melting analysis for SNP genotyping and mutation scanning at the TP53 locus. BioMed Central Genetics **10:5** pp. 12

GCDT (2006) Brassica. Global Crop Diversity Trust, Rome, Italy. http://www.croptrust.org/main/priority.php?itemid=155 (Accessed 13 February 2007).

George, E.F. and de Klerk, G-J. (2008) The components of plant tissue culture media 1: Macro- and micro-nutrients. In E.F. George, M.A. Hall and G.J. de Klerk (eds), Plant Propagation by Tissue Culture (Third edition). Springer, Dortrecht, The Netherlands, p. 65 – 114

Gibson, R.S. and Ferguson, E.L. (1998) Nutrition intervention strategies to combat zinc deficiency in developing countries. Nutrition Research Reviews **11:** 115-131

Gilchrist, E. and Haughn, G. (2010) Reverse genetics techniques: engineering loss and gain of gene function in plants. Briefing in Functional Genomics **9:** 103-110

Gizard, F., Lavallée, B., DeWitte, F. and Hum, D.W. (2001) A novel zinc finger protein TReP-132 interacts with CBP/p300 to regulate human CYP11A1 gene expression. Journal of Biological Chemistry **276:** 33881-33892

Goto, F., Yoshihara, T., Shigemoto, N., Toki, S. and Takaiwa, F. (1999) Iron fortification of rice seed by the soybean ferritin gene. Nature Biotechnology **17:** 282-286

Graham, R.D., Ascher, J.S. and Hynes, S.C. (1992) Selecting zinc-efficient cereal genotypes for soils of low zinc status. Plant and Soil **146:** 241-250

Graham, R.D., Welch, R.M. and Bouis, H.E. (2001) Addressing micronutrient malnutrition through enhancing the nutritional quality of staple foods: principles, perspectives and knowledge gaps. Advances in Agronomy **70:** 77-142

Grass, G., Fan, B., Rosen, B. P., Franke, S., Nies, D. H. and Rensing, C. (2001) ZitB (YbgR), a member of the cation diffusion facilitator family, is an additional zinc transporter in Escherichia coli. Journal of Bacteriology **183:** 4664 – 4667

Gravot, A., Lieutaud, A., Verret, F., Auroy, P., Vavasseur, A. and Richaud, P. (2004) AtHMA3, a plant P1B-ATPase, functions as a Cd/Pb transporter in yeast. Federation of European Biochemical Societies Letters **561:** 22-28

Greene, E.A., Codomo, C.A., Taylor, N.E., Henikoff, J.G., Till, B.J., Reynolds, S.H., Enns, L.C., Burtner, C., Johnson, J.E., Odden, A.R., Comai, L. and Henikoff, S. (2003) Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. Genetics **164:** 731-740

Grusak, M.A. and Cakmak, I. (2005) Methods to improve the crop-delivery of minerals to humans and livestock. In: M.R. Broadley and P.J. White (eds), Plant nutritional genomics. Oxford, UK: Blackwell, pp 265-286

Guan Z.Q., Chai, T.Y., Zhang, Y.X., Xu, J., Wei, W., Han, L. and Cong L. (2008) Gene manipulation of a heavy metal hyperaccumulator species Thlaspi caerulescens L. via Agrobacterium-mediated transformation. Molecular Biotechnology **40:** 77-86.

Guerinot, M.L. and Salt, D.E. (2001) Fortified foods and phytoremediation. Two sides of the same coin. Plant Physiology **125:** 164-167

Hacisalihoglu, G. and Kochian. L.V. (2003) How do some plants tolerate low levels of soil zinc? Mechanisms of zinc efficiency in crop plants. New Phytologist **159:** 341-350

Hacisalihoglu, G., Hart, J.J. and Kochian, L.V. (2001) High- and low-affinity zinc transport systems and their possible role in zinc efficiency in bread wheat. Plant Physiology **125:** 456-463

Haines, B.J. (2002) Zincophilic root foraging in Thlaspi caerulescens. New Phytologist **155:** 363-372

Hall, J.L. and Williams, L.E. (2003) Transition metal transporters in plants. Journal of Experimental Botany **54:** 2601-2613

Hall, R.D. (1999) An introduction to plant cell culture: Pointers to success. In R.D. Hall (ed), Methods in Molecular Biology: Plant Cell Culture Protocols. Humana Press, New Jersey, USA, pp 1-18

Hamlin, R.L., Schatz, C. and Barker, A.V. (2003) Zinc accumulation in Indian mustard as influenced by nitrogen and phosphorus nutrition. Journal of Plant Nutrition **26:** 177-190

Hammer, D. and Keller, C. (2003) Phytoextraction of Cd and Zn with Thlaspi caerulescens in field trials. Soil Use and Management **19:** 144-149

Hammond, J.P., Bowen, H.C., White, P.J., Mills, V., Pyke, K.A., Baker, A.J.M., Whiting, S.N., May, S.T. and Broadley M.R. (2006) A comparison of the Thlaspi caerulescens and Thlaspi arvense shoot transcriptomes. New Phytologist **170:** 239-260

Hanikenne, M., Talke, I.N., Haydon, M.J., Lanz, C., Nolte, A., Motte, P., Kroymann, J., Weigel, D. and Krämer, U. (2008) Evolution of metal hyperaccumulation required cis-regulatory changes and triplication of HMA4. Nature **453**: 391-395.

Hart, J.J., Norvell, W.A., Welch, R.M., Sullivan, L.A. and Kochian, L.V. (1998) Characterization of zinc uptake, binding, and translocation in intact seedlings of bread and durum wheat cultivars. Plant Physiology **118:** 219-226

Hart, J.J., Welch, R.M., Norvell, W.A. and Kochian, L.V. (2002) Transport interactions between cadmium and zinc in roots of bread and durum wheat seedlings. Physiologia Plantarum **116:** 73-78

Haydon, M.J. and Cobbett, C.S. (2007) A novel major facilitator superfamily protein at the tonoplast influences zinc tolerance and accumulation in Arabidopsis. Plant Physiology **143:** 1705-1719

Hellens, R., Mullineaux, P. and Klee, H. (2000) Technical focus: A guide to Agrobacterium binary Ti vectors. Trends in Plant Science **5:** 446-451

Henderson, W.E. and Kinnersley, A.M. (1988) Corn starch as an alternative gelling agent for plant tissue culture. Plant Cell, Tissue and Organ Culture **15:** 17-22

Hoagland, D.R. (1948) Lectures on the inorganic nutrition of plants (second edition). Chronica Botanica Company, Waltham, MA, USA

Hossain, B., Hirata, N., Nagatomo, Y., Akashi, R. and Takaki, H. (1997) Internal zinc accumulation is correlated with increased growth in rice suspension culture. Journal of Plant Growth Regulation **16:** 239 - 243

Hotz, C. and Brown, K.H. (2004) International Zinc Nutrition Consultative Group (IZiNCG) technical document #1. Assessment of the risk of zinc deficiency in populations and options for its control. Food and Nutrition Bulletin **25:** S94-S203

Hussain, D., Haydon, M.J., Wang, Y., Wong, E., Sherson, S.M., Young, J., Camakaris, J., Harper, J.F. and Cobbett, C.S. (2004) P-type ATPase heavy metal transporters with roles in essential zinc homeostasis in Arabidopsis. The Plant Cell **16:** 1327-1339

Ibrikci, H., Knewtson, S.J.B. and Grusak, M.A. (2003) Chickpea leaves as a vegetable green for humans: evaluation of mineral composition. Journal of the Science of Food and Agriculture **83:** 945-950

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:** 1159–1165.

Inan, G., Zhang, Q., Li, P.H., Wang, Z.L., Cao, Z.Y., Zhang, H., Zhang, C.Q., Quist, T.M., Goodwin, S.M., Zhu, J.H., Shi, H.Z., Damsz, B., Charbaji, T., Gong, Q.Q., Ma, S.S., Fredricksen, M., Galbraith, D.W., Jenks, M.A., Rhodes, D., Hasegawa, P.M., Bohnert, H.J., Joly, R.J., Bressan, R.A. and Zhu, J.K. (2004) Salt cress. A halophyte and cryophyte Arabidopsis relative model system and its applicability to molecular genetic analyses of growth and development of extremophiles. Plant Physiology **135:** 1718-1737

Iniguez-luy, F.L., Lukens, L., Farnham, M.W., Amasino, R.M. and Osborn, T.C. (2009) Development of public immortal mapping populations, molecular markers and linkage maps for rapid cycling Brassica rapa and B. oleracea. Theoretical and Applied Genetics **120:** 31-43

Islam, M.R., Rodova, M. and Calvet, J.P. (2002) A fast and efficient method of DNA fragment isolation from agarose gels without using commercial kits. American Biotechnology Laboratory **20:** 18

IZiNCG (2006) Advocacy Statement: Did you know that zinc can save children's lives? International Zinc Nutrition Consultative Group, Davis, CA, USA, pp 4.

Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. (1987) GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. The EMBO Journal **6:** 3901-3907

Jimenez-Ambriz, G., Petit, C., Bourrie, I., Dubois, S., Olivieri, I. and Ronce, O. (2007) Life history variation in the heavy metal tolerant plant Thlaspi caerulescens growing in a network of contaminated and noncontaminated sites in southern France: role of gene flow, selection and phenotypic plasticity. New Phytologist, **173:** 199-215

Jin, S.G., Komari, T., Gordon, M.P. and Nester, E.W. (1987) Genes responsible for the supervirulence phenotype of Agrobacterium Tumefaciens A281. Journal of Bacteriology **169:** 4417-4425

Johns, T. and Eyzaguirre, P.B. (2007) Biofortification, biodiversity and diet: A search for complementary applications against poverty and malnutrition. Food Policy **32:** 1-24

Jones, G., Steketee, R.W., Black, R.E., Bhutta, Z.A. and Morris, S.S. (2003) How many child deaths can we prevent this year? Lancet **362:** 65-71

Karimi, M., Inzé, D. and Depicker, A. (2002) GATEWAY™ vectors for Agrobacterium-mediated plant transformation. Trends in Plant Science **7:** 193-195.

Katayama, A., Tsujii, A., Wada, A., Nishino, T. and Ishihama, A. (2002) Systematic search for zinc-binding proteins in Escherichia coli. European Journal of Biochemistry **269:** 2403-2413

Kesler, S.E. (2007) Mineral supply and demand into the 21st century. In J.A. Briskey and K.J. Schulz (eds), U.S. Geological Survey circular 1294: proceedings for a workshop on deposit modeling, mineral resource assessment, and their role in sustainable development. Reston, VA, USA: U.S. Geological Survey, 55-62

Khan, H.R., McDonald, G.K. and Rengel, Z. (2003) Zn fertilization improves water use efficiency, grain yield and seed Zn content in chickpea. Plant and Soil **249:** 389-400

Kim, J.S., Chung, T.Y., King, G.J., Jin, M., Yang, T.J., Jin Y.M., Kim, H.-I. and Park, B.S. (2006) A sequence-tagged linkage map of Brassica rapa. Genetics **174:** 29-39

Kim, U.-J., Shizuya, H., de Jong, P.J., Birren, B., Simon, M.I. (1992) Stable propagation of cosmid sized human DNA inserts in an F factor based vector. Nucleic Acids Research **20:**1083–1085

King, G.J. (2005) Brassica: Harvesting the genome, diversity and products. In G.J. King (ed), The Multinational Brassica Genome Project. BBSRC, Rothamstead Research Centre, Norwich, UK **1.10**: 1-18

Klug, A. (1999) Zinc finger peptides for the regulation of gene expression. Journal of Molecular Biology **293:** 215-218

Knight, B., Zhao, F.J., McGrath, S.P. and Shen, Z.G. (1997) Zinc and cadmium uptake by the hyperaccumulator Thlaspi caerulescens in contaminated soils and its effects on the concentration and chemical speciation of metals in soil solution. Plant and Soil **197:** 71-78

Koch, M. and Al-Shehbaz, I.A. (2004) Taxonomic and phylogenetic evaluation of the American "Thlaspi" species: identity and relationship to the Eurasian genus Noccaea (Brassicaceae). Systematic Botany **29:** 375-384

Koch, M. and Mummenhoff, K. (2001) Thlaspi s.str. (Brassicaceae) versus Thlaspi s.l.: morphological and anatomical characters in the light of its nrDNA sequence data. Plant Systematics and Evolution **227:** 209-225

Koleli, N., Eker, S. and Cakmak, I. (2004) Effect of zinc fertilization on cadmium toxicity in durum and bread wheat grown in zinc-deficient soil. Environmental Pollution **131:** 453-459

Koncz, C. and Schell, J. (1986) The promoter of the T_L-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of Agrobacterium binary vector. Molecular and General Genetics **204:** 383–396.

Koornneef, M., Dellaert, L.W.M. and Vanderveen, J.H. (1982) EMS-induced and radiation-induced mutation frequencies at individual loci in Arabidopsis thaliana (L.) Heynh. Mutation Research **93:** 109-123

Kopsell, D.A., Kopsell, D.E., Lefsrud, M.G., Curran-Celentano, J. and Dukach, L.E. (2004) Variation in lutein, β-carotene, and chlorophyll concentrations among Brassica oleracea cultigens and seasons. HortScience **39:** 361–364

Kopsell, D.E., Kopsell, D.A., Lefsrud, M.G. and Curran-Celentano, J. (2004). Variability in elemental accumulations among leafy Brassica oleracea cultivars and selections. Journal of Plant Nutrition, **27:** 1813-1826

Krämer, U. (2010) Metal hyperaccumulation in plants. Annual Review of Plant Biology **61:** 517-534

Küpper, H., Lombi, E., Zhao, F.-J. and McGrath, S.P. (2000) Cellular compartmentation of cadmium and zinc in relation to other elements in the hyperaccumulator. Planta **212:** 75-84

Küpper, H., Zhao, F.J. and McGrath, S.P. (1999) Cellular compartmentation of zinc in leaves of the hyperaccumulator Thlaspi caerulescens. Plant Physiology **119:** 305-311

Labana, K.S. and Gupta, M.L. (1993) Importance and origin. In K.S. Labana, S.S. Banga and S.K. Banga, (eds), Breeding Oilseed Brassicas. Springer-Verlag Berlin, Germany pp 1–20

Lagercrantz, U. and Lydiate, D.J. (1996) Comparative genome mapping in Brassica. Genetics **144:** 1903 – 1910.

Lagercrantz, U., Putterill, J., Coupland, G. and Lydiate, D. (1996) Comparative mapping in Arabidopsis and Brassica, fine scale genome collinearity and congruence of genes controlling flowering time. Plant Journal **9:** 13-20

LANL (2007) Zinc. Los Alamos National Laboritory, University of California, Los Alamos. http://periodic.lanl.gov/elements/30.html (Accessed, 12 January, 2007)

Lasat, M.M. and Kochian, L.V. (2000) Physiology of Zn hyperaccumulation in Thlaspi caerulescens. In N. Terry and G. Bañuelos, (eds), Phytoremediation of contaminated soil and water. CRC Press LLC, Boca Raton, FL, USA pp 159-169

Lasat, M.M., Baker, A.J.M. and Kochian, L.V. (1996) Physiological characterization of root Zn^{2+} absorption and translocation to shoots in Zn hyperaccumulator and nonaccumulator species of Thlaspi. Plant Physiology **112:** 1715-1722

Lehto, N.J., Davison, W., Zhang, H. and Tych, W. (2006) Analysis of micro-nutrient behaviour in the rhizosphere using a DGT parameterised dynamic plant uptake model. Plant and Soil **282:** 227-238

Li, G., Gao, M., Yang, B. and Quiros, C.F. (2003) Gene for gene alignment between the Brassica and Arabidopsis genomes by direct transcriptome mapping. Theoretical and Applied Genetics **107:** 168-180

Lindsay, W.L. (1979) Chemical Equilibria in Soils. John Wiley & Sons, Inc., New York, USA

Liu, M.Q., Yanai, J., Jiang, R.F., Zhang, F., McGrath, S.P. and Zhao, F.J. (2008) Does cadmium play a physiological role in the hyperaccumulator Thlaspi caerulescens? Chemosphere **71:** 1276-1283

Lombi, E., Zhao, F.J., Dunham, S.J. and McGrath, S.P. (2000) Cadmium accumulation in populations of Thlaspi caerulescens and Thlaspi goesingense. New Phytologist **145:** 11-20

Lombi, E., Zhao, F.J., McGrath, S.P., Young, S.D. and Sacchi, G.A. (2001) Physiological evidence for a high-affinity cadmium transporter highly expressed in a Thlaspi caerulescens ecotype. New Phytologist **149:** 53-60

Love, C.G., Graham, N.S., Ó Lochlainn, S., Bowen, H.C., May, S.T., White, P.J., Broadley, M.R., Hammond, J.P. and King, G.J. (2010) A Brassica exon array for whole-transcript gene expression profiling. PLoS ONE **5:** e12812. doi:10.1371/journal.pone.0012812

Lukens, L., Zou, F., Lydiate, D., Parkin, I. and Osborn, T. (2003) Comparison of a Brassica oleracea genetic map with the genome of Arabidopsis thaliana. Genetics **164:** 359-372

Lundin, S., Stranneheim, H., Pettersson, E., Klevebring, D. and Lundeberg, J. (2010) Increased throughput by parallelization of library preparation for massive sequencing. PLoS One **5:** 1-7

Luo, Z.B., He, X.J., Chen, L., Tang, L., Gao, S. and Chen F (2010) Effects of zinc on growth and antioxidant responses in Jatropha curcas seedlings. International Journal of Agriculture and Biology **12:** 119-124

Ma, J.F., Ueno, D., Zhao F.-J. and McGrath, S.P. (2005) Subcellular localisation of Cd and Zn in the leaves of a Cd-hyperaccumulating ecotype of Thlaspi caerulescens. Planta **220:** 731-736

MacLachlan, J. (1985) Macroalgae (Seaweeds): Industrial sources and their utilizations. Plant Soil **89:** 137-157

Macnair, M.R. (2003) The hyperaccumulation of metals by plants. Advances in Botanical Research **40:** 63–105

Maret, W. (2005) Zinc coordination environments in proteins determine zinc functions. Journal of Trace Elements in Medicine and Biology **19:** 7-12

Marquès, L., Cossegal, M., Bodin, S., Czernic, P. and Lebrun, M. (2004) Heavy metal specificity of cellular tolerance in two hyperaccumulating plants, Arabidopsis halleri and Thlaspi caerulescens. New Phytologist **164:** 289-295

Marschner, H. (1995) Mineral Nutrition of Higher Plants (second edition). Academic Press, London

Marschner, H., Romheld, V. and Cakmak, I. (1987) Root-induced changes of nutrient availability in the rhizosphere. Journal of Plant Nutrition **10:** 1175-1184

Martin, T., Wöhner, R.-V., Hummel, S., Willmitzer, L. and Frommer, W.B. (1992) The GUS reporter system as a tool to study plant gene expression. In S.R. Gallagher, (ed), GUS Protocols: Using the GUS Gene as a Reporter of Gene Expression. Academic Press, San Diego, USA, pp 23-43

Martinez-Navarrete, N., Camacho, M.M., Martinez-Lahuerta, J., Martinez-Monzo, J. and Fito, P (2002) Iron deficiency and iron fortified foods - a review. Food Research International **35:** 225-231

Martinez-Trujillo, M., Limones-Briones, V., Cabrera-Ponce, J.L., and Herrera-Estrella, L. (2004) Improving transformation efficiency of Arabidopsis thaliana by modifying the floral dip method. Plant Molecular Biology Reporter **22:** 63-70

Martino, A., Mancuso, T. and Rossi, A.M. (2010) Application of high-resolution melting to large-scale, high-throughput SNP genotyping: A comparison with the TaqMan® method. Journal of Biomolecular Screening **15:** 623-629

Mason, P. (2006) Physiological and medicinal zinc. Pharmaceutical Journal **276:** 271-274

Mazé, P. (1915) The determination of rare mineral elements necessary in the development of maize. Comptes Rendus Hebdomadaires Des Seances De L'Academie Des Sciences **160:** 211-214

McCallum, C.M., Comai, L., Greene, E.A. and Henikoff, S. (2000) Targeted screening for induced mutations. Nature Biotechnology **18:** 455-457

McCallum, C.M., Comai, L., Greene, E.A. and Henikoff, S. (2000) Targeting induced local lesions in genomes (TILLING) for plant functional genomics. Plant Physiology **123:** 439-442

McGrath, S.P., Shen, Z.G. and Zhao, F.J. (1997) Heavy metal uptake and chemical changes in the rhizosphere of Thlaspi caerulescens and Thlaspi ochroleucum grown in contaminated soils. Plant and Soil **188:** 153-159

Meerts, P. and Grommesch, C. (2001) Soil seed banks in a heavy-metal polluted grassland at Prayon (Belgium). Plant Ecology **155:** 35-45

Meerts, P., Duchêne, P., Gruber, W. and Lefèbvre, C. (2003) Metal accumulation and competitive ability in metallicolous and non-metallicolous Thlaspi caerulescens fed with different Zn salts. Plant and Soil **249:** 1-8

Meyer, M., Stenzel, U. and Hofreiter, M. (2008) Parallel tagged sequencing on the 454 platform. Nature Protocols **3(2):** 267-278

Michaels, S.D. and Amasino, R.M. (1999) FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. The Plant Cell **11:** 949-956

Mills, R.F., Francini, A., Ferreira da Rocha, P.S.C., Baccarini, P.J., Aylett, M., Krijger, G.C. and Williams, L.E. (2005) The plant P1B-type ATPase

AtHMA4 transports Zn and Cd and plays a role in detoxification of transition metals supplied at elevated levels. Federation of European Biochemical Societies (FEBS) Letters **579:** 783-791

Mills, R.F., Krijger, G.C., Baccarini, P.J., Hall, J.L. and Williams, L.E. (2003) Functional expression of AtHMA4, a P-1B-type ATPase of the Zn/Co/Cd/Pb subclass. Plant Journal **35:** 164–176.

Milner, M.J. and Kochian, L.V. (2008) Investigating heavy-metal hyperaccumulation using Thlaspi caerulescens as a model system. Annals of Botany **102:** 3-13

Mirouze, M., Sels, J., Richard, O., Czernic, P., Loubet, S., Jacquier, A., Francois I., Cammue, B.P.A., Lebrun, M., Berthomieu, P. and Marquès, L. (2006) A putative novel role for plant defensins: a defensin from the zinc hyper-accumulating plant, Arabidopsis halleri, confers zinc tolerance. The Plant Journal **47:** 329-342

Mita, S., Suzuki-Fujii, K. and Nakamura, K. (1995) Sugar-inducible expression of a gene for [beta]-amylase in Arabidopsis thaliana. Plant Physiology **107:** 895-904

Monis, P.T., Giglio, S., and Saint, C.P. (2005) Comparison of SYT09 and SYBR Green I for real-time polymerase chain reaction and investigation of the effect of the dye concentration on amplification and DNA melting curve analysis. Analytical Biochemistry **340:** 24-34

Moreno, D.A., López-Berenguer, C. and García-Viguera, C. (2007) Effects of stir-fry cooking with different edible oils on the phytochemical composition of broccoli. Journal of Food Science **72:** S64-S68

Mummenhoff, K., Franzke, A. and Koch, M. (1997) Molecular data reveal convergence in fruit characters used in the classification of Thlaspi s.l. (Brassicaceae). Botanical Journal of the Linnean Society **125:** 183-199

Nakagawa, T., Kurose, T., Hino, T., Tanaka, K., Kawamukai, M., Niwa, Y., Toyooka, K., Matsuoka, K., Jinbo, T. and Kimura, T. (2007) Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. Journal of Bioscience and Bioengineering **104:** 34-41.

Nambiar, E.K.S. (1976) Uptake of Zn-65 by oats in relation to soil-water content and root growth. Australian Journal of Soil Research **14:** 67-74

Nestel, P., Bouis, H.E., Meenakshi, J.V. and Pfeiffer, W. (2006) Biofortification of staple food crops. Journal of Nutrition **136:** 1064-1067

Neue, H.U., Quijano, C., Senadhira, D. and Setter, T. (1998) Strategies for dealing with micronutrient disorders and salinity in lowland rice systems. Field Crops Research **56:** 139-155

Nielsen, F.H. (2000) Evolutionary events culminating in specific minerals becoming essential for life. European Journal of Nutrition **39:** 62-66

Nord, E.A. and Lynch, J.P. (2009) Plant phenology: a critical controller of soil resource acquisition. Journal of Experimental Botany **60:** 1927-1937

Nyrén, P. (2007) The History of Pyrosequencing®. Methods in Molecular Biology **373:** 1-13

Ó Lochlainn, S., Fray, R. G., Hammond, J. P., King, G. J., White, P. J., Young, S. D. and Broadley, M. R. (2011) Generation of nonvernal-obligate, faster-cycling Noccaea caerulescens lines through fast neutron mutagenesis. New Phytologist **189:** 409-414

Ohl, S.A., Sijmons, P.C., Klein-Van der Lee, F.M., Goddijn, O. and Klap, J. (2002) Nematode-inducible regulatory DNA sequences, Zeneca Limited, London UK.

O'Neill, C.M. and Bancroft, I. (2000) Comparative physical mapping of segments of the genome of Brassica oleracea var. alboglabra that are homoeologous to sequenced regions of chromosomes 4 and 5 of Arabidopsis thaliana. Plant Journal **23:** 233-243

Outten, C.E. and O'Halloran, T.V. (2001) Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. Science **292:** 2488-2492

Palmgren, M.G., Clemens, S., Williams, L.E., Krämer, U., Borg, S., Schjørring, J.K. and Sanders, D. (2008) Zinc biofortification of cereals: problems and solutions. Trends in Plant Science **13:** 464-473

Palozza, P. and Krinsky, N.I. (1992) [Beta]-carotene and [alpha]-tocopherol are synergistic antioxidants. Archives of Biochemistry and Biophysics **297:** 184-187

Papoyan, A. and Kochian, L.V. (2004) Identification of Thlaspi caerulescens genes that may be involved in heavy metal hyperaccumulation and tolerance. Characterization of a novel heavy metal transporting ATPase. Plant Physiology **136:** 3814-3823

Park, S., Cheng, N.H., Pittman, J.K., Yoo, K.S., Park, J., Smith, R.H. and Hirschi, K.D. (2005) Increased calcium levels and prolonged shelf life in tomatoes expressing Arabidopsis H^+/Ca^{2+} transporters. Plant Physiology **139:** 1194-1206.

Parkin, I.A.P., Gulden, S.M., Sharpe, A.G., Lukens, L., Trick, M., Osborn, T.C. and Lydiate, D.J. (2005) Segmental structure of the Brassica napus genome based on comparative analysis with Arabidopsis thaliana. Genetics **171:** 765-781

Parsons, B.L. and Heflich, R.H. (1997) Genotypic selection methods for the direct analysis of point mutations. Mutation Research **387:** 97-121

Paterson, A.H., Lan, T.H., Amasino, R., Osborn, T.C. and Quiros, C. (2001) Brassica genomics: a complement to, and early beneficiary of, the Arabidopsis sequence. Genome Biology **2:** 1011.1-1011.4

Peer, W.A., Mamoudian, M., Freeman, J.L., Lahner, B., Richards, E.L., Reeves, R.D., Murphy, A.S. and Salt, D.E. (2006) Assessment of plants from the Brassicaceae family as genetic models for the study of nickel and zinc hyperaccumulation. New Phytologist **172:** 248-260

Peer, W.A., Mamoudian, M., Lahner, B., Reeves, R.D., Murphy, A.S. and Salt, D.E. (2003) Identifying model metal hyperaccumulating plants: germplasm analysis of 20 Brassicaceae accessions from a wide geographical area. New Phytologist **159:** 421-430

Pence, N.S., Larsen, P.B., Ebbs, S.D., Letham, D.L.D., Lasat, M.M., Garvin, D.F., Eide, D. and Kochian, L.V. (2000) The molecular physiology of heavy metal transport in the Zn/Cd hyperaccumulator Thlaspi caerulescens. Proceedings of the National Academy of Sciences of the United States of America **97:** 4956-4960

Persans, M.W., Nieman, K. and Salt, D.E. (2001) Functional activity and role of cation-efflux family members in Ni hyperaccumulation in Thlaspi goesingense. Proceedings of the National Academy of Sciences of the United States of America **98:** 9995-10000

Pettersson, E., Lundeberg, J. and Ahmadian, A. (2009) Generations of sequencing technologies. Genomics **93:** 105-11

Pfeiffer, C.C. and Braverman, E.R. (1982) Epochal trace elements and evolution. Inflammation Research **12:** 412-415

Piñeros, M.A., Shaff, J.E. and Kochian, L.V. (1998) Development, characterization, and application of a cadmium-selective microelectrode for the measurement of cadmium fluxes in roots of Thlaspi species and wheat. Plant Physiology **116:** 1393-1401

Pittman, J.K., Shigaki, T. and Hirschi, K.D. (2005) Evidence of differential pH regulation of the Arabidopsis vacuolar Ca^{2+}/H^{+} antiporters CAX1 and CAX2. Federation of European Biochemical Societies (FEBS) Letters **569:** 2648-2656

Pollard, A.J. and Baker, A.J.M. (1996) Quantitative genetics of zinc hyperaccumulation in Thlaspi caerulescens. New Phytologist **132:** 113-118

Prasad, A.S. (1991) Discovery of human zinc deficiency and studies in an experimental human model. American Journal of Clinical Nutrition **53:** 403-412

Qian, X.D., Eguchi, T., Yoshida, S. and Chikushi, J. (2005) Analytical model for zinc uptake by root system of Thlaspi caerulescens. Journal of the Faculty of Agriculture Kyushu University **50:** 443-458

Rana, D., Boogaart, T., O'Neill, C.M., Hynes, L., Bent, E., Macpherson, L., Park, J.Y., Lim, Y.P. and Bancroft, I. (2004) Conservation of the microstructure of genome segments in Brassica napus and its diploid relatives. The Plant Journal **40:** 725-733.

Reeves, P.A., He, Y., Schmitz, R.J., Amasino, R.M., Panella, L.W. and Richards, C.M. (2007) Evolutionary conservation of the FLOWERING LOCUS C-mediated vernalization response: Evidence from the sugar beet (Beta vulgaris). Genetics **176:** 295-307

Reeves, R.D. (2003) Tropical hyperaccumulators of metals and their potential for phytoextraction. Plant and Soil **249:** 57-65

Reeves, R.D. and Baker, A.J.M. (2000) Metal-accumulating plants. In I. Raskin and B.D. Ensley (eds), Phytoremediation of Toxic Metals: Using Plants to Clean Up the Environment. John Wiley & Sons, New York, NY, USA, pp 193–229

Reeves, R.D. and Brooks, R.R. (1983) European species of Thlaspi L (Cruciferae) as indicators of nickel and zinc. Journal of Geochemical Exploration **18:** 275-283

Reeves, R.D., Schwartz, C., Morel, J.L. and Edmondson, J. (2001) Distribution and metal-accumulating behavior of Thlaspi caerulescens and associated metallophytes in France. International Journal of Phytoremediation **3:** 145-172

Regvar, M., Vogel, K., Irgel, N., Wraber, T., Hildebrandt, U., Wilde, P. and Bothe, H. (2003) Colonization of pennycresses (Thlaspi spp.) of the Brassicaceae by arbuscular mycorrhizal fungi. Journal of Plant Physiology **160:** 615-626

Rengel, Z. (2001) Genotypic differences in micronutrient use efficiency in crops. Communications in Soil Science and Plant Analysis **32:** 1163-1186

Rengel, Z. and Graham, R.D. (1995) Importance of seed Zn content for wheat growth on Zn-deficient soil. Plant and Soil **173:** 259-266

Rengel, Z. and Wheal, M.S. (1997) Kinetic parameters of Zn uptake by wheat are affected by the herbicide chlorsulfuron. Journal of Experimental Botany **48:** 935-941

Richau , K.H. and Schat, H. (2009) Intraspecific variation of nickel and zinc accumulation and tolerance in the hyperaccumulator Thlaspi caerulescens. Plant Soil **314:** 253-262.

Rigola, D., Fiers, M., Vurro, E. and Aarts, M.G.M. (2006) The heavy metal hyperaccumulator Thlaspi caerulescens expresses many species-specific genes, as identified by comparative expressed sequence tag analysis. New Phytologist **170:** 753-765

Robinson, B.H., Leblanc, M., Petit, D., Brooks, R.R., Kirkman, J.H. and Gregg, P.E.H. (1998) The potential of Thlaspi caerulescens for phytoremediation of contaminated soils. Plant and Soil **203:** 47-56

Rogers, C., Wen, J., Chen, R. and Oldroyd, G. (2009) Deletion based reverse genetics in Medicago truncatula. Plant Physiology **151:** 1077-086

Roosens, N., Verbruggen, N., Meerts, P., Ximenez-Embun, P., and Smith, J.A.C. (2003) Natural variation in cadmium tolerance and its relationship to metal hyperaccumulation for seven populations of Thlaspi caerulescens from western Europe. Plant, Cell and Environment **26:** 1657-1672

Rosenzweig, A.C. (2002) Metallochaperones: bind and deliver. Chemistry and Biology **9:** 673-677

Salgueiro, M.J., Zubillaga, M., Lysionek, A., Caro, R., Weill, R. and Boccio, J. (2002) Fortification strategies to combat zinc and iron deficiency. Nutrition Reviews **60(2):** 52-58

Salt, D.E., Baxter, I. and Lahner, B. (2008) Ionomics and the study of the plant ionome. Annual Review of Plant Biology **59:** 709-733

Salt, D.E., Kato, N., Krämer, U., Smith, R.D. and Raskin, I. (2000) The role of root exudates in nickel hyperaccumulation and tolerance in accumulator and nonaccumulator species of Thlaspi. In N. Terry and G. Bañuelos (eds), Phytoremediation of contaminated soil and water. Lewis Publishers, Boca Raton, FL, USA pp 189–200

Salt, D.E., Prince, R.C., Baker, A.J.M., Raskin, I. and Pickering, I.J. (1999) Zinc ligands in the metal hyperaccumulator Thlaspi caerulescens as determined using X-ray absorption spectroscopy. Environmental Science & Technology **33:** 713-717

Sambrook, J. and Russell, D.W. (2001) Preparation of plasmid DNA by alkaline lysis with SDS maxipreparation. In J. Sambrook and D.W. Russell (eds), Molecular cloning: a laboratory manual, Volume 3 (third edition). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., USA, pp 1.31

Sandström, B. (2001) Micronutrient interactions: effects on absorption and bioavailability. British Journal of Nutrition **85:** S181-S185

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Science of the United States of America **74:** 5463–5467.

Sarret, G., Saumitou-Laprade, P., Bert, V., Proux, O., Hazemann, J.-L., Traverse, A.S., Marcus, M.A. and Manceau, A. (2002) Forms of zinc accumulated in the hyperaccumulator Arabidopsis halleri. Plant Physiology **130:** 1815-1826

Sattelmacher, B. (2001) The apoplast and its significance for plant mineral nutrition. New Phytologist **149:** 167-192

Scaife, A. and Turner, M. (1983) Diagnosis of mineral disorders in plants, Vol 2. Vegetables. Her Majesty's Stationary Office, London, UK

Schlichting, C.D. and Pigliucci, M. (1998) Phenotypic Evolution: A Reaction Norm Perspective. Sinauer, Sunderland, MA, USA, pp 340

Schranz, M.E., Quijada, P., Sung, S.-B., Lukens, L., Amasino, R. and Osborn, T.C. (2002) Characterization and effects of the replicated Flowering Time Gene FLC in Brassica rapa. Genetics **162:** 1457-1468

Schwambach J., Fadanelli C. and Fett-Neto, A.G. (2005) Mineral nutrition and adventitious rooting in microcuttings of Eucalyptus globulus. Tree Physiology **25:** 487-494

Schwartz, C., Morel, J.L., Saumier, S., Whiting, S.N. and Baker, A.J.M. (1999) Root development of the Zinc-hyperaccumulator plant Thlaspi caerulescens as affected by metal origin, content and localization in soil. Plant and Soil **208:** 103-115

Sebastian, R.L., Howell, E.C., King, G.J., Marshall, D.F. and Kearsey, M.J. (2000) An integrated AFLP and RFLP Brassica oleracea linkage map from two morphologically distinct doubled-haploid mapping populations. Theoretical and Applied Genetics **100:** 75-81

Shankar, A.H. and Prasad, A.S. (1998) Zinc and immune function: the biological basis of altered resistance to infection. The American Journal of Clinical Nutrition **68:** 447S-463S

Sharma, C.P. (2006) Plant micronutrients. Science Publishers, Enfield, NH, USA

Shen, Z.G., Zhao, F.J. and McGrath, S.P. (1997) Uptake and transport of zinc in the hyperaccumulator Thlaspi caerulescens and the non-hyperaccumulator Thlaspi ochroleucum. Plant, Cell and Environment **20:** 898-906

Shigaki, T., Barkla, B.J., Miranda-Vergara, M.C., Zhao, J., Pantoja, O. and Hirschi, K.D. (2005) Identification of a crucial histidine involved in metal transport activity in the Arabidopsis Cation/H^+ Exchanger CAX1. The Journal of Biological Chemistry **280:** 30136-30142

Shigaki, T., Pittman, J.K. and Hirschi, K.D. (2003) Manganese specificity determinants in the Arabidopsis metal/H+ antiporter CAX2. The Journal of Biological Chemistry **278:** 6610-6617

Shrimpton, R., Gross. R., Darnton-Hill. I. and Young, M. (2005) Zinc deficiency: what are the most appropriate interventions? British Medical Journal **330:** 347-349

Siemianowski, O., Mills, R.F., Williams, L.E. and Antosiewicz, D.M. (2010) Expression of the P_{1B}-type ATPase AtHMA4 in tobacco modifies Zn and Cd root to shoot partitioning and metal tolerance. Plant Biotechnology Journal **10:** 1-11

Sijmons, P.C., Dekker, B.M.M., Schrammeijer, B., Verwoerd, T.C., Vandenelzen, P.J.M. and Hoekema, A. (1990) Production of correctly processed human serum-albumin in transgenic plants. Bio-Technology **8:** 217-221

Sinclair, S.A., Sherson, S.M., Jarvis, R., Camakaris, J. and Cobbett, C.S. (2007) The use of the zinc-fluorophore, Zinpyr-1, in the study of zinc homeostasis in Arabidopsis roots. New Phytologist **174:** 39-45

Singh, J., Upadhyay, A.K., Prasad, K., Bahadur, A. and Rai, M. (2007) Variability of carotenes, vitamin C, E and phenolics in Brassica vegetables. Journal of Food Composition and Analysis **20:** 106-112

Skoog, F. (1940) Relationships between zinc and auxin in the growth of higher plants. American Journal of Botany **27:** 939-951

Snowdon, R.J. (2007) Cytogenetics and genome analysis in Brassica crops. Chromosome Research **15:** 85-95

Sommer, A.L. and Lipman, C.B. (1926) Evidence on the indispensable nature of zinc and boron for higher green plants. Plant Physiology **1:** 231-249

Stein, A.J., Nestel, P., Meenakshi, J., Qaim, M., Sachdev, H.P.S. and Bhutta, Z.A. (2007) Plant breeding to control zinc deficiency in India: how cost-effective is biofortification? Public Health Nutrition **10:** 492-501

Stephenson, P., Baker, D., Girin, T., Perez, A., Amoah, S., King, G.J. and Ostergaard, L. (2010) A rich TILLING resource for studying gene function in Brassica rapa. BioMed Central Plant Biology **10:62** pp. 10

Sugino, R.P. and Innan, H. (2006) Selection for more of the same product as a force to enhance concerted evolution of duplicated genes. Trends in Genetics **22:** 642-644

Sung, S. and Amasino, R.M. (2004) Vernilization in Arabidopsis thaliana is mediated by the PHD finger protein VIND3. Nature **427:** 159-164

Talke, I.N., Hanikenne, M. and Krämer, U. (2006) Zinc-dependent global transcriptional control, transcriptional deregulation, and higher gene copy number for genes in metal homeostasis of the hyperaccumulator Arabidopsis halleri. Plant Physiology **142:** 148-167

Tayler, S.I. (2004) Evolution of Zinc Hyperaccumulation in Thlaspi. PhD. University of Exeter, Exeter

Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C., Odden, A.R., Young, K., Taylor, N.E., Henikoff, J.G., Comai, L. and Henikoff, S. (2003) Large-scale discovery of induced point mutations with high-throughput TILLING. Genome Research **13:** 524-530

Timmer, C.P. (2003) Biotechnology and food systems in developing countries. Journal of Nutrition **133:** 3319-3322

Tittarelli, A., Milla, L., Vargas, F., Morals, A., Neupert, C., Meisel, L.A., Salvo-G, H., Peñaloza, E., Muñoz, G., Corcuera, L.J. and Silva, H. (2007) Isolation and comparative analysis of the wheat TaPT2 promoter: identification in silico of new putative regulatory motifs conserved between monocots and dicots. Journal of Experimental Botany **58:** 2573-2582

Tottey, S., Rondet, S.A.M., Borrelly, G.P.M., Robinson, P.J., Rich, P.R. and Robinson, N.J. (2002) A copper metallochaperone for photosynthesis and respiration reveals metal-specific targets, interaction with an importer, and alternative sites for copper acquisition. Journal of Biological Chemistry **277:** 5490-5497

Trick, M., Kwon, S-J., Choi, S.R., Fraser, F., Soumpourou, E., Drou, N., Wang, Z., Lee, S.Y., Yang, T-J., Mun, J-H., Paterson, A.H., Town, C.D., Pires, J.C., Pyo Lim, Y., Park, B.S. and Bancroft, I. (2009) Complexity of genome evolution by segmental rearrangement in Brassica rapa revealed by sequence-level analysis. BioMed Central Genomics **10:539** pp. 16

U, N. (1935) Genome analysis in Brassica with special reference to the experimental formation of B. napus and peculiar mode of fertilisation. Japanese Journal of Botany. **7:** 389-452

van de Mortel, J.E., Almar Villanueva, L., Schat, H., Kwekkeboom, J., Coughlan, S., Moerland, P.D., Ver Loren van Themaat, E., Koornneef, M. and Aarts, M.G.M. (2006) Large expression differences in genes for iron and zinc homeostasis, stress response, and lignin biosynthesis distinguish

roots of Arabidopsis thaliana and the related metal hyperaccumulator Thlaspi caerulescens. Plant Physiology **142:** 1127-1147

van de Mortel, J.E., Schat, H., Moerland, P.E., Ver Loren van Themaat, E., van Der Ent, S., Blankestijn, H., Ghandilyan, A., Tsiatsiani, S. and Aarts, M.G.M. (2008) Expression differences for genes involved in lignin, glutathione and sulphate metabolism in response to cadmium in Arabidopsis thaliana and the related Zn/Cd-hyperaccumulator Thlaspi caerulescens. Plant, Cell and Environment **31:** 301-324

van de Mortel, J.E., Villanueva, L.A., Schat, H., Kwekkeboom, J., Coughlan, S., Moerland, P.D., van Themaat, E.V.L., Koornneef, M. and Aarts, M.G.M. (2006) Large expression differences in genes for iron and zinc homeostatis, stress response, and lignin biosynthesis distinguish roots of Arabidopsis thaliana and the related metal hyperaccumulator Thlaspi caerulescens. Plant Physiology **142:** 1127-1147

Vancanneyt, G., Schmidt, R., Oconnorsanchez, A., Willmitzer, L. and Rochasosa, M. (1990) Construction of an intron-containing marker gene - splicing of the intron in transgenic plants and its use in monitoring early events in Agrobacterium-mediated plant transformation. Molecular & General Genetics **220:** 245-250

Vasconcelos, M., Datta, K., Oliva, N., Khalekuzzaman, M., Torrizo, L., Krishnan, S., Oliveira, M., Goto, F. and Datta, S.K. (2003) Enhanced iron and zinc accumulation in transgenic rice with the ferritin gene. Plant Science **164:** 371-378

Vazquez, A.M. and Linacero, R.(2010) Stress and somaclonal variation. In E-C. Pua and M.R. Davey (eds), Plant Developmental Biology – Biotechnological Perspectives: Volume 2 p. 45-64.

Vazquez, M.D., Barcelo, J., Poschenrieder, C., Madico, J., Hatton, P., Baker, A.J.M. and Cope, G.H. (1992) Localization of zinc and cadmium in Thlaspi caerulescens (Brassicaceae), a metallophyte that can hyperaccumulate both metals. Journal of Plant Physiology **140:** 350-355

Vera-Estrella, R., Miranda-Vergara, M.C. and Barkla, B.J. (2009) Zinc tolerance and accumulation in stable cell suspension cultures and in vitro regenerated plants of the emerging model plant Arabidopsis halleri (Brassicaceae). Planta **229:** 977-986

Verbruggen, N., Hermans, C. and Schat, H. (2009) Molecular mechanisms of metal hyperaccumulation in plants. New Phytologist **181:** 759-776.

Verret, F., Gravot, A., Auroy, P., Leonhardt, N., David, P., Nussaume, L., Vavasseur, A. and Richaud, P. (2004) Overexpression of AtHMA4 enhances root-to-shoot translocation of zinc and cadmium and plant metal tolerance. Federation of European Biochemical Societies (FEBS) Letters **576:** 306-312

Verret, F., Gravot, A., Auroy, P., Preveral, S., Forestier, C., Vavasseur, A. and Richaud, P. (2005) Heavy metal transport by AtHMA4 involves the N-terminal degenerated metal binding domain and the C-terminal His(11) stretch. Federation of European Biochemical Societies (FEBS) Letters **579:** 1515–1522.

Vreugdenhil, D., Aarts, M.G.M. and Koornneef, M. (2005) Exploring natural genetic variation to improve plant nutrient content. In Broadley and P.J. White, P.J. (eds) Plant Nutritional Genomics. Oxford, UK: Blackwell Publishing, p. 201-219

Wang, R., Farrona, S., Vincent, C., Joecker, A., Schoof, H., Turck, F., Alonso-Blanco, C., Coupland, G. and Albani, M.C. (2009) PEP1 regulates perennial flowering in Arabis alpina. Nature **459:** 423-427

Wang, X., Zhang, Q., Sun, X., Chen, Y., Zhai, T., Zhuang, W., Qi, J. and Wang, Z. (2009) Fosmid library construction and initial analysis of end sequences in female half-smooth tongue sole (Cynoglossus semilaevis) Mar Biotechnology **11:** 236-242

Watanabe, T., Broadley, M.R., Jansen, S., White, P.J., Takada, J., Satake, K., Takamatsu, T., Tuah, S.J. and Osaki, M. (2007) Evolutionary control of leaf element composition in plants. New Phytologist **174:** 516-523

Weeks, M.E. and Leicester, H.M. (1968) Discovery of the elements. Journal of Chemical Education (seventh edition) pp. 625-626

Wei, W., Zhang, Y., Han, L., Guan, Z. and Chai, T. (2008) A novel WRKY transcriptional factor from Thlaspi caerulescens negatively regulates the osmotic stress tolerance of transgenic tobacco. Plant Cell Reports **27:** 795-803

Welch, R.M. and Graham, R.D. (2002) Breeding crops for enhanced micronutrient content. Plant and Soil **245:** 205-214

Welch, R.M. and Graham, R.D. (2004) Breeding for micronutrients in staple food crops from a human nutrition perspective. Journal of Experimental Botany **55:** 353-364

Welch, R.M., Webb, M.J. and Loneragan, J.F. (1982) Zinc in membrane function and its role in phosphorus toxicity [Crops]. In A. Scaife (ed), Proceedings of 9th International plant Nutrition Colloquium, Warwick University, England, pp 710 - 715

Wheal, M. and Rengel, Z. (1997) Chlorsulfuron reduces rates of zinc uptake by wheat seedlings from solution culture. Plant and Soil **188:** 309-317

White, J.G. and Zasoski, R.J. (1999) Mapping soil micronutrients. Field Crops Research **60:** 11-26

White, P.J. (1998) Calcium channels in the plasma membrane of root cells. Annals of Botany **81:** 173-183

White, P.J. (2001) The pathways of calcium movement to the xylem. Journal of Experimental Botany **52:** 891-899

White, P.J. and Broadley, M.R. (2005) Biofortifying crops with essential mineral elements. Trends in Plant Science **10:** 586-593

White, P.J. and Broadley, M.R. (2009) Biofortification of crops with seven mineral elements often lacking in human diets – iron, zinc, copper, calcium, magnesium, selenium and iodine. New Phytologist **182:** 49-84

White, P.J., Whiting, S.N., Baker, A.J.M. and Broadley, M.R. (2002) Does zinc move apoplastically to the xylem in roots of Thlaspi caerulescens? New Phytologist **153:** 201-207

Whiting, S.N., Broadley, M.R. and White, P.J. (2003) Applying a solute transfer model to phytoextraction: Zinc acquisition by Thlaspi caerulescens. Plant and Soil **249:** 45-56

Whiting, S.N., Leake, J.R., McGrath, S.P. and Baker, A.J.M. (2000) Positive responses to Zn and Cd by roots of the Zn and Cd hyperaccumulator Thlaspi caerulescens. New Phytologist **145:** 199-210

WHO (2002) The World Health Report 2002: Reducing Risks, Promoting Healthy Life. World Health Organization, Geneva, Switzerland pp.230.

WHO (2006) Promoting fruit and vegetable consumption around the world, World Health Organization, Geneva, Switzerland http://www.who.int/dietphysicalactivity/fruit/en/index.html (Accessed 13 February 2007).

Wild, J., Hradecna, Z. and Szybalski, W. (2002) Conditionally amplifiable BACs: Switching from single-copy to high-copy vectors and genomic clones. Genome Research **12:** 1434-1444

Williams, L.E., Pittman, J.K. and Hall, J.L. (2000) Emerging mechanisms for heavy metal transport in plants. Biochimica et Biophysica Acta Biomembranes **1465:** 104-126

Williams, P.H. and Hill, C.B. (1986) Rapid-cycling populations of Brassica. Science **232:** 1385-1389

Wither, E.D. and Brooks, R.R. (1977) Hyper-accumulation of nickel by some plants of South-East Asia. Journal of Geochemical Exploration **8:** 579-583

Wolfe, K.H., Gouy, M., Yang, Y.W., Sharp, P.M. and Li, W.H. (1989) Date of the monocot–dicot divergence estimated from the chloroplast DNA sequence data. Proceedings of the National Acadamy of Sciences of the United States of America **86:** 6201 – 6205.

Wood, D.W., Setubal, J.C., Kaul, R., Monks, D.E., Kitajima, J.P., Okura, V.K., Zhou, Y., Chen, L., Wood, G.E., Almeida, N.F., Woo, L., Chen, Y.C., Paulsen, I.T., Eisen, J.A., Karp, P.D., Bovee, D., Chapman, P., Clendenning, J., Deatherage, G., Gillet, W., Grant, C., Kutyavin, T., Levy, R., Li, M.J., McClelland, E., Palmieri, A., Raymond, C., Rouse, G., Saenphimmachak, C., Wu, Z.N., Romero, P., Gordon, D., Zhang, S.P., Yoo, H.Y., Tao, Y.M., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z.Y., Dolan, M., Chumley, F., Tingey, S.V., Tomb, J.F., Gordon, M.P., Olson, M.V. and Nester, E.W. (2001) The genome of the natural genetic engineer Agrobacterium tumefaciens C58. Science **294:** 2317-2323

Wu, J., Schat, H., Sun, R., Koornneef, M., Wang, X. and Aarts, M.G.M. (2007) Characterization of natural variation for zinc, iron and manganese accumulation and zinc exposure response in Brassica rapa L. Plant and Soil **291:** 167-180.

Wu, J., Yuan, Y-X., Zhang, X-W., Zhao, J., Song, X., Li, Y., Li, X., Sun, R., Koornneef, M., Aarts, M.G.M. and Wang, X-W. (2008) Mapping QTLs for mineral accumulation and shoot dry biomass under different Zn nutritional conditions in Chinese cabbage (Brassica rapa L. ssp. pekinensis). Plant and Soil **310:** 25-40.

Yang, Y.-W., Lai, K.-N., Tai, P.-Y. and Li, W.-H. (1999) Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between Brassica and other angiosperm lineages. Journal of Molecular Evolution **48:** 597 – 604.

Ye, X.D., Al-Babili, S., Kloti, A., Zhang, J., Lucca, P., Beyer, P. and Potrykus, I. (2000) Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. Science **287:** 303-305

Yilmaz, A., Ekiz, H., Gultekin, I., Torun, B., Barut, H., Karanlik, S. and Cakmak, I. (1998) Effect of seed zinc content on grain yield and zinc concentration of wheat grown in zinc-deficient calcareous soils. Journal of Plant Nutrition **21:** 2257-2264

Zha, H.G., Jiang, R.F., Zhao, F.J., Vooijs, R., Schat, H., Barker, J.H.A. and McGrath, S.P. (2004) Co-segregation analysis of cadmium and zinc accumulation in Thlaspi caerulescens interecotypic crosses. New Phytologist **163:** 299-312

Zhang, F.S., Romheld, V. and Marschner, H. (1991) Release of zinc mobilizing root exudates in different plant species as affected by zinc nutritional status. Journal of Plant Nutrition **14:** 675-686

Zhang, X., Henriques, R., Lin, S.S., Niu Q.W. and Chua N.H (2006) Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. Nature Protocols **1:** 641-646

Zhang, X.G. (1996) Corrosion and Electrochemistry of Zinc. Plenum Publishing New York, USA

Zhao, F.J., Hamon, R.E. and McLaughlin, M.J. (2001) Root exudates of the hyperaccumulator Thlaspi caerulescens do not enhance metal mobilization. New Phytologist **151:** 613-620

Zhao, F.J., Hamon, R.E., Lombi, E., McLaughlin, M.J. and McGrath, S.P. (2002) Characteristics of cadmium uptake in two contrasting ecotypes of the hyperaccumulator Thlaspi caerulescens. Journal of Experimental Botany **53:** 535-543

Zhao, F.J., Lombi, E. and McGrath, S.P. (2003) Assessing the potential for zinc and cadmium phytoremediation with the hyperaccumulator Thlaspi caerulescens. Plant and Soil **249:** 37-43

Zhao, F.J., Lombi, E., Breedon, T. and McGrath, S.P. (2000) Zinc hyperaccumulation and cellular distribution in Arabidopsis halleri. Plant, Cell and Environment **23:** 507-514

Zhao, J.J., Kulkarni, V., Liu, N.N., Del Carpio, D.P., Bucher, J. and Bonnema, G. (2010) BrFLC2 (FLOWERING LOCUS C) as a candidate gene for a vernalization response QTL in Brassica rapa. Journal of Experimental Botany **61:** 1817-1825

Zhong, S., Khodursky, A., Dykhuizen, D.E. and Dean, A.M. (2004) Evolutionary genomics of ecological specialization. Proceedings of the National Academy of Science of the United States of America **101:** 11719-11724

Zhou, H., Xia, X.G. and Xu, Z. (2005) An RNA polymerase II construct synthesizes short-hairpin RNA with a quantitative indicator and mediates highly efficient RNAi. Nucleic Acids Research **33:** 1-8