Functional Characterisation of ScIRT1 and ScIREG2 Transport Proteins in the Nickel Hyperaccumulator, *Senecio coronatus*

Masters Dissertation

By: Llewelyn van der Pas

Supervised by: A/Prof. Rob A Ingle

Plagiarism Declaration

I know the meaning of plagiarism and declare that all of the work in the dissertation, save for that which is properly acknowledged, is my own. I have used the Harvard style of referencing and each contribution to, and quotation in, this dissertation from the work(s) of other people has been attributed and has been cited and referenced.

Signed by candidate

Acknowledgements

First and foremost, I need to thank my supervisor Rob Ingle for this project. His passion for Ni hyperaccumulation and the genetics involved was the driving force that kept the project afloat. I would also like to thank Lara Donaldson for her valuable input and troubleshooting skills during the creation of the destination vector and for always being around to listen and share ideas. This work would not have been possible without the assistance from the lab members of Lab 430. I personally want to thank Amy Bruce for being a pillar on which I could lean and for showing me ropes when I came into the lab. My friend Dean Reikert for helping me out when I needed it most. My mother deserves a special thanks for always being in my corner. Last but not least, I want to thank my husband for always providing the support I needed and for pretending to know what I was doing in the lab.

List of Abbreviations used

As	Arsenic
Ala	Alanine
bp	Base pairs
Ca	Calcium
CBB	Calvin-Benson-Bassham cycle
Cd	Cadmium
Co	Cobalt
Cu	Copper
DMG	Dimethylglyoxime
Fe	Iron
His	Histidine
K	Potassium
kb	kilobases
Mg	Magnesium
Mn	Manganese
N	Nitrogen
NA	Nicotianamine
Ni	Nickel
P	Phosphorus
Pb	Lead
PSII	Photosystem II
Sb	Antimony
Se	Selenium
Tl	Thallium
WT	Wildtype
Zn	Zinc

Abstract

Nickel hyperaccumulation is a unique plant adaption that has led to roughly 390 plant taxa being able to not only withstand the toxicity associated with Ni but actively translocate it to aerial tissues. However, the underlining molecular mechanisms that drive Ni hyperaccumulation remain unclear. *Senecio coronatus*, a Ni hyperaccumulator, is a novel species as both hyperaccumulating and non-accumulating populations can be found on the serpentine soils of the Barberton Greenstone Belt, South Africa. A comparative RNA-seq analysis on these populations of *S. coronatus* revealed that *ScIRT1* and *ScIREG2* , putative homologues of the *Arabidopsis* transporters, *AtIRT1* and *AtIREG2* which are capable of transporting Ni, showed much higher expression in the hyperaccumulating populations compared to the non-hyperaccumulating populations, suggesting a potential role in Ni hyperaccumulation. It was thus necessary to investigate whether *ScIRT1* and *ScIREG2* encode functional homologues of these *Arabidopsis* transporters. To accomplish this, *irt1* and *ireg2* mutants were obtained from a T-DNA insertion seed collection and their homozygosity was then determined by PCR genotyping. Since a lack of iron induces *IRT1* and *IREG2* expression, loss of gene expression of homozygous *irt1* and *ireg2* mutants by means of reverse transcriptase PCR on plant roots grown hydroponically in the absence of Fe was then done to establish full knock-out status. From this, homozygous mutants were identified, however, absence of gene expression for *irt1* and *ireg2* mutants was not clear. In addition to validating homozygosity, phenotypic characterisation, with the aim of developing reliable assays to be used in complementation analysis, was done by growing homozygous mutants and Col-0 in hydroponic media deficient in Fe and supplemented with Ni. The assays revealed that under Fe-deficient and Ni-supplemented conditions, a reduction in root biomass was a more reliable phenotypic characteristic for *ireg2* mutants than root length or shoot biomass. In contrast, for *irt1,* no observable phenotype was established under Fe-deficiency conditions. In parallel, Gateway cloning was employed to create expression clones where *ScIRT1* and *ScIREG2* protein coding expression was to be driven by native *Arabidopsis* promoters *pAtIRT1* and *pAtIREG2 (*i.e. *pAtIRT1:ScIRT1* and *pAtIREG2:ScIREG2*) respectively for complementation of the *Arabidopsis irt1* and *ireg2* mutants. The open reading frames of the *S. coronatus* genes and the *Arabidopsis* promoters were PCR amplified, cloned into appropriate pDONR221 vectors, and sequence verified. The *ScIREG2* clone however, revealed point mutations and could not be used. *pAtIRT1* was successfully recombined with *ScIRT1* to generate a two-fragment expression clone which was verified by DNA sequencing. Thus herein, the foundations for *ScIRT1* and *ScIREG2* complementation experiments have been established.

Conventions used

All enzyme names were cross referenced with the International Union of Biochemistry and Molecular Biology (IUBMB) on their website http://www.sbcs.qmul.ac.uk/iubmb/ and relevant enzyme information was also retrieved from *BRENDA* at http://www.brenda-enzymes.org/index.php . Enzyme commission numbers were correct at the time of writing this dissertation.

Table of Contents

List of Tables and Figures

Tables

Figures

INTRODUCTION

1.1 Hyperaccumulation: Terms and Definitions

As sessile organisms, plants must contend with an array of environmental aggressors. To deal with these aggressors, plants have developed remarkable adaptations. One such adaptation is the ability of some plants to not only tolerate excessive amounts of heavy metals, but to actively absorb and translocate them to their above ground tissues (Baker & Brooks, 1989). These plants constitute a small portion of the plant kingdom and are collectively known as metal hyperaccumulators. The term 'hyperaccumulator' was first coined by Jaffré, et al., (1976) who at that time, described the Ni content of the latex from *Sibertia acuminata* as being the highest Ni concentration (25.74%) ever to be found. Little is known about the molecular basis of Ni hyperaccumulation despite a lot of research interest in heavy metal hyperaccumulation. A plant can be considered a Ni hyperaccumulator when Ni concentration in dry weight is in excess of 1000 $mg.kg^{-1}$ in aerial tissues (Brooks, et al., 1977).

Heavy metal hyperaccumulation, known to occur in over 700 described plant taxa, with 532 being Ni hyperaccumulators (Reeves, et al., 2018), is a remarkable phenomenon which has attracted a lot of research in the past two decades (Krämer, 2010) due to its potential application in phytoremediation and phytomining (Rascio & Navari-Izzo, 2011). There is, however, a distinction between plants that are capable of hyperaccumulating and those that are simply hypertolerant. Heavy metal hyperaccumulators are plants that actively absorb and translocate heavy metals to aerial tissues. Hypertolerant plants by contrast, are able to survive on metal rich soils, but do not 'actively' absorb metals. That being said, some hypertolerant species maintain a low concentration of heavy metals in their aerial tissues (Baker & Brooks, 1989). In hyperaccumulating literature, there are obligate hyperaccumulators (or obligate metallophytes), which are species that are restricted to metalliferous soils and account for the largest proportion of hyperaccumulators (Pollard, et al., 2014); and a smaller group of hyperaccumulators called facultative hyperaccumulators which accumulate whenever growing on metalliferous soils (Pollard, et al., 2014), but also grow on non-metal-rich substrates. Whilst 'heavy' metal in chemical terms refers to transition metals which have a mass greater than 20 (Rascio & Navari-Izzo, 2011), in biological terms, 'heavy' metal refers to metals or metalloids than can be potentially phytotoxic to organisms (Rascio & Navari-Izzo, 2011).

1.2 Metal ions in plants

Plant growth is governed by many factors including the presence of at least 17 nutrients. According to Epstein (1965), an element is considered essential when 1) the absence thereof prevents the plant from completing its lifecycle and 2) it forms part of some essential plant metabolite. Plant nutrients are broadly classified into two categories; macro- and micronutrients. The macronutrients include elements essential for the building of structural components of molecules and for energy generation. They are typically required in large concentrations in excess of 10 mmole.kg^{-1} of dry weight (Hopkins & Hünter, 2009). The micronutrients by contrast, are required in much smaller concentrations, typically less than 10 mmole.kg^{-1} of dry weight and are often found to be required for catalytic or regulatory roles (Hopkins & Hünter, 2009). For instance, Cu is a necessary cofactor for many oxidative enzymes such as polyphenol oxidases and superoxide dismutase; whereas Zn is an activator for many enzymes such as alcohol dehydrogenase and carbonic anhydrase, and a component of Zn-finger transcription factors (Hopkins & Hünter, 2009). Cobalt is considered more of a beneficial element rather than an essential element as it is primarily required by nitrogen-fixing bacteria, such as *Azotobacter*, which form symbiotic relationships with many members from the Fabaceae where it increases nitrogen fixation rates (Robson, et al., 1979) by, for instance, enhancing nodulation formation (Dilwroth, et al., 1979). Manganese forms part of a manganoprotein which is part of the oxygen-evolving complex of photosystem (PS) II but can also replace magnesium for reactions that require the hydrolysis of ATP. Critical toxicity for Mn is considerably higher than the other metals, but plants typically experience deficiency of Mn rather than toxicity, especially under low soil pH conditions. Ni, which is required in the range of 0.05 – 5.0 mmol.kg^{-1}, is known to be a constituent of at least eight enzymes (Ryan, et al., 2010) of which most occur within microscopic organisms. However, in plants urease (EC 3.5.1.5), which catalyses the hydrolysis of urea into CO_2 and NH_3, is the only known enzyme to contain Ni (Carter, et al., 2009).

Plant nutrient requirements vary from nutrient to nutrient and are essentially determined by the plant's lower and upper critical concentrations, shown in Figure 1-1. Below the lower critical concentration, plant growth is retarded due to nutrient limitation and deficiency symptoms are evident. Between the lower and upper critical concentrations, the nutrient supply is sufficient for optimal growth and any increase in nutrient concentration has little effect on growth. Above the upper critical concentration, the nutrient can become toxic and can lead to severe toxicity with consequent effects on growth.

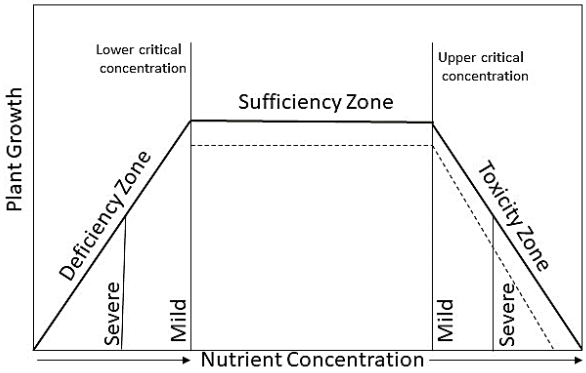


Figure 1-1: Plant growth as a function of Nutrient Concentration*: Hypothetical graph showing plant growth as a function of increasing nutrient concentration. Solid line represents essential nutrients and dotted line represents non-essential nutrients. Figure adapted from Alloway, (2013)*

Of the 17 nutrients required for plant growth, six are transition metals, of which Fe is often in such limiting supply that it is hardly accumulated to any large extent. Five of these; Co, Cu, Mn, Zn, and Ni, are essential nutrients and have been shown to be metals that are hyperaccumulated (though not by a single species).

Table 1-1: *Essential and non-essential metals with critical toxicity and number of plant taxa and families which hyperaccumulate the respective metal*

		Number of Hyperaccumulating[1]	
Metal	**Critical toxicity[2;3] (µg. g^{-1})**	**Plant taxa**	**Plant Families**
Essential Metals			
Co	0.4	26	11
Cu	20 – 30	35	15
Mn	200 – 3 500	10	6
Ni	10 – 50	390	42
Zn	100 – 300	15	6
Non-essential metals/non-metals			
As	<2 – 80	15	2
Cd	6 – 10	5	2
Pb	0.6 – 28	14	7
Se[4]	3 – 100	20	7
Sb	<2	2	2
Tl	20	1	1

[1]Values from (Krämer, 2010)

[2]Values from (Becket 1977; Hong, et al. 2008; Kabata-Pendias and Pendias 2001; Römheld 2012)

[3]Defined as a 10% reduction in plant growth or yield

[4]Some members of Astragalus (Fabaceae) are known to tolerate high concentrations of selenium and its accumulation is thought to serve some essential role in these plants (Hopkins & Hünter, 2009)

Other non-essential metals and non-metals that have been shown to be hyperaccumulated are As, Cd, Pb, Se, Sb, and Tl. Table 1-1 above shows the essential and non-essential

metals and non-metals that are hyperaccumulated along with their respective critical toxicity values and the number of plant taxa and families that are hyperaccumulators. A noticeable feature of Table 1-1 is the relatively low concentration required for Ni to be toxic, yet roughly 73% of known hyperaccumulators hyperaccumulate Ni (Krämer, 2010). This percentage begs the question; if 73% of the 450 known hyperaccumulators can tolerate Ni, what happens to the rest of the plant kingdom if they are exposed to Ni?

1.3 Nickel Toxicity

Heavy metal phytotoxicity arises due to heavy metals interfering with cellular/molecular process by for instance, rendering some enzymes inactive, displacing essential elements due to chemical similarity of the divalent cations, and/or obstructing vital functional groups (Rascio & Navari-Izzo, 2011). The accumulation of heavy metals and subsequent heavy metal poisoning commonly results in the production of reactive oxygen species (ROS) (Halliwell & Gutteridge, 1984; Erdei, et al., 2002; Pandey, et al., 2009) due to for instance, interference with electron transport processes (Seregin & Kozhevnikova, 2006; Rascio & Navari-Izzo, 2011). Typical consequences of ROS and oxidative stress include peroxidation of lipids which can lead to membranes disassembling, and oxidative weakening of macromolecules such as DNA via strand cleavage (Rascio & Navari-Izzo, 2011) which can lead to single or double strand DNA breakage leading to genotoxicity and genome instability (Dutta, et al., 2018). Besides the aforementioned, heavy metal accumulation can have other nonspecific effects such as inhibiting mineral absorption; alteration of water management by altering stomatal conductance and transpiration rates; reduced or declined photosynthesis which alters downstream application of photosynthates such as a reduction in respiration rates; and growth and morphological developmental inhibition due to anatomical changes (Seregin & Kozhevnikova, 2006).

1.3.1 Effect of Nickel on Enzymatic Activity

Heavy metals can interfere with enzymatic activity by for instance interacting with thiol groups, which can lead to conformational changes (Seregin & Kozhevnikova, 2006). For example, inactivation of some enzymes, such as those that participate in committed steps in metabolic pathways, can lead to metabolic disorders. It is worthwhile to mention that *in vitro* metal toxicity on enzymes does not necessarily reflect *in vivo* effects (Seregin & Kozhevnikova, 2006).

Key enzymes such as Rubisco (Ribulose 1,5-bisphosphate carboxylase/oxygenase, EC 4.1.1.39) show decreased activity at 0.5 and 1mM Ni (Shoeran , et al., 1990). Enzymes in the Calvin-Benson-Bassham (CBB) cycle such as glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12), 3-phosphoglycerate kinase (EC 2.7.2.3), fructose 1,6-bisphosphatase (EC

3.1.3.11), and fructose-bisphosphate aldolase (EC 4.1.2.13) all show decreased activity in *Cajanus cajan* at 0.5 and 1 mM Ni (Shoeran , et al., 1990). At 1 mM Ni, nitrate reductase (EC 1.7.1.2) and glutamine synthetase (EC 6.3.1.2) in *Beta vulgaris* show decreased activity (Kevresan, et al., 1998). *Glycine max* showed decreased activity of alanine aminotransferase (EC 2.6.1.2) when exposed to 0.2 mM Ni (El-Shintinawy & El-Ansary, 2000).

1.3.2 Effect of Nickel on Photosynthesis

Any alteration of photosynthetic enzymes of carbon fixation and the CBB cycle as well as interference or displacement of metals required by light harvesting complexes will have serious impacts on overall photosynthetic productivity. This in turn will have downstream effects on other processes such as nitrogen assimilation and cellular respiration which are closely tied to photosynthetic productivity. Figure 1-2 below shows the various stages where Ni can interfere with the electron transport chain of photosynthesis. Ni has inhibitory effects on electron transport by changing the structure of plastoquinone Q_B or the reaction centre proteins (Mohanty, et al., 1989; Krupta & Baszynski, 1995). Ni further decreases the content of ferredoxin, plastocyanin and the cytochromes b_{6f} and b_{559} which directly causes electron transport to be reduced (Veeranjaneyulu & Das, 1982).

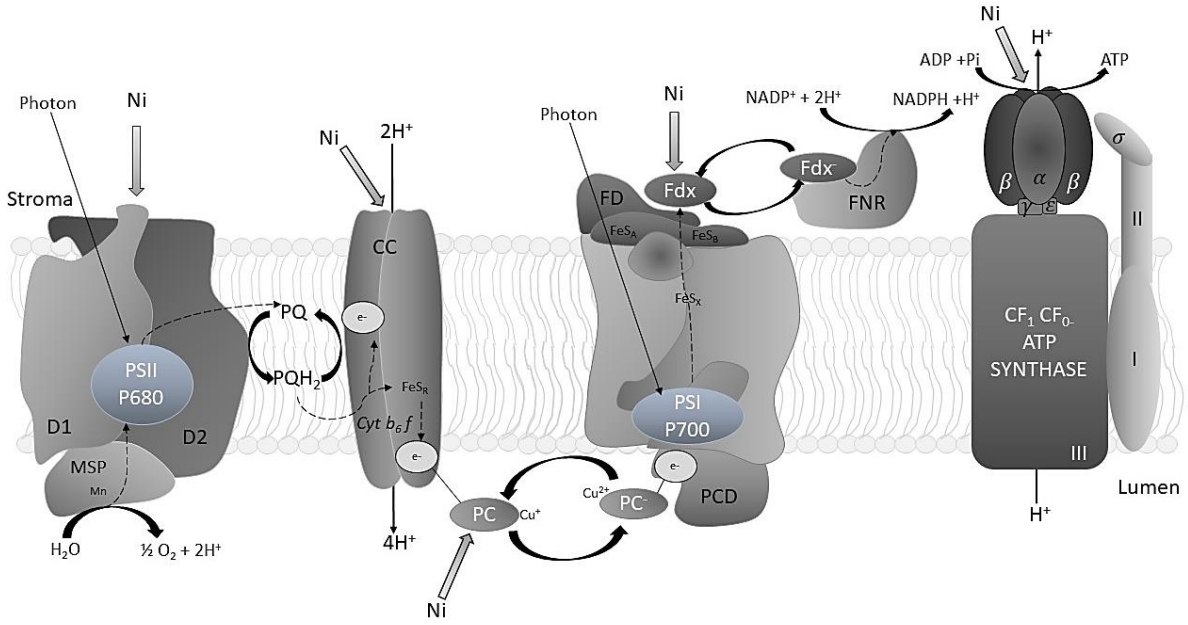


Figure 1-2: Representation of the electron transport chain of photosynthesis *with points of interest where Ni interferes. Dotted lines indicate the flow of electrons. MSP= manganese stabilising protein; D1 and D2= PSII reaction proteins; PQ= oxidised plastoquinone; PQH_2= reduced plastoquinone; CC= cytochrome complex; $FeS_{R,X,A,B}$= Fe-containing Rieske proteins; PC= plastocyanin; PCD= plastocyanin docking; FD= ferredoxin docking; Fdx= ferredoxin; FNR= ferredoxin-NADP+ oxidoreductase. Figure modified from (Seregin & Kozhevnikova, 2006; Campbell & Farrell, 2012; Graham, et al., 2014)*

Alteration in the proper functioning of many of these aforementioned electron transport complexes can lead to the generation of ROS. Furthermore, Ni has been shown to cause

disruptions to chloroplast structures, prevention of chlorophyll synthesis, and disordering of electron transport complexes (Seregin & Ivanov, 2001). It has been shown by numerous authors (Seregin & Ivanov, 2001; Veeranjaneyulu & Das, 1982; Mohanty, et al., 1989; Krupta & Baszynski, 1995; Maksymiec, 1997) that Ni directly affects various stages during photosynthesis. PSII activity is further reduced due to the accumulation of ATP and NADPH from blocked CBB cycle reactions, which creates a high pH gradient across the thylakoid membrane (Krupta & Baszynski, 1995).

1.4 Evolution of Nickel Hyperaccumulation

Ni hyperaccumulation has evolved independently multiple times across various families, though is most prevalent in the Asteraceae, Brassicaceae, and Euphorbiaceae families (Krämer, 2010; Seregin & Kozhevnikova, 2006; Borhldi, 2001). Given the broad distribution of Ni hyperaccumulation across unrelated families it is clear that it is of polyphyletic origin (Macnair, 2003; Mengoni, et al., 2003). Although the Ni-hyperaccumulating trait is most prevalent among the dicotyledons, there are at least two monocotyledons (Poaceae and Juncacea) and one pteridophyte (Salvinaceae) that hyperaccumulate Ni, suggesting that the acquisition of the trait was random (Borhldi, 2001) and may well have been in response to being in serpentine conditions. The global distribution of Ni hyperaccumulators further reflects the global distribution of serpentine sites (Krämer, 2010) suggesting that in some part the occurrence of Ni hyperaccumulators may have evolved in response to serpentine soil. For example, some 'hotspots' for Ni hyperaccumulators include Cuba, New Caledonia, the Balkan peninsula, California, and Turkey, which are also known to be areas of high serpentine soils. Serpentine soils are characterised by having a high magnesium to calcium ratio, with many having larger amounts of heavy metals such as Cr, Co, and Ni derived from Fe and Mg rich ultramafic rocks (Callahan, et al., a 2006).

Within the Brassicaceae family, the genus *Alyssum*, has been used to study the evolutionary history of Ni hyperaccumulation. (Mengoni, et al., 2003) used the internal transcribed spaces (ITSs) from the 18S – 26S nrDNA (nuclear ribosomal DNA) to study the evolutionary dynamics of *Alyssum* and identified accumulators and non-accumulators within each of the four clades within the *ITS1* phylogeny, implying multiple evolutionary origins for Ni hyperaccumulation. From this they were able to deduce that the hyperaccumulators and non-hyperaccumulators belong to the same taxon. Several studies have also revealed significant variation among populations of hyperaccumulating species (Verbruggen, et al., 2009). For instance in a review by Baker & Brooks, (1989), where they listed most of the hyperaccumulating species globally,

the Ni hyperaccumulating genus *Alyssum* had a mean maximum Ni content of 12 214.4 ± 6 278.2 µg.mg^{-1} whereas the genus *Thlaspi* had 9 000.4 ± 7 915.4 µg.mg^{-1}. Whilst it clear from just the two examples given that there is significant variation in Ni content, all of the species represented are able to accumulate at least 1 000 mg.kg^{-1} Ni.

1.5 Evolutionary advantage of Ni hyperaccumulation

In a review done by Boyd and Martens (1992), they summarised four postulated benefits of metal hyperaccumulation. The first hypothesis, the tolerance/disposal hypothesis, suggests that hyperaccumulation evolved as a mechanism to sequester metals in tissues (tolerance) and, in some plants, disposal by shedding those tissues that contain the metal (Boyd & Martens, 1992; Boyd, 2004). The second hypothesis, the interference hypothesis, (renamed the elemental allelopathy hypothesis by Boyd & Martens, [1998]), suggests that perennial hyperaccumulators enrich the soil surface below their canopies with high-metal litter as a deterent for the establishment of less tolerant species. The third hypothesis, the drought resistance hypothesis, suggests that hyperaccumulation developed as a means to be more tolerant towards drought, though this has not been shown to be case. The fourth hypothesis, the elemental defence hypothesis, states that by concentrating heavy metals in aerial tissue, herbivory and pathogenic infections can be deterred. It is the latter hypothesis which has attracted the most attention and also seems the most plausible with several lines of evidence providing support.

Martins & Boyd, (1994) proposed that elemental defence differs from secondary metabolite (alkaloids, terpenes, glucosinolates) defence in that the toxic component is taken up from the soil rather than produced via photosynthate derivatives. In addition, they further state that the herbivores/pathogens which encounter the hyperaccumulator lack the proper biochemistry to eliminate the heavy metal (Martens & Boyd, 1994). There are numerous studies that have demonstrated support for the elemental defence hypothesis. Most notably, Jhee, et al. (2006) showed that the diamondback moth, *Plutella xylostella*, laid 2.5-fold more eggs on the Ni hyperaccumulator *Strepthanthus polygaloides* which were grown under low-Ni conditions. They concluded that Ni hyperaccumulation served as a viable defense mechanism that not only improved plant fitness, evident in better growth, but also provided sufficient toxicity for diamondback moth larvae. More recently, (Mincey & Boyd, 2018) showed that seeds from *S. polygaloides*, which contained approximately 300 µg Ni g^{-1}, had a greater effect on mortality rates of the generalist granivore beetle *Tribolium confusum* when compared to a non-hyperaccumulating relative *S. insignis*, which contained approximately 5 µg Ni g^{-1}. The authors

further observed that Ni concentrations were greater in the embryo than in the seed coat which could assist in better reproductive success by protecting the seed and developing seedling against herbivory (Mincey & Boyd, 2018).

Since evolution is dynamic, it is perhaps not that surprising that there are some cases of co-adaptation and co-evolution of hyperaccumulation and herbivores/pathogens (Boyd, 2004). For example, Martens & Boyd, (2002) conducted a field study in California of *S. polygaloides* and found that mammalian herbivores damaged Ni-containing plants. They proposed that the generalist diet of these herbivores diluted Ni-containing foliage with that of low-Ni foliage (Martens & Boyd, 2002). In addition. there are some specialist herbivores which seem to have evolved in association with hyperaccumulators, for example the mirid bug *Melanotrichus boydi* which feeds on *S. polygaloides* (Schwartz & Wall, 2001), and *Chrysolina pardalina*, a chrysomelid beetle, which feeds on *Berkheya coddii* (Mesjasz-Przybylowicz & Przybylowicz, 2001). Owing to the toxic effects of hyperaccumulators, it has been assumed that mycorrhizal symbioses are less likely to occur. This is mostly true for the Brassicaceae members because these members are characteristically non-mycorrhizal (Smith & Read, 1997; Wang & Qui, 2006). However, in 2003 (Turnau & Mesjasz-Przybylowicz, 2003) reported that *B. coddii* and three other Ni hyperaccumulators from the Asteraceae in South Africa were colonised by arbuscular mycorrhizal fungi (AMF). In a much more in-depth study, (Amir, et al., 2007) looked at nine different hyperaccumulating species across different families in New Caledonia, which hosts a great diversity of Ni-hyperaccumulators, and found that all nine of them had AMF colonisation. Although above a certain threshold AMF spore germination was inhibited, evidence points to these AMFs having developed some degree of tolerance (Amir, et al., 2007).

1.6 Physiology of Metal Hyperaccumulation

Rascio and Navari-Izzo (2011) propose that hyperaccumulation is dependent on three characteristics; 1) plants need to have a strong capacity for the uptake of heavy metals, 2) they need to have a very effective root-to-shoot translocation mechanism, and 3) they should have a mechanism(s) to detoxify and sequester heavy metals. Should a plant fail to meet these characteristics it cannot be deemed a hyperaccumulator. In the following section a generalised mechanism by which the above characteristics occur will be discussed.

1.6.1 Nickel uptake

The existence of a Ni-specific transporter in roots has not been identified yet. It is, however, thought that Ni uptake occurs as a result of cation transporters that have low selectivity and

specificity such as those belonging to the ZRT/IRT-like (ZIP) family of transport proteins. One such protein is IRT1 (iron-regulated transporter). *IRT1*, which is primarily expressed in root epidermal cells, was found to be induced under Fe deficiency stress (Vert, et al., 2002) but it has a broad specificity towards many cations such as zinc, cobalt, cadmium, manganese and Ni. Though the family comprises many members, all thought to be involved in initial uptake and translocation, numerous authors have shown altered gene regulation between hyperaccumulating and non-hyperaccumulating relatives such as the differential expression between the hyperaccumulator *N. caerulescens* compared to *A. thaliana* (van de Mortel, et al., 2008), *IRT1* has shown the most promise as being involved in the initial uptake of Ni. Ni is a divalent cation and can, for instance, compete with Fe during this initial uptake process. However, this can only be achieved as noted earlier when Fe is limiting.

Given the hypothetical entry of Ni through these low-specificity transporters (Taylor & Macnair, 2006) reported that Ni accumulation in the Ni hyperaccumulator *Noccaea pindicum* was inhibited when hydroponic media was supplemented with equimolar concentrations of Zn and Ni compared to plants supplemented with Ni only, whereas for *N. caerulescens*, a significant reduction in Ni uptake was noted (Mohseni, et al., 2018) under similar condition, suggesting that at least for *Noccaea*, uptake of Ni may occur through transport of Zn transporters. Contrary to this, (Mohseni, et al., 2018) found that Ni uptake in *Alyssum bracteatum* was rather 'dependent' on Fe concentration, whereas *A. inflatum* Ni uptake was unaffected by Fe concentration. This suggests that even within the same genus, Ni does not enter the plant via the same non-selective transporter. That being said, as noted above, the altered expression of members of the ZIP family is primarily thought to be involved in this initial Ni uptake. Comparative RNA-seq of *Senecio coronatus*, a Ni hyperaccumulator from South Africa, showed high expression of *IRT1* in the roots of hyperaccumulator versus non-hyperaccumulator population on serpentine soil (Meier, et al., 2018). It is thus possible to suggest that non-specific transporters are primary sources for Ni uptake, but they vary among hyperaccumulators; within *Noccaea*, Zn transporters are implicated whereas within *Alyssum*, Fe transporters are implicated. Once within the plant body of a Ni hyperaccumulator, Ni is primarily depostisted within the shoot epiermis (Jaffré , et al., 2018; Küpper , et al., 2011), trichomes (Ghasemi, et al., 2009), and the vacuole (Schaaf, et al., 2006), but can also be found in other plant organs such reproductive the organs (Meindi, et al., 2014), though to varying degrees in different species.

1.6.2 Nickel chelation and xylem loading

Chelation of metals from soil to plant is an important aspect for the acquisition of many metals for plants. For example, Fe is relatively insoluble largely due to the formation of hydrous oxides

as pH increases (Callahan, et al., a 2006). Ligands and chelates play essential roles in the overall absorption, sequestration, transportation, and storage of metals in plants (Callahan, et al., a 2006). Some plants such as many dicotyledons and non-grassy monocotyledons, use the exudation of H^+ as a means to increase the solubility of certain metals, such as Fe^{2+}, whilst other plants, notably graminaceous monocotyledons, exude phytosiderophores which help co-ordinate Fe^{3+} in the rhizosphere. There is a plethora of potential chelating agents and ligands that could be involved in enhancing tolerance towards Ni, however, their active role in hyperaccumulation remains unclear. Histidine, the α-amino acid with its imidazole side chain, is able to act as a ligand owing to the tridentate nature of its carboxy, amino, and imidazole groups. Interestingly, histidine has a relatively high association constant with Ni (8.7 Ig K) (Martel & Smith, 1974). Elevated histidine was noted in Brassicaseae hyperaccumulators such as *N. geosingense* and *A. lesbiacum* when compared to their non-accumulating relatives (Ingle, et al., 2005; Persans, et al., 1999). Histidine biosynthesis follows a regulated pathway and it is likely that the initial enzyme involved in this pathway, ATP-phosphoribosyl transferase (ATP-PRT), is likely to be responsible for increasing the amount of histidine in hyperaccumulators. Ingle, et al., (2005); Rees, et al., (2009); and Wycisk, et al., (2004), for example, found that when *ATP-PRT* is overexpressed in Arabidopsis plants, histidine content is elevated and tolerance towards Ni is noted. It is not just the initial enzyme of the histidine biosynthetic pathway that shows increase expression in hyperaccumulators versus their non-hyperaccumulating relatives. De la Torre, et al., (2018) found that orthologs of histidinol dehydrogenase was more expressed in Ni hyperaccumulators across various plant families. They claim that this provides further support for histidine's role in Ni hyperaccumulation, though presence of higher transcription does not necessarily correspond to increased enzyme activity, or concentrations of free His. A study by Kerkeb & Krämer, (2003) proposes that histidine increases the mobilisation of Ni towards and into xylem once Ni is within the root symplasm and that Ni uptake was independent of histidine uptake. Whilst there has been no conclusive evidence for the presence of Ni specific transport protein(s) that are involved in xylem loading, evidence from Ingle, et al., (2005) and Krämer, et al., (1996) suggests that at least among the *Alyssum* hyperaccumulators a potential transport protein exists which seems absent in their non-hyperaccumulating relatives.

Phytosiderophores are ligands which are exuded by graminaceous plants and play a role in metal chelation. Siderophores by contrasts are chelating agents used by microorganisms. Nicotianamine (NA) is a precursor to some phytosiderophores such as mugineic and avenic acid (Callahan, et al., a 2006). Nicotianamine synthase (NAS, EC 2.5.1.4.3) catalyses the condensation of three S-adenosyl-methionine molecules to produce NA. NA is then converted to mugineic acid derivatives via nicotianamine aminotransferase (NAAT, EC 2.6.1.80). NA,

compared to histidine, has a much greater association constant for Ni (16.1 lg K) (Benes, et al., 1983). (Callahan, et al., b 2006), reported a positive correlation between NA content and shoot Ni in *N. caerulescens* and other species, whereas the overexpression of NAS enhanced Ni tolerance (Callahan, et al., b 2006). However no conclusive evidence has been found that directly links higher NA content with hyperaccumulation.

Other chelating agents which may be involved in Ni hyperaccumulation include several organic acids such as citric, isocitric, oxalic, tartaric, malic, malonic, and aconitic acid (Callahan, et al., a 2006). Of these, elevated levels of citrate (the anion of citric acid) has been reported in several hyperaccumulators (Lee, et al., 1977; Lee, et al., 1978). Among the many hyperaccumulating Brassicaceae Ni hyperaccumulators, the carboxylic acid derivative malate seems to be the chief Ni ligand (Montargès-Pelletier, et al., 2008; Pelosi, et al., 1976). The South African hyperaccumulators *Berkheya coddii* and *S. coronatus* have been shown to have Ni complexed with citrate and malate (Montargès-Pelletier, et al., 2011). As with the aforementioned chelating agents and ligands, little evidence is available that shows their importance or dependence in hyperaccumulation, but their role however may be restricted to vacuolar sequestration owing to the more acidic environment of the vacuole. The relatively low association constant of citric acid, for example with Ni (5.4 lg K) (Martel & Smith, 1974), further illustrates that they may not necessarily play such important roles in hyperaccumulation.

Deng, et al. (2017) proposed that there are nine physiological processes which occur during Ni hyperaccumulation. These processes, which share similarity to Zn hyperaccumulation, are schematically represented in Figure 1-3 and can be summarised as follows: 1) the first process being the initial absorption at the root level. Here, Ni^{2+} is absorbed via a low-affinity transport system into the root symplast. The low-affinity transporters belong to a bigger group of transport proteins, the ZIP-family, which transport Zn and Fe, though this is speculative and has not been conclusively shown. 2) During the second process Ni could be chelated by organic compounds such as His, NA, and/or citric acid (Kerkeb & Krämer, 2003).

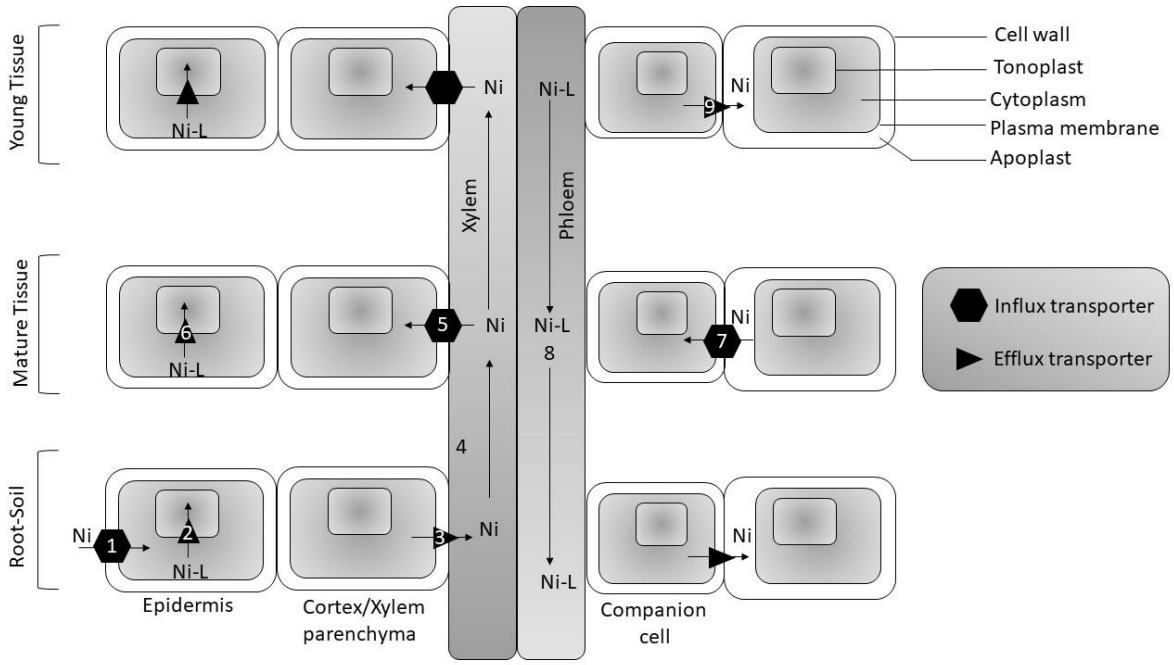


Figure 1-3: Schematic representation of Ni hyperaccumulation physiological processes 1-9*. Ni: free Ni; Ni-L: Ni-ligand complexes. Figure adapted from Deng, et al. (2017)*

3) Next, Ni is transported to xylem via xylem parenchyma. 4) Ni then moves towards the leaves where the majority is transported into mature leaves due to stronger transpirational pull. During the xylem transportation process, Ni is in a free hydrated cation form. 5) At the leaf vein interface, Ni first accumulates in the apoplast before being unloaded into the symplast, with some Ni remaining in the apoplast. 6) Once in the leaf, Ni is then sequestered into the vacuoles presumably by a transport protein. Typically, the vacuoles of epidermal cells are targetted for sequestration. 7) At the phloem interface, companion cells may express influx transporters which could absorb large amounts of Ni into the cytosol where Ni could be chelated once more. 8) From the companion cells, Ni-ligand complexes could be transferred to the phloem where they would passively be translocated to sink organs. 9) At sink sites, such as young leaves, Ni could be transported apoplastically from phloem. However, neither step 8 nor 9 has definitively been demonstrated to occur.

1.7 *Senecio coronatus*

Senecio coronatus is a eudicotyledon in the Asteraceae family. Common names for *S. coronatus* include: Sybossie, Indlebe Yebokwe, Ubulibazi, and Moremoholo, to name a few. The genus contains roughly 1 250 species and is one of the largest flowering genera with a global distribution. There are about 300 species which occur in South Africa. Seregin & Kozhevnikova, (2006) list *S. pauperculus* as a Ni-hyperaccumulator with a dry weight Ni content of 1900 $mg.kg^{-1}$, but its taxonomy is unclear as it is also described as *Packera*

paupercula. Although taxonomically they may not be in the same genus, *S. coronatus* and *P/S. paupercula* belong to the same family, Asteraceae. *S. coronatus* was first described in *Flora Capensis* Vol 3 in 1894 by W. H. Harvey and is mostly distributed in the North-eastern parts of South Africa. Recently, *S. conrathii* was identified as another Ni hyperaccumulator from the Barberton Greenstone belt, South Africa (Siebert, et al., 2017)

Whilst most research has focused on a hand-full of known hyperaccumulators, *S. coronatus* provides a novel means to investigate the phenomenon of hyperaccumulating variation owing to the fact that there is variation in shoot Ni concentrations irrespective of soil Ni content. Boyd, et al., (2008) reported two hyperaccumulating populations (8800 and 12 000 mg Ni kg^{-1} DW) and two non-accumulating populations (<130 mg Ni kg^{-1} DW) of *S. coronatus* from four serpentine sites (>1000 mg Ni kg^{-1} DW [Morrey, et al., 1989]) in the Barberton Greenstone Belt, South Africa. This observation makes *S. coronatus* a novel species to investigate the molecular and genetic mechanisms of Ni hyperaccumulation.

1.8 The Genes of Interest

Meier, et al., (2018) used RNA-sequencing to identify genes which were differentially expressed in hyperaccumulating versus non-hyperaccumulating populations of *S. coronatus*. From their results, 394 root and 399 shoot genes were identified as being overexpressed, of which 22 were putative homologues of transport proteins including proteins such as Nramp1 and MTP3. *IRT1* and *IREG2* had increased expression for all the *S. coronatus* populations, but both had much greater up regulation in shoots than roots for the hyperaccumulating populations with very low expression observed in the non-hyperaccumulating populations. They postulated that high expression of *IRT1* in the roots enhanced Ni uptake, of which a fraction is sequestered into root vacuoles by *IREG2* (Meier, et al., 2018). Furthermore, they suggested that the higher expression in the shoots may act as a driver for root-to-shoot transport (Meier, et al., 2018). Whilst up-or down-regulation of genes is merely an indication of which genes are most or least active respectively, it provide a starting point for potential candidate genes which, in this case, may play a role in the accumulation trait of *S.coronatus*. As such, it naturally follows that functional analysis of these genes should be performed.

1.8.1 The *IRT1* Gene

The *IRT* (**i**ron-**r**egulated **t**ransporter) gene belongs to the ZIP family of transport genes of which three *IRT* genes have been described or discussed in literature. *IRT1* is a plasma membrane protein which is predominantly expressed in root epidermal cells (Vert, et al., 2002; Eide, et al., 1996) and has been shown to play a critical role during Fe-deficiency response (Nishida, et al., 2011; Vert, et al., 2002; Eide, et al., 1996). *IRT1* assists in the acquisition of Fe^{2+} during deficiency responses, but as a consequence of its low specificity (Vert, et al., 2001), it also takes up other divalent cations such as Ni^{2+}, Cd^{2+}, Co^{2+}, and Mn^{2+}. *IRT2* by

contrast, has a much stronger substrate affinity and does not appear to transport heavy metals as readily as does *IRT1* (Vert, et al., 2001). Furthermore, Vert, et al., (2001) suggested that *IRT2* is more likely to be involved in 'housekeeping' during Fe acquisition, whereas *IRT1* plays a greater role in 'scavenging' during Fe starvation.

Results from Nishida et al., (2011), who employed the use of an *irt1 A. thaliana mutant* grown on Fe-sufficient and Fe-deficient media can be summarised as follows: 1) *IRT1* is a pathway for Ni uptake; 2) Ni^{2+} competes with Fe^{2+} for uptake; 3) up-regulation of *IRT1* resulted in reduced growth due to increased Ni uptake which led to a decrease in shoot Fe despite enhancing Fe uptake; 4) loss of *IRT1* results in less Ni being accumulated; and 5) quantitative RT-PCR revealed that Ni exposure induces *IRT1* expression (possibly indirectly due to competition with Fe). The results from Nishida et al., (2011) clearly demonstrate that *IRT1* is directly involved in the accumulation of heavy metals and that exposure to heavy metals under Fe starvation conditions further enhances it. Whether *IRT1* is required for hyperaccumulation to occur remains unclear, but the fact that *IRT1* has been shown to be overexpressed in hyperaccumulators (van de Mortel, et al., 2006) suggests that it may play a role.

1.8.2 The *IREG2* Gene

IREG2 (IRON REGULATED), investigated by Schaaf, et al., (2006) as a potential component in metal homeostasis as a means to avoid Ni toxicity in Fe-deficiency response in *A. thaliana*, shares 34.8% similarity to *IREG1*, found on mammalian doudenal enterocyte membranes. *IREG1* mediates Fe transport and is expressed in a variety of cells in which Fe export seems to be a prominent function (Schaaf, et al., 2006). McKie, et al., (2000) demonstrated that the mammalian *IREG1* is a doudenal Fe export protein which mediates Fe efflux under conditions such as Fe deficiency which led Schaaf, at al., (2006) to consider *IREG2* as being Fe-regulated. The IREG1 or ferroportin (FPN) 1 localises to the plasma membrane and is predominantly expressed in the stele (Morrissey, et al., 2009) and is thought to be involved in vascular loading during iron deficiency. Its role, however, was investigated within the context of Co, but it is possible that it may also be involved in Ni vascular loading. The results from Schaaf, et al., (2006) can be summarised as follows: 1) Fe-uptake defective yeast mutants grown on 4 – 10 µM Fe showed little growth differences when compared to the control; 2) *AtIREG2* expression in heavy metal sensitive yeast strains such as *cot1* conferred yeast growth on high Ni concentrations; 3) *IREG2* localises to the tonoplast and is pH dependent (proton dependent export); 4) overexpression of *IREG2* in transgenic plants enhanced their tolerance towards Ni; and 5) Ni uptake rates are related to Fe-deficiency.

Based on these results, Schaaf, et al., (2006) hypothesised that Fe^{2+} is not a suitable sustrate for *AtIREG2*, unlike its close mammalian relative *IREG1*. Of the observed radio-labelled metals which were being transported (^{65}Zn, ^{63}Ni, and ^{59}Fe), only ^{63}Ni accumulation was significantly

altered (Schaaf, et al., 2006). The localisation at the tonoplast suggests that *AtIREG2* serves a detoxification function by vacuolar sequestration of Ni. By means of cross-species RNA-seq comparisons, (de la Torre, et al., 2018) identified IREG/Ferroporin transporters as the 'most robust up-regulated function among distantly related Ni hyperaccumulators'. They found that within the Brassicaceae, Rubiaceae and Euphorbiaceae families (which also represent the families with the most Ni hyperaccumulators) IREG/Ferroporin transporters had a 4 – 800-fold increased expression when compared to their non-accumulator relatives (de la Torre, et al., 2018). Furthermore, New Caledonian members had significantly greater expression in both hyperaccumulator and non-hyperaccumulator representatives from Salicaceae and Cunoniaceae families (de la Torre, et al., 2018). By silencing the endogenous expression of *NcIREG2* from the hyperaccumulator *N. caerulescens* by means of artificial miRNA, they were able to show that the amplitude of Ni accumulation had a strong correlation with *NcIREG2* expression. Moreover, cellular location studies revealed that NcIREG2 also localises to the tonoplast. Merlot, et al., (2014) identified *PgIREG1* from the hyperaccumulator *Psychorita gabriellae* through next generation sequencing as a candidate gene potentially involved in accumulation. Similarly to AtIREG2 from Schaaf, et al., (2006), PgIREG1 localises to the tonoplast and was shown to be involved in Ni tolerance , as evidenced the *ireg2-l Arabidopsis* mutant displaying restored Ni accumulation when transformed with *PgIREG1.* Furthermore, constitutive over expressors of *PgIREG1* in *ireg2-l* transgenic lines accumulate more Ni than wildtype plants and suggests that overexpression of *PgIREG1* could increase Ni storage (Merlot, et al., 2014).

When Fe is limiting, *IRT1* is over-expressed to increase the uptake of Fe. Its low specificity therefore results in the additional accumulation of other cations such as Ni. Ni then further enhances *IRT1* expression (Nishida, et al., 2011) but does not induce *IREG2* expression (Schaaf, et al., 2006). The Fe-deficiency state induces up-regulation of *IREG2* controlled by *FRU/FIT1* transcription factors (Schaaf, et al., 2006), which also controls the transcription of *IRT1. IREG2* therefore sequesters excess Ni into the vacuoles as a means to 'counterbalance' the low specificity of *IRT1* (Schaaf, et al., 2006).

1.9 Research Aims and Objectives

Given the in-depth studies of the roles of *IRT1* (Nishida, et al., 2011) and *IREG2* (Schaaf, et al., 2006), it is hypothesised that homologs identified in *S. coronatus* may play a role in the uptake and vacuolar sequestration of Ni respectively. The aim therefore is to determine whether these genes encode functional homologues of the *Arabidopsis* transporters *AtIRT1* and *AtIREG2*. To do so, *A. thaliana* mutants defective in either *IRT1* or *IREG2* will be transformed with expression constructs containing the *S. coronatus* protein coding sequence under the control of the corresponding *Arabidopsis* promoter i.e. pAtIRT1:ScIRT1 or pAtIREG2:ScIREG2. The project thus has three main aims:

1. Generation of expression constructs by Gateway cloning

Arabidopsis thaliana IRT1 and *IREG2* promoters (*AtIRT1* and *AtIREG2* respectively) and *Senecio coronatus IRT1* and *IREG2* genes (*ScIRT1* and *ScIREG2* respectively) will first be cloned into Gateway compatible DONR vectors before being recombined to create two separate two fragment expression clones in an appropriate destination vector. These will then be used to transform mutant knock-out *irt1* and *ireg2* plants.

2. Isolation of homozygous T-DNA insertion lines for IRT1 *and* IREG2

T-DNA insertion lines for *AtIRT1* and *AtIREG2* will be obtained and homozygous individuals identified by PCR-based genotyping. The position of the insert will be confirmed by DNA sequencing, and RT-PCR used to establish whether the T-DNA insertion indeed results in knock out of target gene expression

3. Establishment of phenotypic assays to determine whether S. coronatus *proteins complement* Arabidopsis *mutants*

Growth assays will be developed to identify robust metal-responsive phenotypes in order to develop an effective assay to assess the degree of complementation.

MATERIALS AND METHODS

2.1 Genotyping and Phenotyping of T-DNA insertion lines

2.1.1 Plant growth

Arabidopsis thaliana seeds were acquired from the SALK institute (Table 2-1) and were grown on a soil mix containing equal amounts of Sphagnum peat (Jiffy-7®) and vermiculite. Plants were grown in a growth room (22°C, 16 h light, 8 h dark, 55% relative humidity) with watering occurring every two to three days a week. During the growing period, plants were supplemented periodically with phostrogen fertilizer (1.4 $g.L^{-1}$).

***Table 2-1**: SALK seed codes with their respective mutant type and collection*

Seed code	Mutant locus	Collection status from SALK	Published characterisation of SALK line
SALK 097869C	*irt1*	Homozygous	None available
SALK 054554C	*irt1*	Segregating line	(Nishida, et al., 2011)
SALK 074442C	*ireg2*	Homozygous	(Schaaf, et al., 2006)
SALK 12071C	*ireg2*	Homozygous	(Schaaf, et al., 2006)
SALK 024525	*irt1*	Segregating line	(Nishida, et al., 2011)

2.1.2 Genotyping of T-DNA insertion lines

2.1.2.1 DNA-based genotyping

To confirm the homozygosity of mutant lines, leaf discs were taken from plants and genomic DNA was isolated using a modified version Edwards, et al., (1991) DNA isolation protocol described below . SALK and gene specific primers, listed in Table 2-2, were created using SALK Institute Genomic Analysis Laboratory's T-DNA Primer Design software (http://signal.salk.edu/tdnaprimers.2.html). Additional primers were also created using SnapGene® Viewer 4.3.4. Primers for genotyping were supplied by *Inqaba Biotech*. Homozygosity was confirmed by running various PCR reactions and then electrophoresing the PCR products on a 1% (w/v) agarose gel supplemented with ethidium bromide (0.2 – 0.5 $\mu g.mL^{-1}$ depending on application). For one set of reactions the T-DNA left border primer (BP) and the gene-specific reverse primer (RP) were used as it would yield a band if the mutant was homozygous, whilst the gene-specific forward (LP) and reverse primers (RP) were used for the second reaction to determine the absence of the allele of interest in homozygous lines and would yield a band in the wildtype Columbia-0 control. The reaction is schematically shown in Figure 2-1 below.

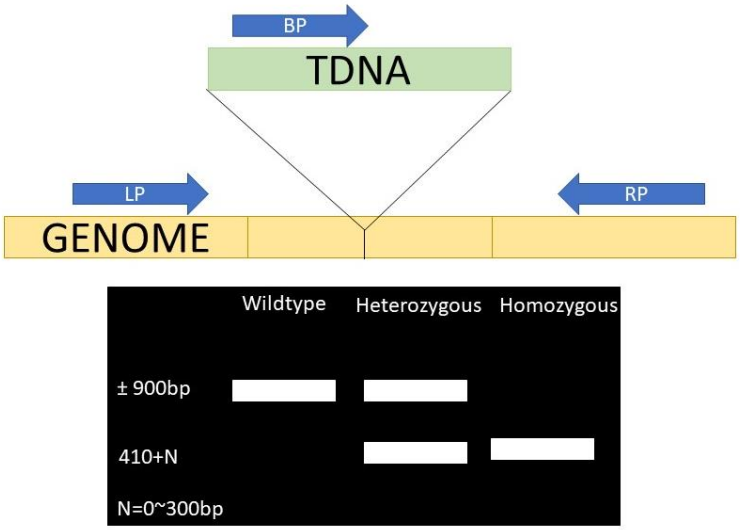


Figure 2-1: Schematic representation of genotyping PCRS*: TDNA is inserted within the genome at the gene of interest and disrupts that gene thereby generating a knock-out mutant. Primers are used to verify homozygosity of mutant knock-out lines. A PCR reaction using the LP and RP primers, which are gene specific, yield a wildtype band (±900 bp) in non-mutant plants but would fail to produce a band in full knock-out lines. A PCR using the BP and RP primers would produce a band (410 +N bp) if there is a T-DNA insertion for either homozygous or heterozygous lines.*

Table 2-2*: Genotyping primers with sequences used to verify homozygosity of knockout mutant seed lines*

Mutant	SALK Line	Primer name	Sequence (5'→3')
ireg2	074442C	Gene specific forward*	GGTACCTGATCATCGTTGACG
		Gene specific reverse*	GATTTGATCAGACCTTGCACC
	12071C	Gene specific forward	AAGCTGTCACGTGTGATGTTG
		Gene specific reverse	ATAAAACCGGTTCCAAGGATG
irt1	097869C	Gene specific forward	ATACACTGAGCACAAGCTGTCA
		Gene specific reverse	ATAAAACCGGTTCCAAGGATG
	054554C	Gene specific forward	CAAGTAAAACCAGCGCTTTTG
		Gene specific reverse	GCTCTTTGCTTCCATCAAATG
Both		Left border T-DNA region	ATTTTGCCGATTTCGGAAC

* Gene specific forward primer also known as left primer and gene specific reverse also known as right border primer

Cycling parameters for genotyping PCRs were as follows: initial denaturation at 95°C for three mins, followed by 30 cycles of 95°C for 20 sec, 55°C for 15 sec for *IREG2* and 50°C for *IRT1*, 72°C for 1 min, and a final extension at 72°C for 1 min.

2.1.2.2 Gene expression of homozygous T-DNA insertion lines

For RNA extraction, Col-0, *irt1* knockout mutants, and *ireg2* knockout mutants were grown hydroponically in plant nutrient media (PN) comprising 2.5 mM KNO_3; 1 mM $Ca(NO_3)_2.4H_2O$; 1 mM $MgSO_4.7H_2O$; 1.25 mM KPO_4 in 2L of MiliQ-H_2O and a 1000x micronutrient stock comprising 70 mM H_3BO_3; 14 mM $MnSO_4.H_2O$; 0.5 mM $CuSO_4.5H_2O$; 1 mM $ZnSO_4.6H_2O$; 0.2 mM Na_2MoO_4; 10 mM NaCl and 0.01 mM $CoCl_2.6H_2O$. These plants were grown in the presence of 100 µM Fe-Na-EDTA for two weeks with the nutrient media being changed weekly. Following this, new media deficient in Fe-Na-EDTA was used to induce expression of *IRT1* and *IREG2* and plants were allowed to grow for a further ten days. Total RNA from mutant and wildtype seedlings was extracted from root tissue using an inhouse protocol described below from which cDNA was synthesised. Expression from cDNA was examined by means of RT-PCR using the cDNA primers as listed in Table 2-3.

***Table 2-3**: RT-PCR primers with sequences used to assess gene expression of identified homozygous mutants*

Gene	Primer name	Sequence (5'→3')
IREG2	*A. thaliana IREG2* forward	AGACTCTGGCTTGTAACTCAA
	A. thaliana IREG2 everse	GATTACCGCAAGATCGAACATC
IRT1	*A. thaliana IRT1* forward	ATGGCTTCAAATTCAGCACTTCTC
	A. thaliana IRT1 reverse	AAGCCCATTTGGCGATAATCGA

Mutant plants would have no expression if they are true homozygous knockouts for their respective genes, whereas the wildtype would be expected to show induction of *IRT1* (Vert, et al., 2002) and *IREG2* (Schaaf, et al., 2006) expression under Fe-starvation conditions. Prior to commencing with RT-PCRs, the newly synthesised cDNA was diluted (1:20) and used as template in a PCR reaction using the following primers (SAND Forward: 5'-CAGACAAGGCGATGGCGATA-3' and SAND Reverse: 5'-CTTTCTCTCAAGGGTTT CTGGGT-3' for At2G28390.1) to assess quality of cDNA (initial denaturation at 94°C for 5 min; followed by 30 cycles of 95°C for 15 sec; 55°C for 30 sec; 72°C for 30 sec; and final extension at 72°C for 10 min).

2.1.3 Phenotyping of T-DNA insertion lines

To determine whether mutant homozygous plant lines displayed characteristic phenotypes, a hydroponics-based Fe-deficiency assay was established. *A. thaliana* Col-0, *irt1*, and *ireg2* seeds were grown hydroponically in PN media for two weeks in the presence of 100 µM Fe-Na-EDTA to ensure successful germination and establishment of knock-out mutants. Two weeks later, the hydroponic boxes were then supplemented with or without 10 µM $NiCl_2$ with or without 100 µM Fe-Na-EDTA. The plants were then allowed to grow for 10 days after which plants were removed and root length, root biomass, and shoot biomass was recorded. Root and shoot tissue from the phenotype assays were also suspended in a solution of dimethylglyoxime (DMG) as a visual representation for the presence of Ni. Additionally, for *irt1* phenotyping, *irt1* mutants and Col-0 seeds were plated on PN phytoagar plates supplemented with Fe at 100 µM and low Fe at 10 µM. These were then allowed to grow for two weeks after which relative survival (regarded as the presence of true leaves) was determined.

2.2 Polymerase Chain Reactions

Primers (*Table 2-4*) were designed using SnapGene® Viewer 4.3.4 and supplied by *Inqaba Biotech*. For all amplicons to be used for cloning, *2x KAPA HiFi HotStart ReadyMix* was used. Primers used for Gateway® cloning were HPLC purified. Cycling parameters varied for each amplicon and will be discussed separately. PCR products were purified either using *Wizard SV Gel and PCR Clean-Up System* (Promega Corporation) as per the manufacturer's protocol or PEG purification as per Gateway® user manual.

***Table 2-4**: Primers for PCR amplification of products for downstream cloning*

Gene	Primer Name	Primer Sequence (5' → 3')	Amplicon Size (bp)
ScIRT1	ATTB5	GGGGACAACTTTGTATACAAAAGTTGTAAAAATGGCTTCAAGTTCAAAAAATGTC	1 108
	ATTB2	GGGGACCACTTTGTACAAGAAAGCTGGGTTTTAAGCCCATTTTG	
AtIRT1	ATTB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTACATTAAGTCAATAACCCCATAATC	1 150
	ATTB5R	GGGGACAACTTTTGTATACAAAGTTGTAGATTGTTTAATGTTTGTGTGAC	
ScIREG2	ATTB5	GGGGACAACTTTGTATACAAAAGTTGTAAAAATGGAGGAGGGGTATTATT	1 561
	ATTB2	GGGGACCACTTTGTACAAGAAAGCTGGGTTTTAAAGAAACAGGGCAAGC	
AtIREG2	ATTB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAGAGAGTAGTCGGAATGTTCTGG	1 834
	ATTB5R	GGGGACAACTTTTGTATACAAAGTTGTCTTCTGACTACTTTGATTCTTTCCTC	

Amplicons were verified by means of a 1 X TAE 1% agarose gel electrophoresis supplemented with ethidium bromide.

2.2.1 Amplification of *ScIRT1* and *ScIREG2* for Gateway® BP®-Cloning

Senecio coronatus cDNA was used as template DNA for the amplification of the protein coding regions of *ScIRT1* and *ScIREG2* using the *att*B5 and *att*B2 primers sets as listed in Table 2-4 above. For Gateway® BP®-cloning, *att*B5 sites were added to the 5' ends and *att*B2 sites were added to the 3' ends of the amplicons. Cycling parameters were as follows: initial 98˚C for 3

min followed by 30 cycles of amplification of 98˚C for 20 sec, 62˚C for 15 sec, 72˚C for 72 sec, and a final extension at 72˚C for 72 sec.

2.2.2 Amplification of *Arabidopsis thaliana AtIRT1* and At*IREG2* promoters for Gateway® BP®-Cloning

Arabidopsis thaliana, ecotype Columbia (Col-0), DNA was used as template DNA for the amplification of the *A. thaliana* promoters from *AtIRT1* and *AtIREG2*. DNA was extracted using a modified version of the CTAB extraction protocol described below. For Gateway® BP-cloning, *att*B1 sites were added to the 5' ends and *att*B5R sites were added to the 3' ends of the amplicons. Given AT-richness of gene regulatory regions, a two-step PCR was followed as suggested by Dhatterwal, et al., (2017). Cycling parameters were as follows: initial denaturation at 98˚C for 3 min; 35 cycles of amplification of 98˚C for 20 sec and 65˚C for 3 min, and a final extension at 65˚C for 7 min.

2.3 Gateway® cloning for complementation

For the complementation part of the project, Gateway® cloning technology was employed to construct expression vectors which have the *Arabidopsis thaliana* promoter driving the expression of *Senecio coronatus* protein coding regions of the orthologue, for example *AtIRT1* promoter driving the *ScIRT1* protein expression. The process is schematically represented below in Figure 2-2. First, a BP-Clonase® II recombination reaction occurs whereby the individual PCR fragments, which have flanking attB sites, are recombined into a donor vector to create an entry clone. The *attB* sites can only recombine with corresponding *attP* sites on the donor vector and subsequently produce attL sites. Following successful recombination of the PCR fragment with the donor vector, a *ccdB* containing by-product is produced. Subsequently, an LR-Clonase® II recombination reaction is used whereby two donor vectors recombine to create the expression clone. For the LR-Clonase® II reaction, *attL* sites of the donor clone recombine with *attR* sites on the destination vector to create an *attB*-containing expression clone and an *attP-ccdB*-containing by-product.

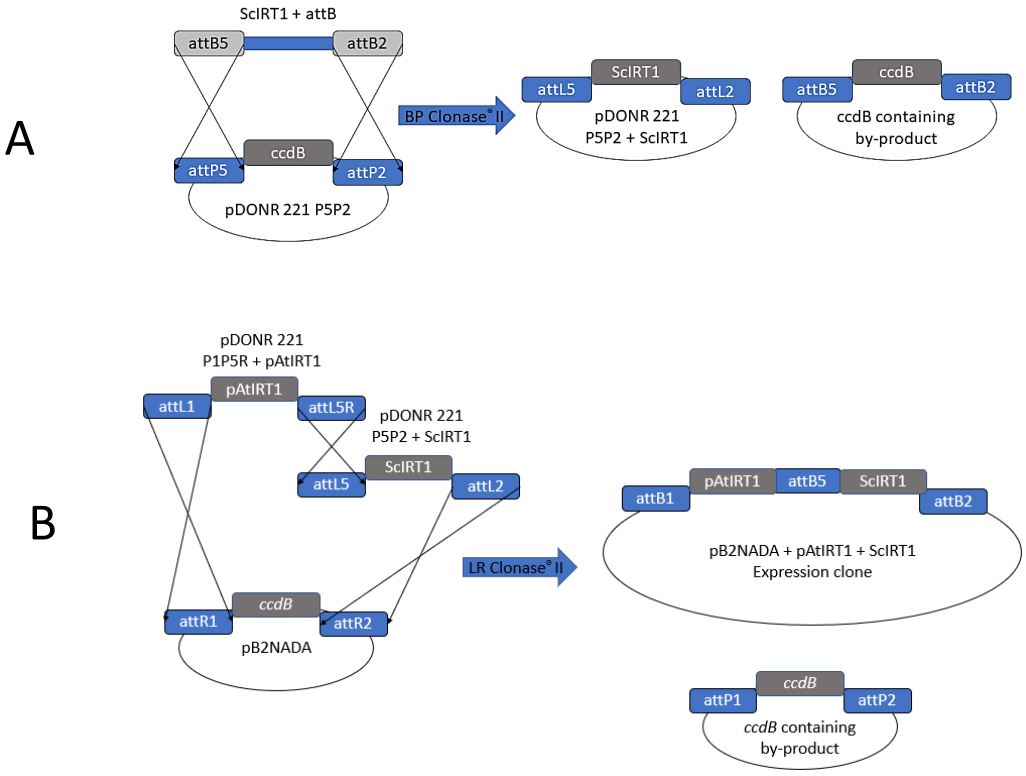


Figure 2-2: Schematic representation of Gateway® BP-Clonase® II (A) and LR-Clonase® II (B) *for two fragment cloning. Schematic shows the BP-Clonase® II and LR-Clonase® II reactions for the generation of the expression clone containing the AtIRT1 and ScIRT1 amplicons.*

2.3.1 BP-Clonase ® II recombination

As per the Multisite Gateway® Pro cloning protocol, BP-Clonase® II recombination reaction was performed with the *att*B-flanking PCR products and a donor vector, pDONR™221. The *A. thaliana* promoters, flanked by *att*B1 and *att*B5r, were recombined with pDONR™221 P1-P5r whereas the *S. coronatus* inserts, flanked by *att*B5 and *att*B2, were inserted into pDONR™221 P5-P2. The BP-Clonase® II reactions were done overnight at 25˚C and terminated using proteinase K. A total of 4 µL of the BP-Clonase® II reaction was transformed into 50 µL chemically competent *Escherichia coli* DH5α cells. Cloned products were screened on LA (1% [w/v] tryptone; 0.5% [w/v] yeast extract; 0.5% [w/v] NaCl; 1.5% [w/v] bacterioagar) kanamycin plates (50 mg.mL^{-1}) and subjected to colony PCR using gene specific *att*B-containing primers. Following successful colony PCR transformants, plasmid DNA was isolated as described below. Restriction digests and PCRs on plasmids were performed to validate successful clones before sending clones to Stellenbosch University's Central Analytical Facility (SUNCAF) for sequencing using their M13 forward and M13 reverse primers. Sequences were aligned using BLAST on NCBI's website (Zhang, et al., 2000).

2.3.2 LR-Clonase ® II recombination

Following confirmation of successful BP-Clonase® II clones, the BP-Clonase® II clones were then subjected to a LR-Clonase® II recombination reactions. The viability of various destination vectors was tested to assess their efficacy. Of the three vectors tested namely pB7WG, pEARLYGATE302, and pB2WG7.0, the latter was selected on the basis of its little to no background colony production when transformed into competent *E. coli* DH5α cells, in contrast to the other two destination vectors tested. The vector, however, needed to be modified as it contained a CaMV35S promoter which needed to be removed. This was accomplished by double digestion using *SacI* (New England BioLabs) and *SpeI* (New England BioLabs). *SpeI* produces blunt ends whereas *SacI* produces sticky ends which were blunted with T4 DNA polymerase as per the manufacturer's recommendations. The digested backbone was gel excised and purified using Wizard SV Gel and PCR Clean-Up System (Promega Corporation) as per the manufacturer's protocol. The blunt ends were then ligated together using T4 DNA ligase (New England BioLabs) overnight at 16˚C. The ligation was then transformed into One Shot™ TOP 10 chemically competent *E. coli* cells and plated on 100 mg.mL^{-1} spectinomycin LA plates. Colonies were initially screened using a primer upstream of the *SacI* site (N35S: 5'-GTCAGTTCCAAACGTAAAACG-3') and a primer found within the CaMV 35S terminator sequence (T35 S 5'- ACTGGATTTTGGTTTTAGGAATTAGA-3').

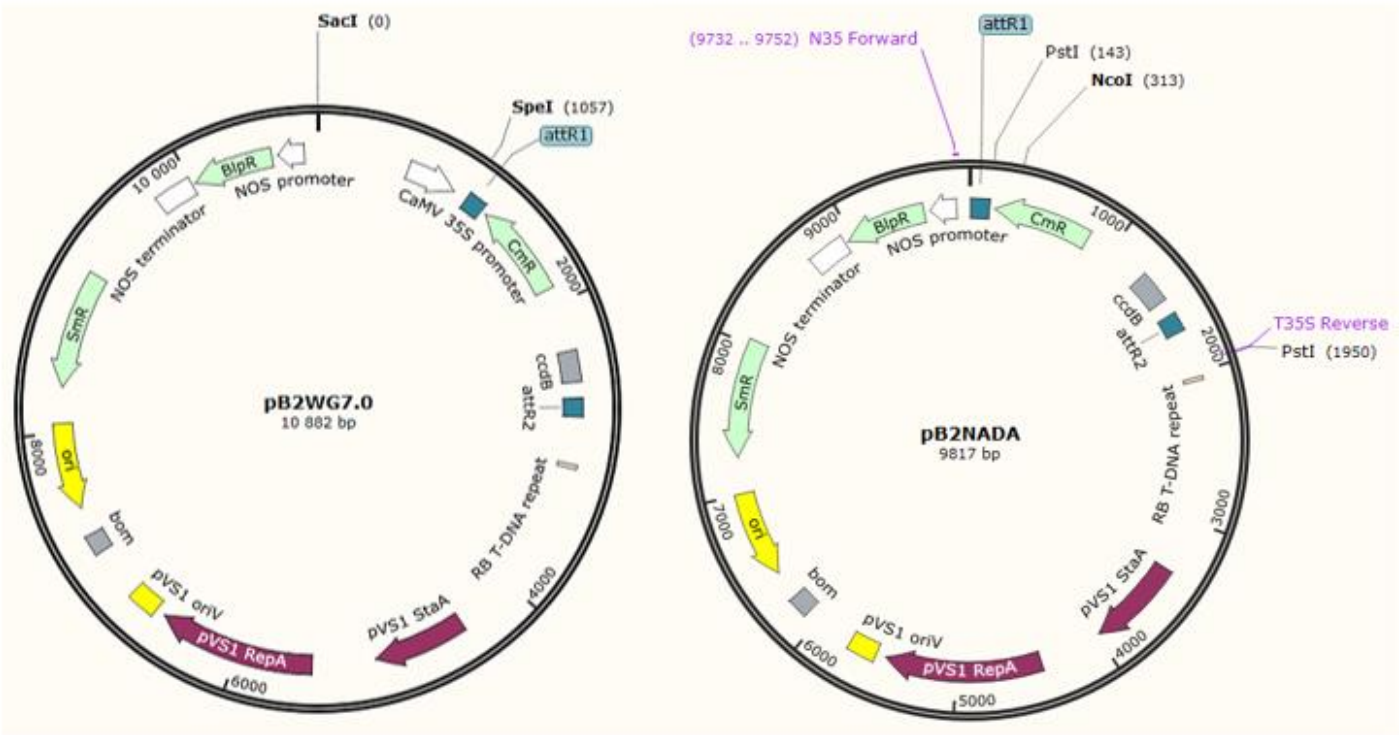


Figure 2-3: Vector map for pB2WG7 and pB2NADA *showing important restriction sites. SacI and SpeI were used to remove the CaMV 35S Promoter. NcoI and PstI were used to confirm removal of the promoter.*

Those colonies that did not yield a PCR amplicon in this reaction were selected for plasmid extraction as described above. Recirculation of pB2WG7 was verified by digesting with *NcoI* (New England BioLabs) and *PstI* (Fermentas). This vector was renamed to pB2NADA and sent for sequencing to validate full removal of CaMV 35S promoter using the upstream forward primer above. Final size for pB2NADA is 9817 bp.

The *att*P1 site of the *A. thaliana* promoter was recombined with an *att*L1 site on the pB2NADA vector whereas the *att*P5r of the *A. thaliana* promoter and the *att*P5 site of the *S. coronatus* gene were then recombined and the *att*P2 site of the *S. coronatus* gene was recombined with an *att*L2 site on the pB2NADA expression vector. This is schematically represented in Figure 2-2 above. The same screening method was employed as above using a combination of primers for the colony PCR. Original *att*B containing primers were used to verify insertion of the individual inserts as well as a combination of the *AtIRT1 att*B5r and *ScIRT1 att*B2, and *AtIREG2 att*B5r and *ScIREG2 att*B2 primers were used to verify the full-length insertion.

2.4 Nucleic acid extraction and manipulation

Routinely used techniques used throughout the project are briefly described below.

2.4.1 DNA Extractions

Modified Edwards, et al., (1991) DNA extraction: Leaf discs of approximately 1 cm^2 in size was gently ground in 400 µL of extraction buffer (200 mM Tris pH 7.5; 250 mM NaCl; 25 mM EDTA; 0.5% [w/v] SDS) and incubated at 60˚C for 10 min. Following incubation, 400 µL of chloroform: isoamylalcohol (24:1) was added and centrifuged for ten mins at 6,160 x *g*. The aqueous layer was transferred to a clean tube to which 2x volume of cold 100% ethanol and 1/10 volume of 3 M sodium acetate (pH5.2) was added and then incubated on ice for ten mins. The DNA was pelleted by centrifuging for 10 min at 6,160 x *g*, after which the supernatant was discarded. Two washes of 70% (v/v) ethanol were done. The pellet was airdried in a fume hood to remove excess ethanol. The pellet was then resuspended in 1x TE buffer (10 mM Tris-HCl pH7.5; 1 mM EDTA).

Modified CTAB DNA Extraction: Roughly 1g of leaf tissue was gently ground using 450 µL TES buffer (0.1 M Tris pH 8.0; 0.01 M EDTA; 2% w/v SDS) to which 25 µL proteinkinase K (Invitrogen) was added and incubated at 60˚C for 1h. Following incubation, 160 µL 5 M NaCl and 70 µL 10% (w/v) CTAB (hexadecyltrimethylammonium bromide) was added and incubated at 65˚C for 1 h. A total of 780 µL of chloroform: isoamylalcohol (24:1) was added and samples incubated on ice for 30 min. The tubes were then centrifuged at 12,074 x *g* for 10 min. The top aqueous layer was transferred to a new tube to which 345 µL isopropanol

was added and centrifuged for five mins. The supernatant was discarded, and the pellet was washed twice with 70% (v/v) ethanol via a 3 min centrifugation at the same speed as above. The pellet was airdried in a fume hood and resuspended in 50 µL 1x TE buffer to which 5 µL RNase (1mg.ml^{-1}) was added and incubated at 37˚C for 1 h.

2.4.2 RNA Extractions and cDNA synthesis

RNA Extraction: Root tissue was dabbed dry and flash frozen in microcentrifuge tubes using liquid nitrogen and kept frozen at -80˚C. Up to three steel ball bearings (precleaned with chloroform) were added to each tube along with 1 mL Trizol (ammonium thiocyanate; guanidium thiocyanate; sodium acetate; glycerol; and phenol) and were then homogenised in a paint shaker for 4 min, allowed to rest for 30 sec and then shaken again for a further 3 min. These were then centrifuged at 10,410 x *g* at 4˚C for 10 min after which 750 µL of the supernatant was transferred into a tube containing 200 µL chloroform and vortexed for 30 sec, incubated at room temperature for 5 min and then centrifuged at 8,870 x *g* for 15 min at 4˚C. Approximately 650 µL of the supernatant was transferred into a tube containing a mixture of 250 µL isopropanol and 250 µL 1.2 M NaCl and 0.8 M sodium citrate and gently inverted three to four times. The tubes were left at room temperature for 10 min after which they were centrifuged for 15 min at 8,870 x *g* at 4˚C. The supernatant was removed, and the pellet was washed twice with 75% ethanol via five-min centrifugation at 8,870 x *g* at 4˚C. The pellet was airdried in a fume hood to evaporate any excess ethanol. The pellet was gently resuspended in 25 µL DEPC-treated water and placed in a heating block at 55˚C for 10 min. Total RNA was analysed by electrophoresing RNA in a 1x TAE 0.8% agarose gel supplemented with ethidium bromide. Prior to cDNA synthesis, RNA samples were DNase treated as follows. A total of 2 µg of RNA was added to 0.1 volume of 10x Turbo DNase buffer (Ambion) and 0.5 µL of Turbo DNase (Ambion) and incubated at 37°C for 30 min. After incubation, 2 µL of DNase inactivator (Invitrogen) was added and incubated at room temperature for 5 min after which it was centrifuged at 6,160 x *g* for 2 min and the supernatant was transferred to a new tube.

cDNA Synthesis: A total of 1 µg of total DNase-treated RNA was added to a microfuge tube containing 1 µL of 50 µM oligo dT 20, 1 µL 10 mM dNTPs (New England BioLabs) and the volume adjusted to 13 µL with nuclease-free water. This tube was incubated at 65 °C for 5 min and then chilled on ice for 1 min. To this, 4 µL 5x first-strand buffer (Invitrogen), 1 µL 0.1 M DTT, and 0.5 µL SuperScript III Reverse transcriptase (Invitrogen) were added and incubated for 5 min at 25°C, followed by a further 1 h incubation at 50°C. The reaction was terminated by incubating at 70°C for 15 min.

2.4.3 Preparation of competent bacteria, bacterial transformation, and plasmid isolation

Preparation of competent *E. coli* cells: *E. coli* DH5α required for competent preparation were streaked onto LA plates containing no antibiotic selection. A single colony was inoculated into 5 mL lysogeny broth (LB) (1% [w/v] tryptone; 0.5% [w/v] yeast extract; 0.5% [w/v] NaCl) and allowed to grow overnight at 37°C with continuous shaking. Two mL of the overnight culture was then inoculated into 250 mL LB supplemented with 20 mM MgSO4 and incubated at 37°C with continuous shaking until the OD600 measured 0.4 – 0.6. The culture was then centrifuged at 5000 x *g* for 5 min at 4°C and the supernatant was discarded. The cells were resuspended in 100 mL *TFB1* buffer (30 mM potassium acetate; 100 mM RbCl; 15% (v/v) glycerol; 10 mM $CaCl_2$; 20 mM $MnCl_2$; pH 5.8 with glacial acetic acid and filter sterilized) and incubated on ice for 5 min and then centrifuged at 5000 x *g* for 5 min at 4°C and the supernatant was discarded. The pellet was resuspended in 10 mL ice cold *TFB2* buffer (10 mM MOPS; 75 mM $CaCl_2$; 10 mM RbCl; and 15% (v/v) glycerol and filter sterilized) and 100 µL aliquots were made. Aliquots were flash-frozen with liquid N and stored at -20°C.

Bacterial Transformation: A heat shock method (30 min on ice; 45 seconds at 45˚C; 2 min on ice) was used to transform bacteria. A total of 50 µL cells was added to 1-10 ng of plasmid DNA and then filled to 1 mL with LB and shaken at 37˚C for 90 mins. The cells were then spun down at 1,540 x *g* and the supernatant removed to a leave a final volume of 100 µL. The cells were then resuspended of which 10 µL and 90 µL was plated on lysogeny agar plates (1.5% [w/v] bacterioagar) (LA) with an appropriate antibiotic selection.

Plasmid Isolation: Bacterial colonies which showed positive results from colony PCRs, were selected for overnight incubation at 37˚C with continuous shaking in LB with appropriate antibiotic selection. The plasmids were then extracted using the Zippy Miniplasmid (Zymo Research)extraction kit as per the manufacturer's guidelines with the following changes. Up to 4 mL of bacterial culture was spun down at 12,074 x *g* and resuspended in nuclease-free water. Following the initial column centrifugation, the flow through was added to the column again before being discarded. For large, low copy number plasmids, the elution buffer was heated to 55°C and then added to the column which was incubated for 10 min before elution. In order to increase the concentration of plasmids, no more than 30 µL of elution buffer was used. For large, low copy number plasmids, this volume was increased to 50 µL as per the manufacturer's recommendations.

RESULTS

3.1 Genotyping of T-DNA insertion lines

3.1.1 SALK_074442C and SALK_097869C are homozygous T-DNA insertion lines

PCR genotyping confirmed that SALK 074442C is a homozygous T-DNA insertion line for *IREG2*, this is depicted in Figure 3-1 A below. Presence of a band at roughly 500 bp for the T-DNA border and gene specific reverse primers confirms T-DNA insertion in all five individuals. Absence of amplification in the PCR reaction with the gene specific forward and reverse primers indicated that the wild-type *IREG2* allele could not be amplified from SALK 074442C plants, while Col-0 yielded the expected 1 kb product.

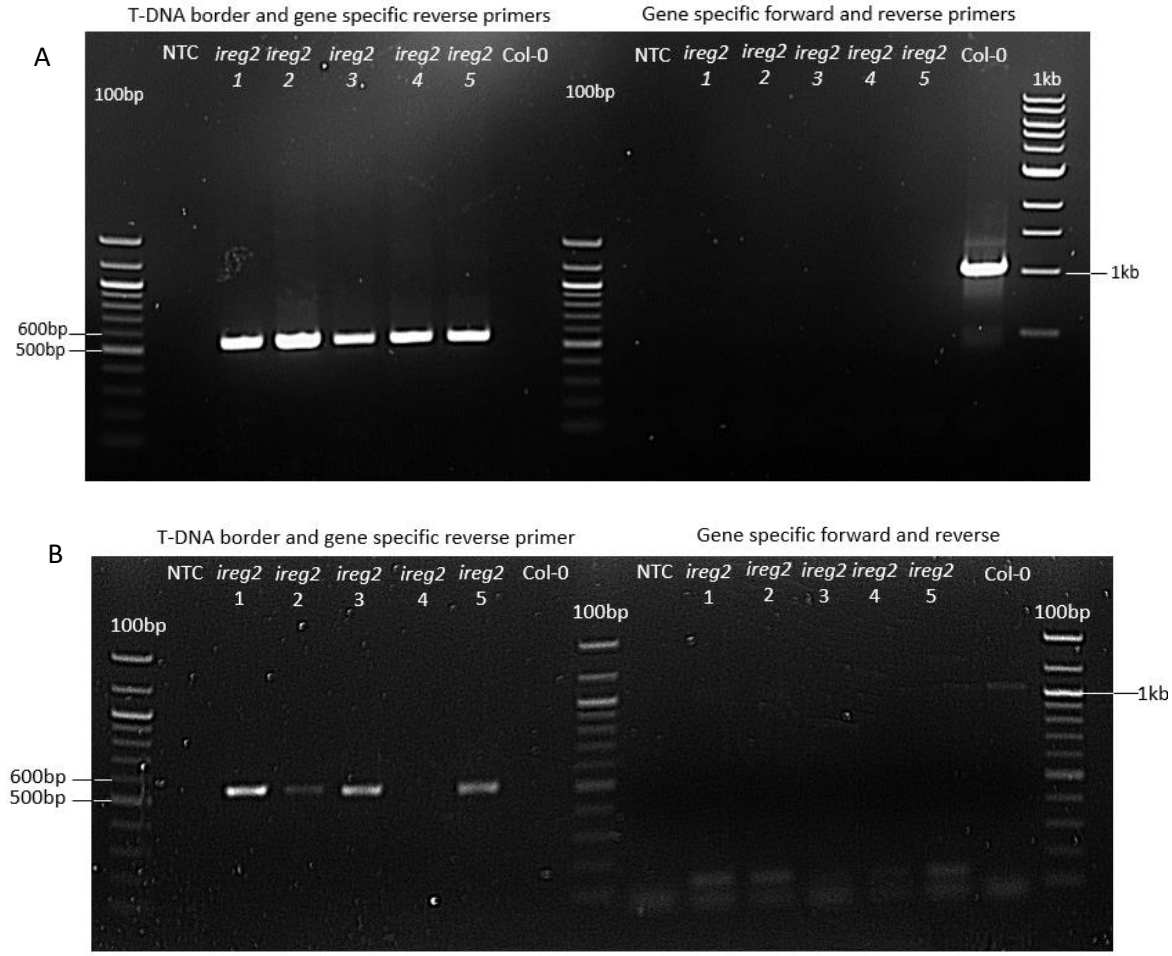


Figure 3-1: ireg2 Genotyping*: Agarose gel electrophoresis showing results from PCR genotyping on* ireg2 *SALK 074442C mutants and wildtype (A) and SALK 127071C and wildtype (B). Molecular weight ladders used were Quick-load® 100 bp DNA ladder (New England BioLabs) and Quick-Load® 1 kb DNA ladder (New England BioLabs)*

The second *ireg2* line, SALK 12071C (Figure 3-1 B) could not be confirmed as being homozygous due to the presence of a faint wildtype band for *ireg2-5* and absence of T-DNA for *ireg2-4* suggests that this line is not homozygous. DNA sequencing of the PCR product obtained with the T-DNA border and gene specific reverse primers for SALK_07742C revealed that the T-DNA insertion was towards the 3' end of the intron between exon 1 and two in SALK_074442C, which is the expected position according to the signal T-DNA express database (Figure 3-2 below). Sequencing was not performed on the segregating SALK_12071C line.

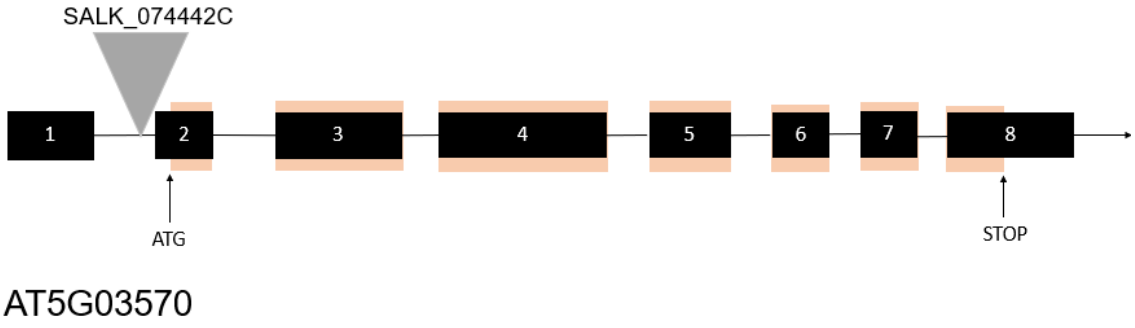


Figure 3-2: SALK 074442C T-DNA insertion*: Schematic representation of where the T-DNA insertion lies for SALK 074442C. Black numbered boxes indicate exons, black lines indicate introns and shaded region around exons represent protein coding regions.*

PCR genotyping also confirmed that SALK 097869C is homozygous T-DNA insertion line for *IRT1*, this is depicted in Figure 3-3 A below. Presence of a band at roughly 700 bp in the PCR reaction with the T-DNA left border primer and the gene specific reverse confirmed the presence of the T-DNA insertion (mutant allele) for all five individuals whereas Col-0 did not have the expected T-DNA band size, though the presence of non-specific amplification was evident. Absence of amplification in the PCR reaction with the gene specific forward and reverse primers indicated that the wild-type *IRT1* allele could not be amplified from the SALK 097869C mutants but was amplified from Col-0 (1 kb) showing the presence of the wild type allele.

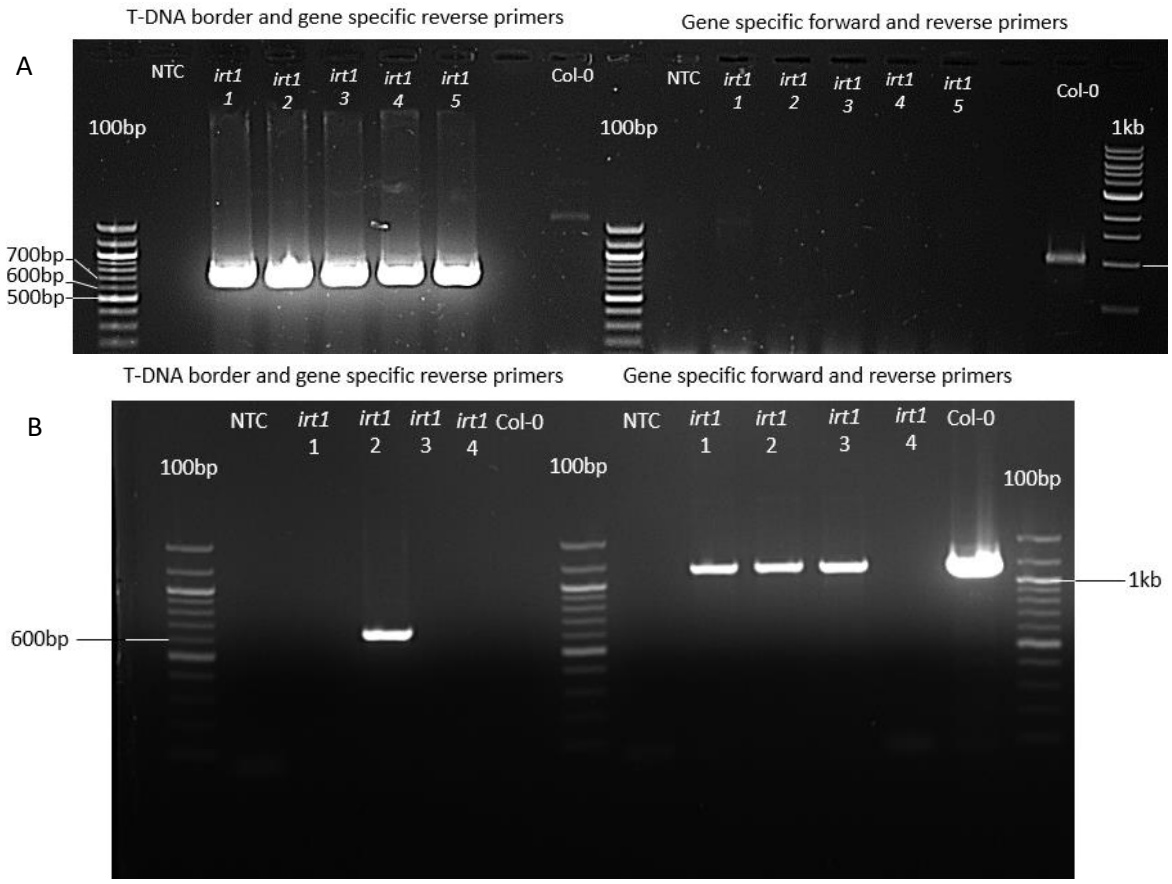


Figure 3-3: irt1 Genotyping: *Agarose gel electrophoresis showing results from PCR Genotyping on* irt1 *SALK 097869C mutants and wildtype (A) and SALK 054554C mutants and wildtype (B). Molecular weight ladders used were Quick-load® 100 bp DNA ladder (New England BioLabs) and Quick-Load® 1 kb DNA ladder (New England BioLabs)*

For SALK 054554C, *irt1-2* gave both wildtype and T-DNA bands of expected sizes (600 bp and 1 kb respectively) suggesting that that individual was heterozygous. The other two *irt1* individuals failed to amplify using the T-DNA border and gene specific reverse primers but amplified with the gene specific forward and reverse primers suggesting that they were wild type. DNA sequencing of the T-DNA PCR product from the confirmed homozygous *irt1* mutant (SALK 097869C) using the T-DNA border primer revealed that the T-DNA insertion is roughly 96 bp upstream of the transcriptional start site (TTS) which is slightly off from the predicted 24 bp downstream of the TTS (Figure 3-4 below). The other *irt1* mutant was not sent for sequencing owing to the fact that it was shown via PCR that they were not homozygous knock-out mutants.

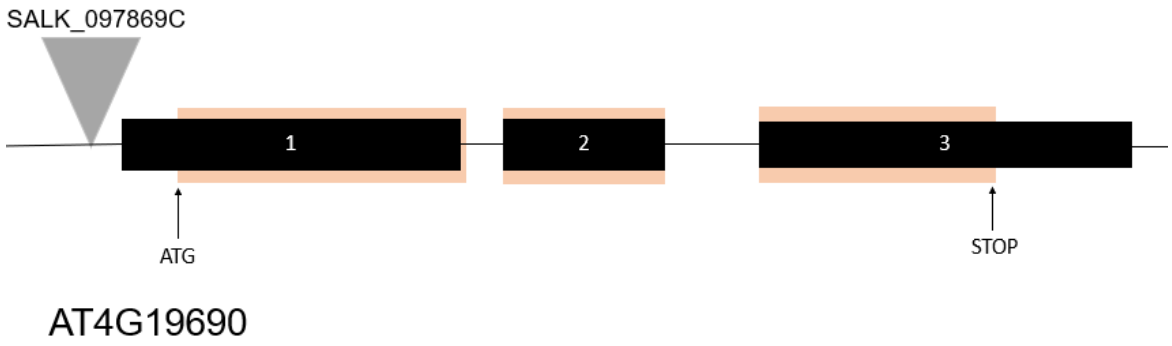


Figure 3-4 SALK 097869c T-DNA insertion*: Schematic representation of where the T-DNA insertion lies for SALK 097869C. Black numbered boxes indicate exons, black lines indicate introns and shaded region around exons represent protein coding regions*

3.1.2 Iron-deficiency induces low *IRT1* and *IREG2* gene expression

To elucidate whether the *irt1* and *ireg2* homozygous T-DNA insertion lines were indeed true knock-outs of these genes, *irt1*, *ireg2* and wildtype plants were grown hydroponically for two weeks in the presence of Fe and then subjected to ten days Fe starvation. The initial two-week Fe growth was to ensure that seeds germinated successfully and that the plants established themselves and were growing healthily before inducing stress. Under Fe-deficient conditions both *IRT1* and *IREG2* are induced in wildtype (Vert, et al., 2002; Schaaf, et al., 2006) plants. Confirmation of loss of gene expression is particularly important in complementation work as knock-out mutants must not have any endogenous expression which could interfere with functional analyses of transgenes. Root tissue was harvested, and total RNA was extracted. All cDNA samples were first subjected to a PCR using SAND (At2g28390) primers, (Figure 3-5 A [*irt1*] and B [*ireg2*]) (SAND Forward: 5'-CAGACAAGGCGATGGCGATA-3' and SAND Reverse: 5'-CTTTCTCTCAAGGGTTT CTGGGT-3'). *SAND*, also known as *MON1* encodes a protein which is involved in membrane trafficking and is constitutively expressed in a wide range of tissue types in *Arabidopsis*. In addition, it has been described as one of the more stable reference genes under various conditions (Zhao, et al., 2016). All the treatments generated a clear amplicon of approximately 300bp indicating that the cDNA was suitable for further use.

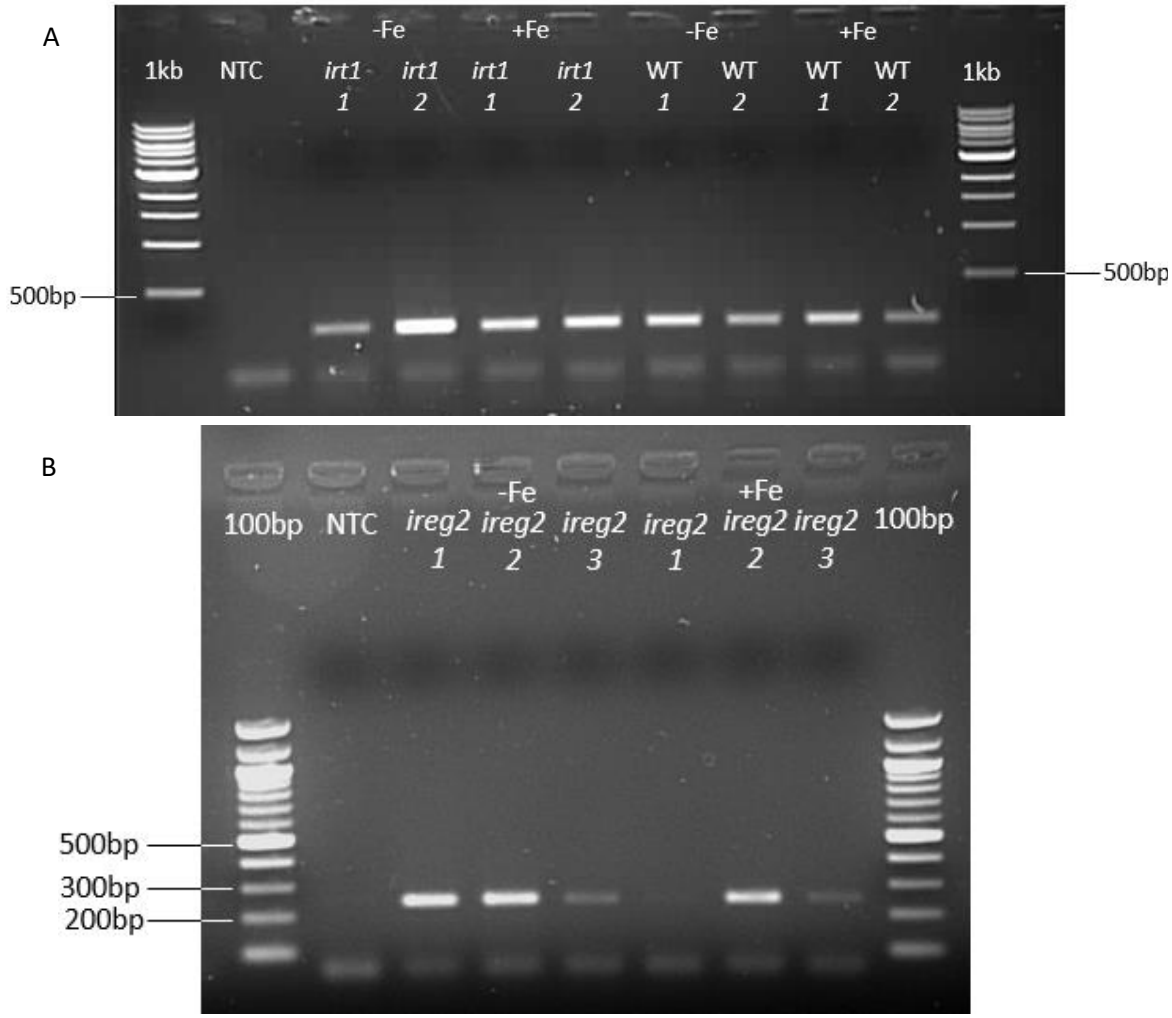


Figure 3-5: SAND PCRs: *Agarose gel electrophoresis showing results from SAND PCRs on* irt1 *homozygous and wildtype cDNA (A) and ireg2 homozygous cDNA (B) from the hydroponics assay. Quick-Load® 1 kb DNA ladder (New England BioLabs) used in A and Quick-load® 100 bp DNA ladder (New England BioLabs) used in B.*

The RT-PCR reactions for *ireg2* homozygous mutant and wildtype plants from the hydroponics assay (Figure 3-6 below) revealed amplification of a roughly 917 bp expected cDNA amplicon for all wildtype samples which were smaller than the DNA control (1204 bp), which contains seven short introns. DNA was used as a control for two reasons: firstly, as a positive control to indicate that the PCR reaction had worked and secondly to determine the presence of any contaminating DNA in the cDNA samples; no such contamination was detectable. The presence of the full-length cDNA amplicon clearly indicated that *IREG2* was expressed under both Fe-sufficient and Fe-deficient conditions. One of the *ireg2* homozygous mutants, however, also generated a faint band of the same size, but none of the other samples yielded any amplification

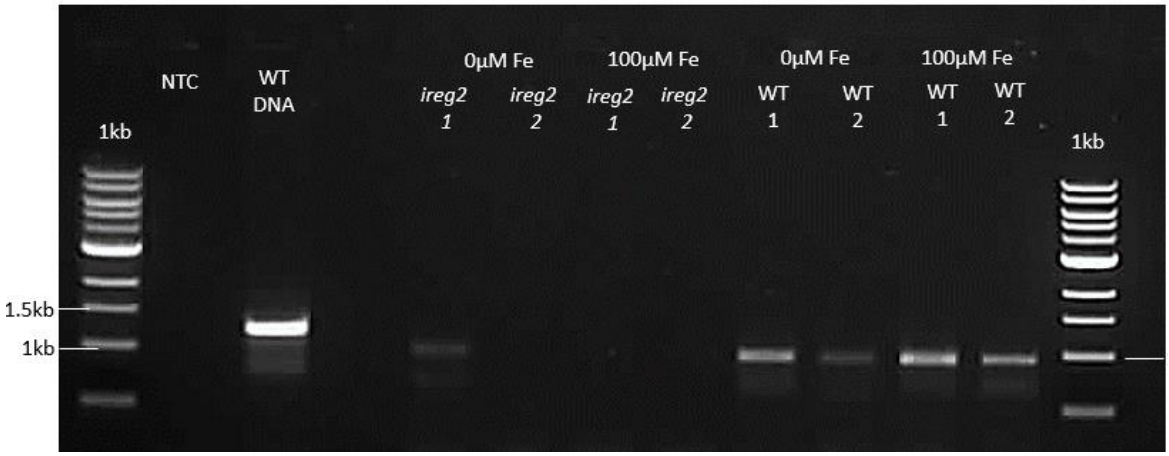


***Figure 3-6: IREG2 RT PCR**: Agarose gel electrophoresis showing results from cDNA RT qPCR for* ireg2 *homozygous mutant lines and wildtype plants following ten days with and without Fe. Molecular weight ladders used are Quick-load® 100 bp DNA ladder (New England BioLabs) and Quick-Load® 1 kb DNA ladder (New England BioLabs)*

The RT-PCR for the *irt1* homozygous mutant and wildtype cDNA proved to be less efficient than the *IREG2* RT PCRs. A faint product was observed in the WT 1 with no Fe whereas no other bands were visible for any of the other WT or *irt1* individuals. The transcript for *IRT1* is expected to be 1041 bp in size compared to the genomic product (1234 bp) which, according to Figure 3-7 is difficult to determine but the band in WT1 for 0 µM Fe is marginally smaller than the DNA control and the lack of DNA amplification in all the other samples is suggestive that the band present for WT 1 for 0 µM Fe is indeed *IRT1* cDNA. This experiment was repeated numerous times, each of which generated the same band for WT 1 to varying degrees of brightness, with no other cDNA samples yielding PCR products.

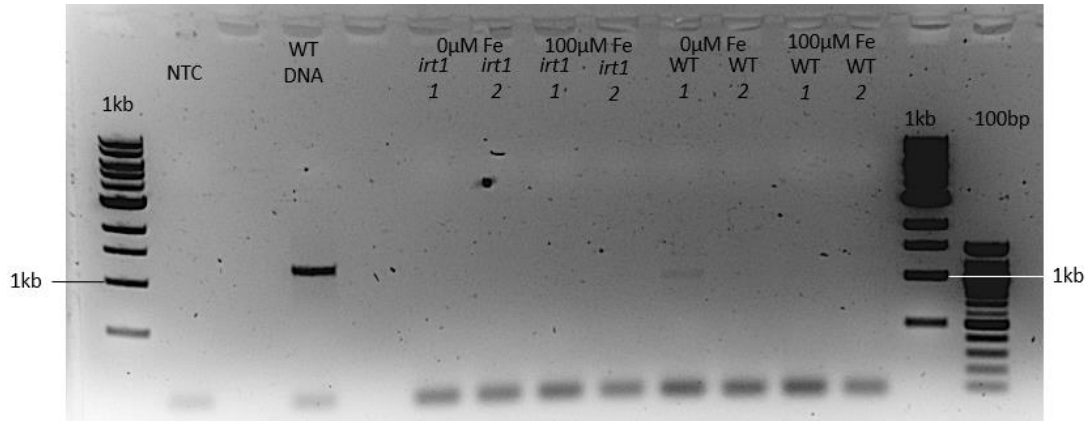


***Figure 3-7: IRT1 RT PCR**: Agarose gel electrophoresis showing results from cDNA RT PCR for* irt1 *homozygous mutant line and wildtype plants following ten days with and without Fe. Molecular weight ladders used are Quick-load® 100 bp DNA ladder (New England BioLabs) and Quick-Load® 1 kb DNA ladder (New England BioLabs)*

3.2 Phenotyping of T-DNA insertion lines

3.2.1 Reduced root biomass and length in the presence of nickel are reliable phenotypic observations for *ireg2* insertion lines

Having found homozygous *ireg2* mutants, an assay needed to be developed that would allow for assessment of *S. coronatus* complementation. As such, a robust phenotypic difference between *ireg2* mutants and wildtype plants had to be determined. If *ScIREG2* has functional homology to wildtype *Arabidopsis AtIREG2,* then it would then be expected to restore the wild-type phenotype. Both wildtype and *ireg2* plants were grown hydroponically for two weeks with Fe and then subsequently grown without Fe and in the presence of 10 µM Ni for 10 days as no Fe induces expression of *IRT1* (Vert, et al., 2002) which acts as a non-specific transporter for Ni entry which is then sequestered by *IREG2* (Schaaf, et al., 2006). In the presence of Ni, it was found that *ireg2* had significantly lower root biomass production compared to wildtype plants (U=7; Z-score = -3.1909; p=0.0142). Between Ni treatments, *ireg2* grown in the presence of Ni had significantly lower root biomass than *ireg2* grown in the absence of Ni (U=16; Z-score= -2.71109; p=0.0067). Although a reduction root biomass production was observed for wildtype plants, this was not significantly different.

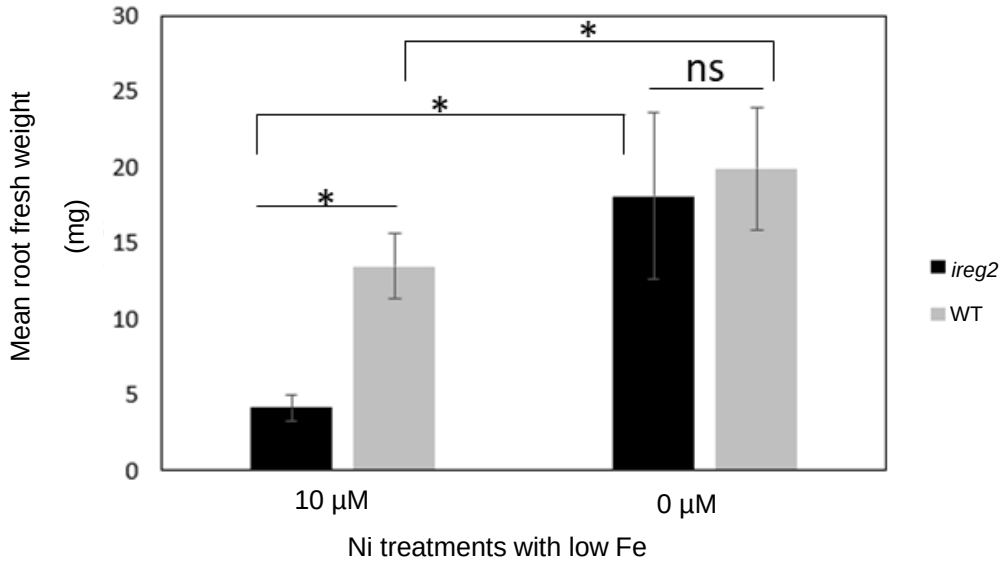


Figure 3-8: ireg2 Phenotyping root biomass assay*: Results from ireg2 phenotyping showing that there is significant difference between ireg2 and wildtype plants under Fe-deficient conditions in the presence of 10 µM $NiCl_2$. Significant difference in root biomass was also observed for ireg2 mutants under Fe-deficient treatments in the presence and absence of 10 µM $NiCl_2$. Significance indicated as* ***** *with p<0.05 based on Mann-Whitney U test. Error bars indicate standard error, N=3 biological repeats.*

Although there was a reduction in mean shoot biomass production for both *ireg2* and wildtype plants between Ni treatments, these were not significant (

Figure ***3-9***). Both, however, displayed severe chlorosis on primary leaves under absence of Fe and presence of Ni treatments.

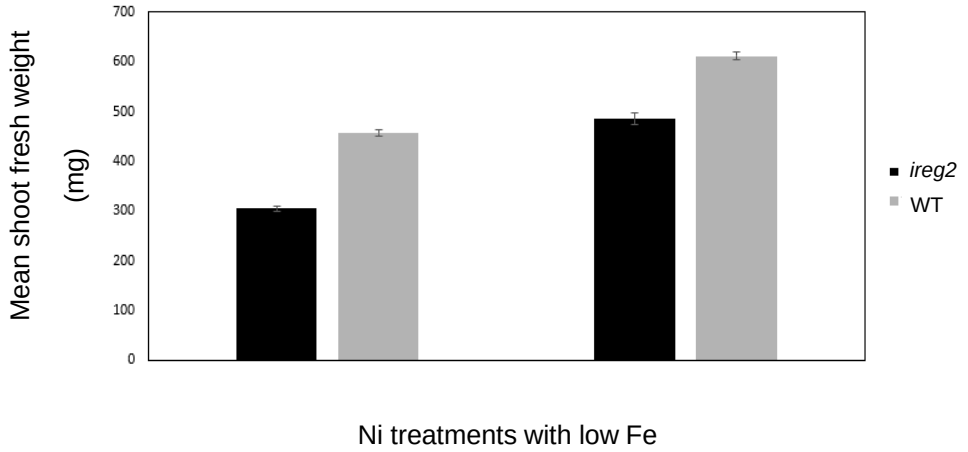


***Figure 3-9: ireg2 Phenotyping shoot biomass assay*:** *Results from ireg2 phenotyping showing that there is no significant difference in mean shoot biomass between ireg2 and wildtype plants under Fe-deficient conditions in the presence or absence of 10 µM $NiCl_2$ based on Mann-Whitney U-test. Error bars indicate standard error, N=3 biological repeats.*

When comparing the root length (mm) of *ireg2* in the presence and absence of Ni, those grown in the presence of Ni had significantly shorter roots than those grown in the absence (U=9.5; Z-score=-3.1688; p=0.00152) (

Figure ***3-10***). Wildtype plants had shorter roots on average in the presence of Ni though they were not significantly shorter than those in the absence of Ni.

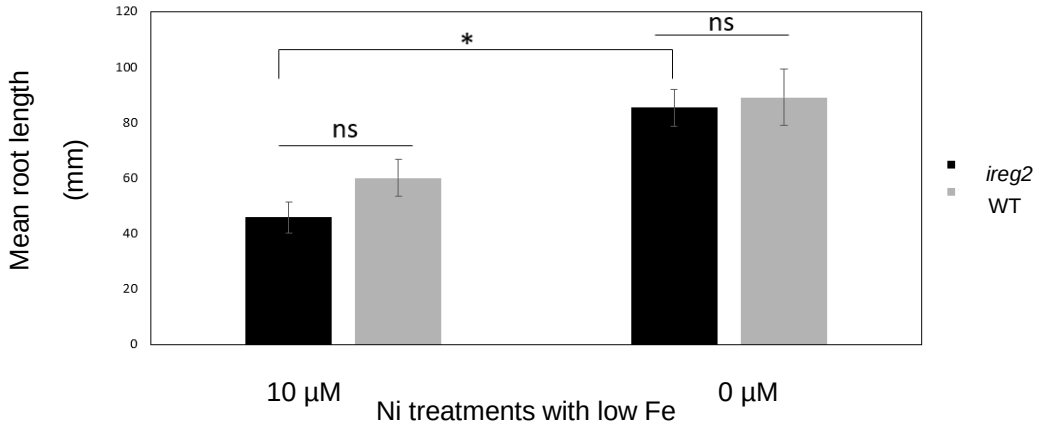


Figure 3-10: ireg2 Phenotyping root length assay*:** *Results from ireg2 phenotyping showing that there is a significant difference in mean root length between ireg2 plants under Fe-deficient conditions in the presence or absence of 10 µM $NiCl_2$. Significance indicated as* ** *with p<0.05 based on based on Mann-Whitney U test. Error bars indicate standard error, N=3 biological repeats.*

Dimethylglyoxime staining of roots and shoots from *ireg2* mutants and wildtype plants showed presence of Ni in wildtype roots under Fe-deficiency and 10 µM $NiCl_2$ whereas no red colour was observed in the roots for *ireg2* mutants for the same treatment, summarised in Table 3-1 and in Figure 3-11 below. Shoot tissue failed to show any definitive differences. Red staining, however, was observed in both wildtype and *ireg2* mutant plants' trichomes (data not shown).

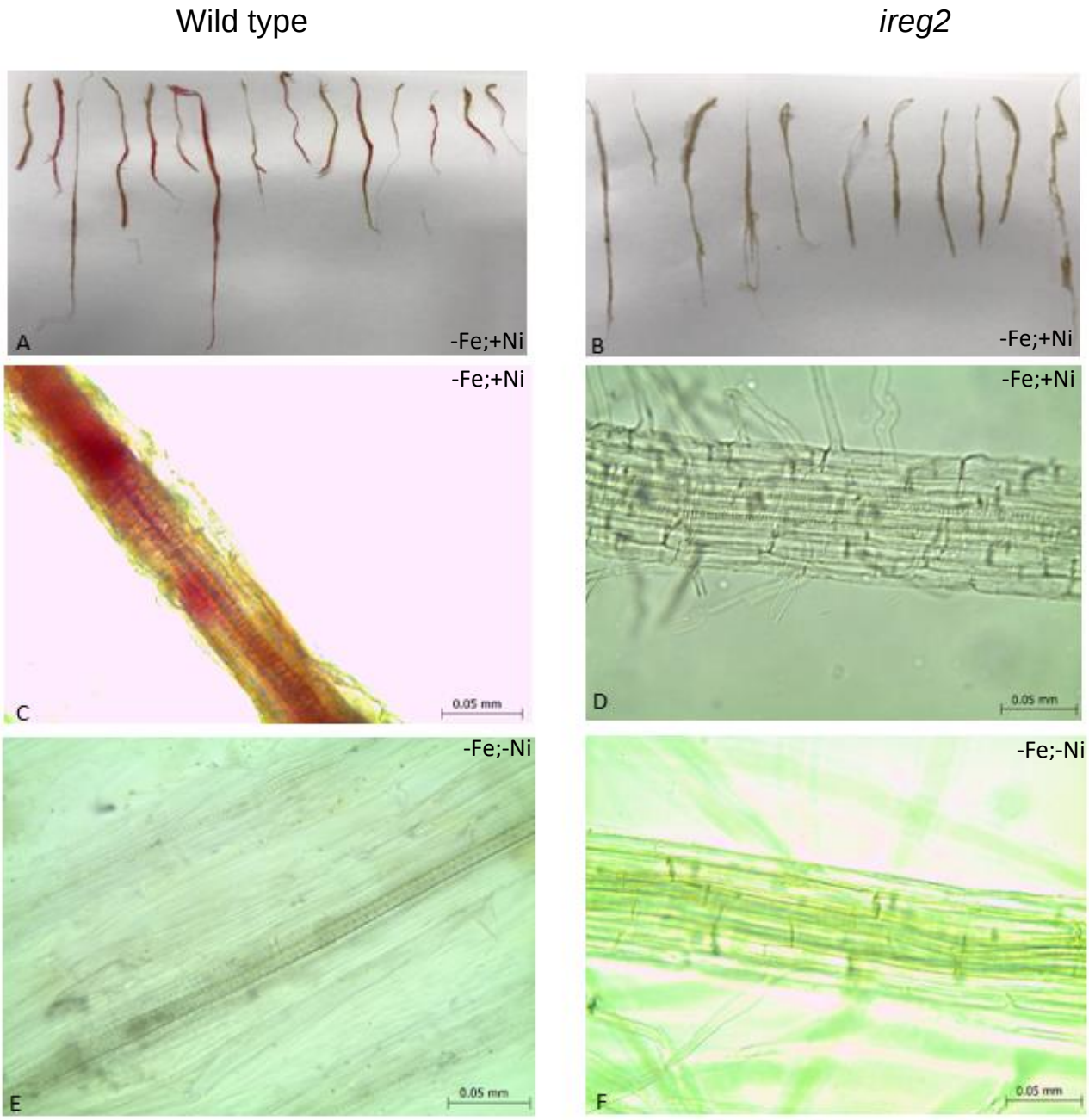


Figure 3-11: Dimethylglyoxime Root staining*: Root staining of WT (A, C, and E) and* ireg2 *(B, D, and F) with D presence of red precipitate for WT, indicated by the darkened areas, and lack of red precipitate in* ireg2 *root tis C-F at 40X magnification*

***Table 3-1**: Dimethylglyoxime staining of* ireg2 *and* irt1 *mutant, and wildtype root tissue*

	10 µM $NiCl_2$	**0 µM $NiCl_2$**
ireg2	Absent	Absent
irt1	Present	Absent
Wildtype	Present	Absent

3.2.2 Non-lethal phenotype for *irt1*

As the primary route for Fe acquisition during Fe starvation, lack of functionality of *IRT1* results in a lack of Fe uptake, which has severe implications on overall plant growth. Previously published results (Vert, et al., 2002) on *irt1* indicate that *irt1* mutants display a lethal phenotype when grown in the absence of Fe. Notable is the severe reduction in overall biomass and a general retardation in growth to the point where *irt1* mutants simply fail to develop beyond the four-to-six leaf stage. To determine whether a lethal phenotype was observed, *irt1* and wildtype seeds were plated on PN agar plates in high (100 µM) and low (10 µM) Fe concentrations and were allowed to grow for two weeks, after which relative survival was determined. No marked difference in survival of the *irt1* seedlings was observed between the two treatments; 44.9% survival on 10 µM Fe versus 57.3% on 100 µM Fe. While this was lower than the 67.4% survival exhibited by wild-type plants on 10 µM Fe, it is certainly not consistent with the expected lethal phenotype.

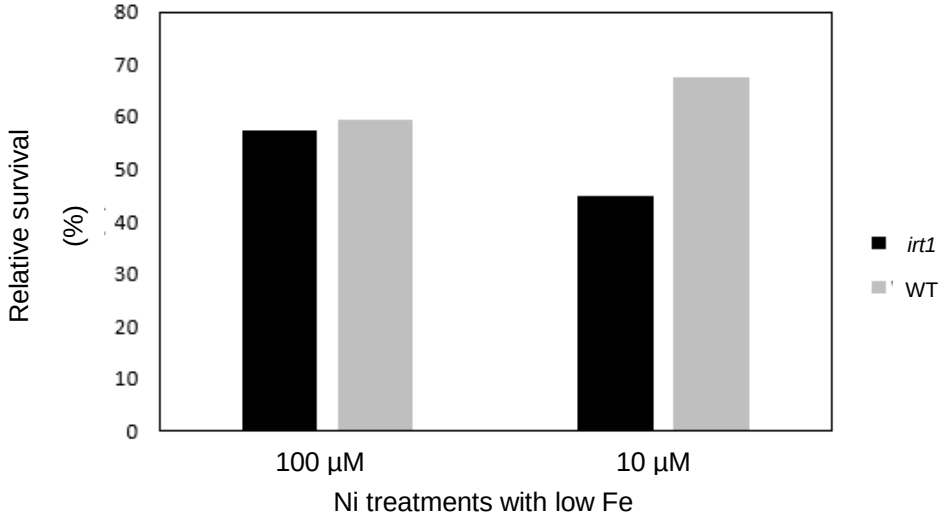


***Figure 3-12: irt1 Survival**: Relative survival as percentage of* irt1 *and WT plants in the presence of 100 µM and 10 µM on PN media grown for two weeks. Plants which had developed true leaves were considered as survivors.*

In hydroponic conditions where *irt1* and wildtype plants were grown in the presence of Fe for two weeks and then starved of Fe for ten days no statistical differences were noted in

either shoot or root biomass (Figure 3-13 and Figure 3-14 below). Although there was a marginal reduction in overall shoot and root biomass when grown in the absence of Fe, this was not statistically significant for either genotype.

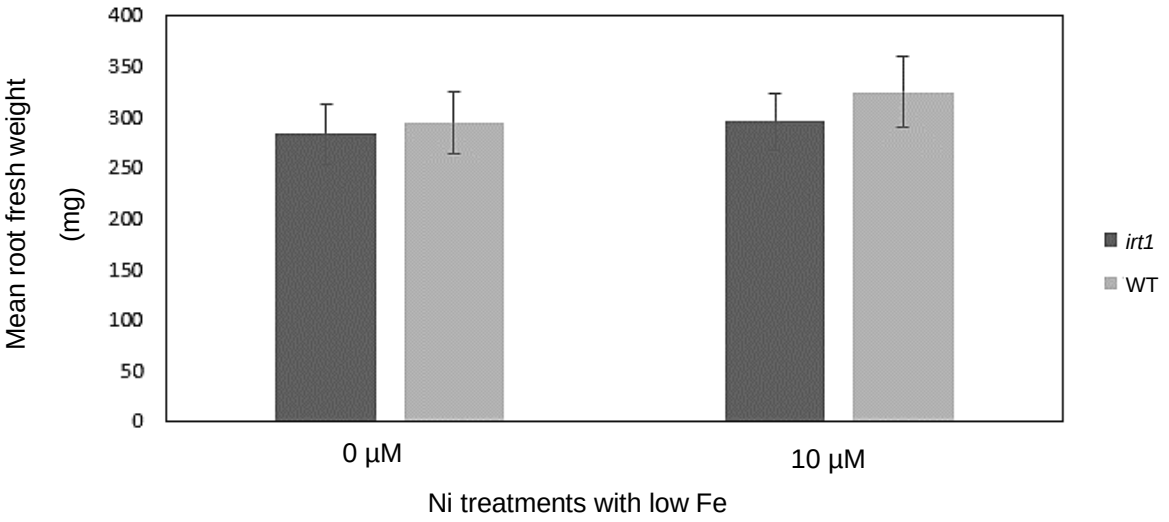


Figure 3-13: Phenotyping of* irt1 *root biomass assay*: Results from* irt1 *phenotyping showed that there is no significant difference in mean root biomass between* irt1 *and wildtype plants under Fe-deficient conditions in the presence or absence of 10µM $NiCl_2$ based on Mann-Whitney U test. Error bars indicate standard error, N=3 biological repeats.*

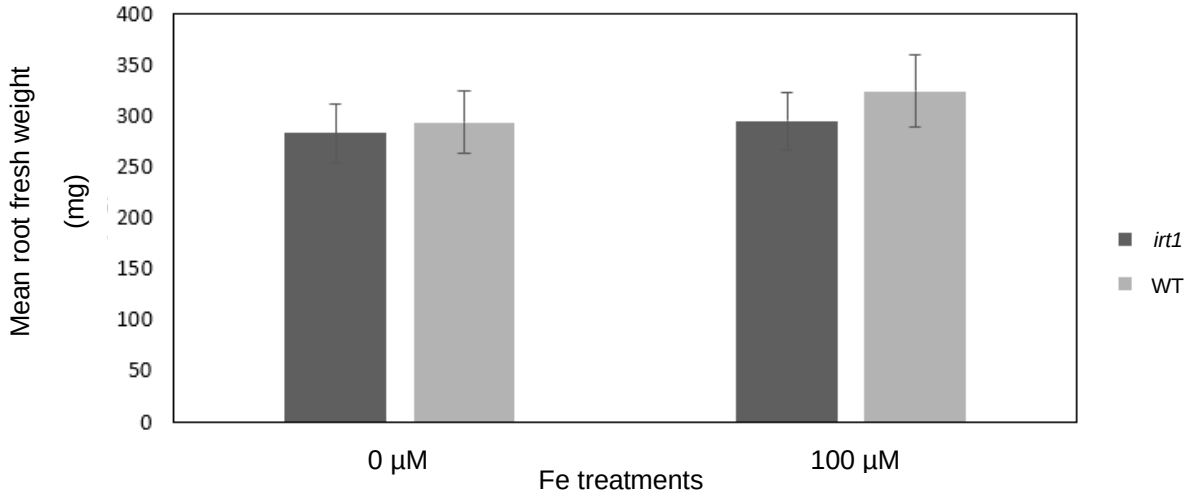


Figure 3-14: Phenotyping of* irt1 *shoot biomass assay: *Results from* irt1 *phenotyping showed that there is no significant difference in mean shoot biomass between* irt1 *and wildtype plants under Fe-deficient conditions in the presence or absence of 10µM $NiCl_2$ based on Mann-Whitney U test. Error bars indicate standard error, N=3 biological repeats.*

In a separate experiment, *irt1* and wildtype plants were grown hydroponically as before with the addition of Ni to the Fe-deficient media. This was done to assess if any Ni was able to

enter root tissue of *irt1* mutants, *IRT1* being thought to be the major route of Ni uptake under Fe deficient conditions in Arabidopsis (Nishida et al. 2011). No difference in DMG staining between wildtype and *irt1* plants was apparent (Table 3-1).

3.3 Gateway cloning for *ScIRT1* and *ScIREG2* complementation

In parallel to the identification and phenotypic characterisation of the T-DNA lines, the generation of entry and subsequent expression clones for p*AtIRT1*:*ScIRT1* and p*AtIREG2:ScIREG2* were being made whereby *attB*-containing PCR amplicons (*pAtIRT1*, *ScIRT1*, *pAtIREG2*, and *ScIREG2*) were first cloned into respective pDONR221 entry vectors via BP Clonase® II reactions and then these BP-clones where then allowed to recombine with one another such that a two-fragment expression clone was generated after an LR Clonase® II+ reaction, for example pAtIRT1 in pDONR221 P1P5r was recombined with ScIRT1 in pDONR221 P5P2 using a suitable Gateway® destination vector. Firstly, the protein coding regions of *S. coronatus IRT1* and *IREG2* were amplified from *S. coronatus* cDNA with primers containing *attB* sites at their 5' and 3' ends. Next the *Arabidopsis* promoters, *pAtIRT1* and *pAtIREG2* were amplified from genomic *Arabidopsis* Col-0 DNA with *attB* containing primers. Given the general AT-richness and tandem repeats of *Arabidopsis* promoters, these were amplified using a two-step PCR. PCR amplicons are shown below in Figure 3-15. All the amplicons were amplified successfully using a high fidelity thermostable polymerase and were subsequently PEG purified as per the Gateway® Multisite Pro manual.

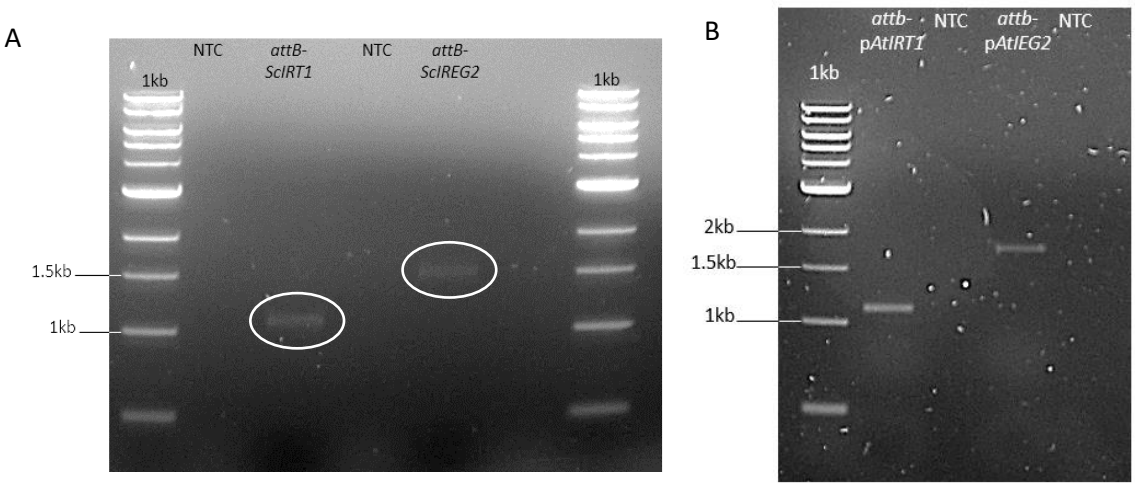


Figure 3-15: PCR amplicons for Gateway® cloning*: Agarose gel electrophoresis of amplicons used for Gateway® cloning. Image A shows successful amplification of the* S. coronatus *protein coding regions whereas image B shows successful amplification the* A. thaliana *promoters. Molecular weight ladder used was Quick-Load® 1kb DNA ladder (New England BioLabs)*

The successfully amplified *attB*-containing amplicons were successfully recombined in their respective *attP*-containing pDONR221 vectors, though this required some optimisation. Following colony PCRs using gene specific *attB*-containing primers, plasmid DNA was extracted from putative positive colonies, and PCR reactions repeated as shown in Figure 3-16 below. Sequences were aligned to reference sequences (see appendix); all of which aligned 100% to the reference sequence except for BP *ScIREG2* (see below).

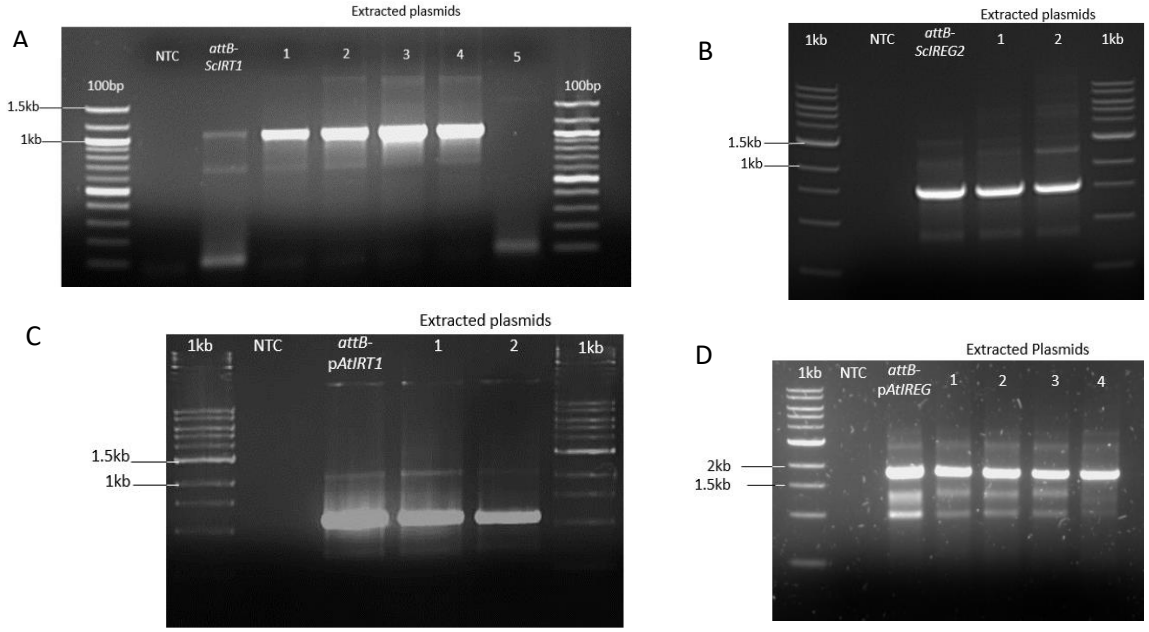


Figure 3-16: BP-Clonase® II plasmid clones*: Agarose gel electrophoresis showing results from PCRs carried out on plasmids extracted from BP-Clonase® II colonies. Gel image A shows* ScIRT1 *in pDONR221 P5P2; image B shows* ScIREG2 *in pDONR221 P5P2; image C shows* pAtIRT1 *in pDONR221 P1P5r; and image D shows* pAtIREG2 *in pDONR221 P1P5r. The original attB-containing PCR amplicon served as positive control as indicated. Molecular weight ladder used was Quick-Load® 1kb DNA ladder (New England BioLabs) and Quick-Load® 100bp DNA ladder (New England BioLabs)*

The BP Clonase® II reaction for *ScIREG2* proved the most difficult to optimise as it required the need to redo this reaction five times, as well as having to reamplify the initial amplicon. Two of those reactions resulted in successful clones, however, sequencing revealed incorrect bases when compared to the reference gene (Figure 3-17). Although most of the sequences aligned perfectly with the reference sequence, the point mutations resulted in amino acid changes shown in Figure 3-17.

DNA multiple sequence alignments for two BP Clonase reactions for ScIREG2

```
BP IREG FEB 2019        ACAGGATTCATCATAAGTTTCGTGTCAGAGATTGCATCTG    666
BP IREG SEP 2018        ACAGGATTCATCATAAGTTTCGTGTCAGAGATTGCATCTG    666
IREG REF GENE           ACAGGATTCATCATAAGTTTCGTGTCAGAGATTGCATCTG    597
Consensus               acaggattcatcataagtttcgtgtcagagattgcatctg

BP IREG FEB 2019        CTGCCGGTTTAGCAATCTTTAATACATTATCAGTTTTCTT    706
BP IREG SEP 2018        CTGCCGGTTTAGCAATCTTTAATACATTATCAGTTTTCTT    706
IREG2 REF GENE          CTGCCAGTTTAGCAATCTTTAATAtATTATCAGTTTTCTT    637
Consensus               ctgcc gtttagcaatctttaata attatcagttttctt

BP IREG FEB 2019        GCAATATTGGCTCTTAAACTCTGTATACAAAGGAATCCCA    746
BP IREG SEP 2018        GCAATATTGGCTCTTAAACTCTGTATACAAAGGAATCCCA    746
IREG REF GENE           GCAATATTGGCTCCTAAACTCTGTATACAAAGGAATCCCA    677
Consensus               gcaatattggctc taaactctgtatacaaaggaatcccca

BP IREG FEB 2019        TCTTTATCAGAAAGAAGCCGACAGCGTTCCCTAAGATTCG    786
BP IREG SEP 2018        TCTTTATCAGAAAGAAGCCGACAGCGTTCCCTAAGATTCG    786
IREG REF GENE           TCTTTATCAGAAAGAAGCCGACAGCGTTCCCTAAGATTCG    717
Consensus               tctttatcagaaagaagccgacagcgttccctaagattcg
```

Amino Acid alignment showing amino acid changes

```
BP IREG     MWEFSVGLYMINVWPNSLLLAATYGVVESASTTLFGPLVGQWIDKSTYPKVLKVWLLTQN  60
IREG REF    MWEFSVGLYMINVWPNSLLLAATYGVVESASTTLFGPLVGQWIDKSTYPKVLKVWLLTQN  60
Consensus   MWEFSVGLYMINVWPNSLLLAATYGVVESASTTLFGPLVGQWIDKSTYPKVLKVWLLTQN

BP IREG     LSFIVAGVTVVGLLISPDLRINNRVLFAVLVVMINLSGALAVLSSLAGTILIEREWVVVI  120
IREG REF    LSFIVAGVTVVGLLISPDLRINNRVLFAVLVVMINLSGALAVLSSLAGTILIEREWVVVI  120
Consensus   LSFIVAGVTVVGLLISPDLRINNRVLFAVLVVMINLSGALAVLSSLAGTILIEREWVVVI

BP IREG     SEGRSSNILTTLNSRIRRIDLVSKLFAPVATGFIISFVSEIASAASLAIFNILSVFLQYW  180
IREG REF    SEGRSSNILTTLNSRIRRIDLVSKLFAPVATGFIISFVSEIASAAGLAIFNTLSVFLQYW  180
Consensus   SEGRSSNILTTLNSRIRRIDLVSKLFAPVATGFIISFVSEIASAA LAIFN LSVFLQYW

BP IREG     LLNSVYKGIPSLSERSRQRSLRFATANDQQQTQTSSTPQEQNDNSEDV  228
IREG REF    LLNSVYKGIPSLSERSRQRSLRFATANDQQQTQTSSTPQEQNDNSKDV  228
Consensus   LLNSVYKGIPSLSERSRQRSLRFATANDQQQTQTSSTPQEQNDNS DV
```

Figure 3-17: Point mutations in BP ScIREG2*: Point mutations seen from two BP Clonase reactions showing highlighted in the DNA multiple sequence alignment and amino acid alignment showing amino acid changes*.

The two-fragment LR-Clonase® II plus reactions became problematic for several reasons. The initial destination vector, pB7WG which was selected on the basis of it being Gateway compatible and carrying spectinomycin resistance gene (as opposed to kanamycin resistance gene which is carried on pDONR vectors) yielded too many unsuccessful

transformants upon initial use (>200 CFU on 100 µL vector only plate, when none would be expected due to presence of *ccdB* gene, compared to 70 CFU on 100 µL experimental LR® Clonase II+ *AtIRT:ScIRT1* plate). The vector's suitability for two-fragment recombination was then tested by performing the LR-Clonase® II + reaction using pENTR L1-plac-lacZα-R5 and pENTR L5-plac-spec-L2 as entry vectors on LA plates containing Xgal (as substrate analogue of lactose for β-galactosidase) and IPTG (as inducer). This method of testing exploited the use of blue-white bacterial selection, blue colonies indicating successful two-fragment recombination into pB7WG, and white colonies the empty pB7WG vector. Although what can be described as 'lawn-like' growth was observed on the LA Xgal and IPTG plates, the results revealed that although pB7WG was able to act as a destination vector for recombination reaction, evident by the presence of some blue colonies, it produced many unsuccessful white colonies relative to the few blue colonies, implying the *ccdB* gene in pB7WG was not working correctly. The feasibility of using pEARLYGATE302 was then also tested in the same manner as above and it yielded many more successful blue colonies than white colonies. However, it and the pDONR221 vectors both confer kanamycin resistance which would have required linearizing of the pDONR221 clones, which was unlikely to be 100% efficient. As such, it became apparent that an alternative destination vector would be required and so pB2NADA was subsequently generated from pB2GW7.0 (a Gateway compatible destination vector containing a spectinomycin resistance gene). Following removal of the CaMV 35S promoter from pB2WG7.0 by restriction enzyme digest, blunt ending and finally ligation, the newly made vector was then digested with *NcoI* and *PstI* to confirm successful execution. The presence of two bands (~7947 bp and >500-1870 bp) for the *PstI* digests and one single band (~9817 bp) for the *NcoI* digests meant success in this instance shown below in Figure 3-18. If unsuccessful then *PstI* digest would produce three bands (9818, 1807, 157 bp) whereas *NcoI* would also produce three bands (9560, 932, 390 bp). Plasmids were selected and sent for sequencing at SUNCAF to validate full removal of CaMV 35S promoter using the aforementioned N35S forward primer (page 32).

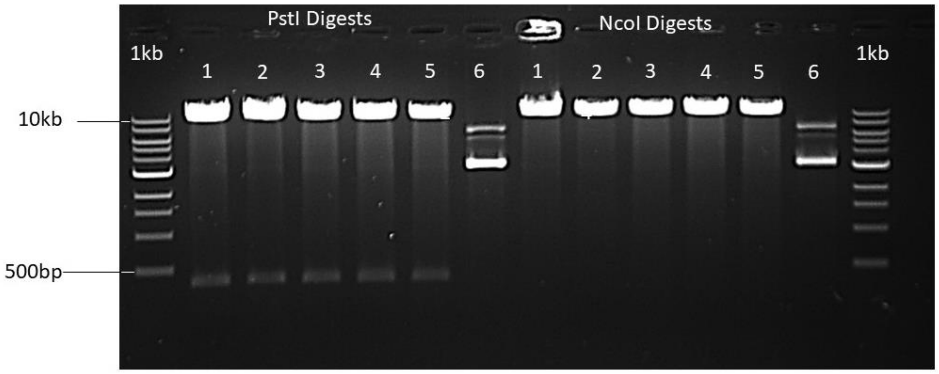
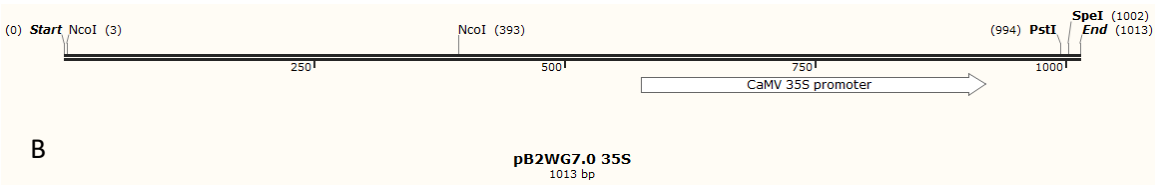


***Figure 3-18: pB2NADA digests**: Agarose gel electrophoresis showing results from* PstI *and* NcoI *digests on pB2NADA vectors. Presence of two bands for* PstI *and presence of one band for* NcoI *digests indicate successful removal of CaMV35 promoter and ligation. Vector six was unsuccessful. Molecular weight ladder used was Quick-Load® 1kb DNA ladder (New England BioLabs)*

Once the sequenced, pB2NADA and pB2WG7.0 DNA sequences were put into *SnapGene Viewer*'s software, the software was able to show that the CaMV 35S promoter had been successfully removed from pB2NADA (Figure 3-19 D) compared to pB2WG7.0 (Figure 3-19 B). When aligned using NCBI's BLAST, pB2NADA failed to return any CaMV 35S promoter regions whereas the original vector returned CaMV 35 promoter regions (Figure 3-19 A and C for pB2WG7.0 and pB2NADA respectively).

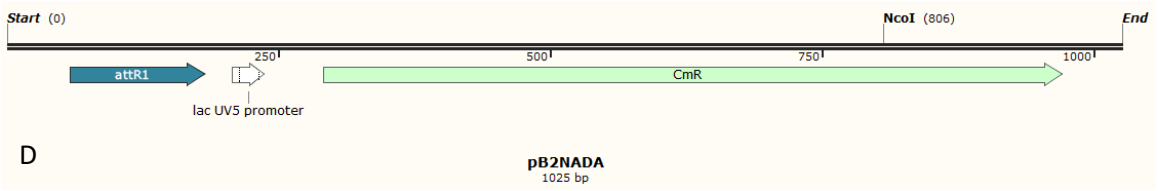
A

	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
☑	Binary vector pHAtC, complete sequence	1749	1835	99%	0.0	99.79%	KU213971.1
☑	Binary vector pBAtC, complete sequence	1749	1835	99%	0.0	99.79%	KU213970.1
☑	Gateway vector p2RLGW7, complete sequence	1735	1735	95%	0.0	99.68%	KY434119.1
☑	Gateway vector p2FLGW7, complete sequence	1735	1735	95%	0.0	99.68%	KY434118.1
☑	Cloning vector p35S::CASP2:GFP, complete sequence	1735	1735	95%	0.0	99.68%	HQ699544.1
☑	Cloning vector p35S::CASP3:GFP, complete sequence	1735	1735	95%	0.0	99.68%	HQ699543.1
☑	Cloning vector p35S::CASP1:GFP, complete sequence	1735	1735	95%	0.0	99.68%	HQ699542.1
☑	Cloning vector p35S::CASP5:GFP, complete sequence	1735	1735	95%	0.0	99.68%	HQ699541.1
☑	Cloning vector p35S::CASP4:GFP, complete sequence	1735	1735	95%	0.0	99.68%	HQ699540.1
☑	Cloning vector p35S-Cas9 DNA, complete sequence	1725	1725	94%	0.0	99.68%	LC189561.1



C

Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
Diatom vector pDEST-C-HA, complete sequence	1783	2103	94%	0.0	99.69%	LT726866.1
Diatom vector pDEST-C-EYFP, complete sequence	1783	2103	94%	0.0	99.69%	LT726865.1
Cloning vector pXF20pemIK-GW, complete sequence	1783	2103	94%	0.0	99.69%	KX036765.1
Binary vector pUB-GW-Hyg DNA, complete sequence	1783	2097	95%	0.0	99.49%	AB303065.1
Binary vector pUB-GW-GFP DNA, complete sequence	1783	2097	95%	0.0	99.49%	AB303064.1
Binary RNAi Gateway destination vector pAGSM552, complete se	1781	2757	95%	0.0	99.49%	KP259613.2
Shuttle vector pGWFDR196, complete sequence	1781	2103	95%	0.0	99.39%	EF515872.1
Moss transformation vector pT1OG DNA, complete sequence	1779	2095	94%	0.0	99.69%	LC126301.1
Transformation vector pBGRY, complete genome	1779	2004	94%	0.0	99.69%	KM016698.1
Transformation vector pBDPR, complete genome	1779	2004	94%	0.0	99.69%	KM016697.1



D

Figure 3-19: BLAST results for pB2WG7.0 and pB2NADA*: NCBI Blast results show top 10 hits from sequenced pB2WG7.0 (A) and pB2NADA (C) using N35S primer. Image B and C shows graphic representation of the sequenced pB2WG7.0 and pB2NADA respectively with important restriction enzyme sites used for creation of pB2NAD and validation of removal of CaMV35S promoter.*

After sequencing confirmed that the newly made pB2NADA was a suitable destination vector, it was transformed into *E.coli* DH5α cells to evaluate the efficacy of its *ccdB* gene and plated on 100 mg.mL^{-1} spectinomycin LA plates, where less than 10 CFU were counted on the 90 µL LA plates. The two-fragment LR Clonase® II+ was then performed using *ScIRT1* in pDONR221 P5P2 and *AtIRT1* in pDONR221 P1P5r as entry clones as per the Multisite Gateway Pro manual and pB2NADA as the destination vector. Figure 3-20 below shows the results from a successful multi-fragment LR Clonase® II+ reaction. Lanes 1-3 show restriction digests using EcoRV as it would yield three fragments (8110, 1621, and 627 bp) for a successful multi-fragment clone whereas only two fragments (8106 and 1711 bp) are generated on the empty pB2NADA vector. From this, clone LR1 successfully generated the expected band sizes. PCR reactions using gene specific *attB*-containing primers were also used to validate that both inserts were there. One set used the *attB*-containing *AtIRT1* forward and *attB*-containing *ScIRT1* reverse primer (NTC and LR1 reactions in lanes 4 and 5 in Figure 3-20 below) and yielded a band that could only be present if *AtIRT1* and *ScIRT1* had successfully been recombined. The next two PCR reactions (NTC – LR1 in lanes 6 to 11 in Figure 3-20 below) used the *attB*-containing *ScIRT1*

forward and reverse primers and *attB*-containing *AtIRT1* forward and reverse primers respectively. The results indicated that both the *Arabidopsis* promoter and *S. coronatus* protein coding region was present. The successfully identified clone (LR1) was sent for DNA sequencing; validating that pAtIRT1 and ScIRT1 were successfully recombined and were present in pB2NADA

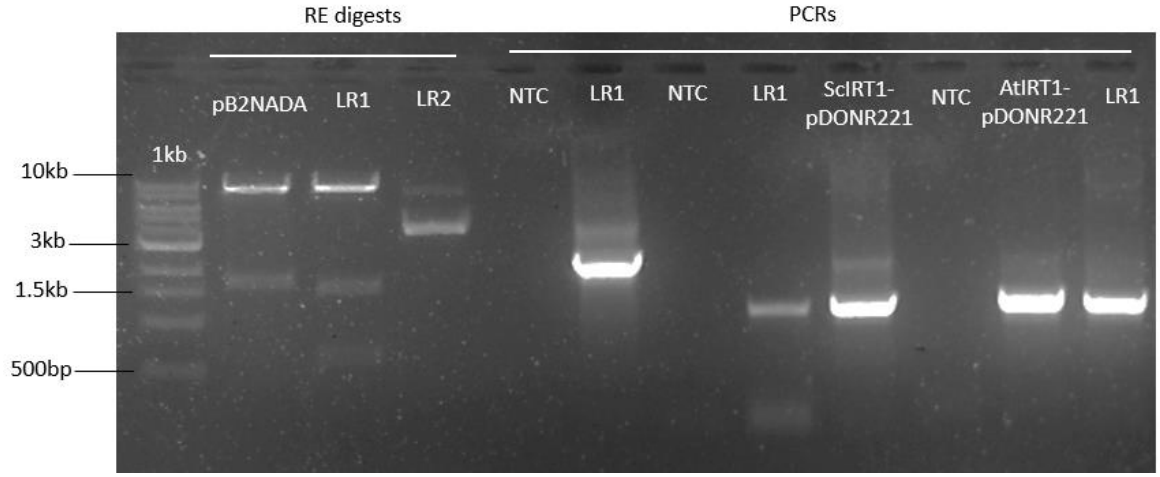


Figure 3-20: Successful IRT1 LR® Clonase II+: *Agarose gel electrophoresis of restriction digests and PCR on IRT LR Clonase II+. Molecular weight ladder used was Quick-Load® 1 kb DNA ladder (New England BioLabs)*

.

DISCUSSION

The research conducted herein set out to evaluate whether *ScIRT1* and *ScIREG2* from the Ni hyperaccumulator *Senecio coronatus* might play a potential role in Ni hyperaccumulation by determining whether they are functional homologs of the *Arabidopsis IRT1* and *IREG2* proteins, which are known to transport Ni (Nishida, et al., 2011; Schaaf, et al., 2006). As such, the primary aims of the research were to identify homozygous *Arabidopsis* mutants and develop robust phenotypic assays for use in future complementation experiments, whilst concurrently generating expression clones with the *S. coronatus* ORFs under the regulation of the respective *Arabidopsis* promoters.

The PCR genotyping of the *ireg2* T-DNA insertion lines confirmed the homozygosity of SALK 07442C but failed to confirm that of SALK 127071C. Both lines had been used by Schaaf, et al., (2006) and whilst the authors reported that both lines were homozygous, here the homozygosity of SALK 127071C could not be validated due to poor amplification of the wildtype allele. Although no wildtype allele was observed for SALK 127071C (Figure 3-1), the absence of T-DNA border primer PCR product for one of the samples brings into question the validity of this line's homozygosity for the T-DNA insert, at least for that individual. This naturally emphasises the need to validate the homozygosity of SALK lines as although they are reported as homozygous, this is not always the case.

Furthermore, homozygosity for T-DNA insertion does not mean that the plants necessarily lack expression of the target gene (Ülker, et al., 2008). This must be validated experimentally as not having complete knock-out lines would obviously impair the reliability of transgenic complementation as the effect one would observe may in fact just be endogenous expression. The use of RT-PCR to assess expression using end-point PCR is a valuable tool to ensure that mutants that are homozygous for a T-DNA insertion do indeed lack endogenous expression of the gene of interest. The RT-PCR results suggest that the SALK 07442c line is likely a true knock-out, as evidenced from the lack of expression observed in three of the four cDNA samples tested (Figure 3-6). However, the presence of a faint band in Figure 3-6 for one of the *ireg2* samples grown in the absence of Fe was noted. It is possible that this may be due to experimental set up error. Both wildtype and *ireg2* mutants were grown in the same hydroponics boxes so as to eliminate any 'box effects' and reduce experimental variability. It is possible that some wildtype root tissue became entangled with the *ireg2* tissue when it was collected. The absence of any expression in the other three *ireg2* samples compared to the clear presence of *IREG2* mRNA obtained for all four wildtype samples suggests that that particular sample may have been contaminated. Independent replication of this experiment is

critical to determine whether the SALK07442C line is indeed a knock-out line, but the results are encouraging. The presence of expression of the wildtype plants under Fe-deficient and Fe-sufficient media suggests that native *AtIREG2* is expressed at a low level even when Fe is present. This observation is also consistent with similar results obtained by Schaaf, et al., (2006), showing some degree of responsiveness to Fe. As *AtIREG2* is regulated by the *FRU/FIT1* transcription factors which are induced during Fe deficiency (Bauer & Guerinot, 2007), one would only expect higher expression under the Fe starvation state. However, it appeared that *AtIREG2* was expressed constitutively in this experiment. Quantitative PCR will need to be employed to determine whether its expression is Fe-responsive under the growth conditions used here.

The phenotyping assays revealed that exposure of the *ireg2* mutant to 10 µM Ni under Fe-deficient conditions resulted in a significant decrease in root biomass production and root length, which was not observed in wild-type plants following three independent experiments (Figure 3-8 and

Figure ***3-10***). Previous studies by Schaaf, et al., (2006), solely relied on root length which was found to be a difficult observation to note here. Although root length for *ireg2* plants grown in the presence of Ni in Fe-deficient conditions was shorter, natural variation in root length as well as the fragility of *Arabidopsis* roots further compounded the difficulty of using this variable. For example, in both wildtype and *ireg2*, root length was not very consistent as some plants would have much longer roots than other which can skew the results. Furthermore, when taking mean root biomass reduction into account, *ireg2* mutants had a 78% reduction whereas wildtype plants had a 33% reduction in the presence of Ni. This is in contrast to mean root length reduction, which showed a 30% reduction for *ireg2* and 4% reduction for wild type. This larger reduction in root biomass production makes it a more robust measure of phenotype characteristics than root length, especially after three independent experiments showed similar results. When comparing results from Schaaf, et al., (2006) to what was seen here, they reported a 10 cm per plant root length for wildtype plants and 6 cm per plant for *ireg2* mutants grown in the absence of Fe and presence of 20µM Ni. Here an average of 8.5 cm and 6.0 cm for wildtype and *ireg2* respectively were observed at 10 µM Ni. The differences in root length between the two sets of results for the wildtype plants are most likely due to an array of experimental conditions including the accuracy of measuring the root tissue, but the similarity of the *ireg2* results suggest that the *ireg2* mutants used here were behaving in the same way as those used by Schaaf, et al., (2006).

The dimethylglyoxime staining provided a further visual assessment of the degree to which Ni had been sequestered, and also revealed an altered phenotype in the *ireg2* mutant. When

compared to the wild type under Fe-starvation and the presence of Ni, *ireg2* knock-out mutants had no red staining, suggesting that whilst Ni was able to enter the roots (presumably via *IRT1*), it was not sequestered there and was most likely translocated to other organs in an attempt to avoid root Ni toxicity. This absence of red staining is what would be expected if *ireg2* is real knock out for the gene. When the leaves were examined for any noticeable differences in red precipitation, no differences were seen as both plants seemed to have sequestered Ni in their trichomes to the same degree, though this was not a quantitative assay, making it difficult to determine if any more or less was sequestered. The assays developed here and the confirmation of homozygosity of SALK 074442C makes it plausible to pursue *IREG2* complementation with the caveat that the RT-PCRs should be repeated to confirm that there is indeed zero expression for the homozygous line.

The genotyping for the *irt1* T-DNA insertion lines clearly confirmed that SALK 097869C is homozygous knockout for *IRT1* whereas SALK 054554C was confirmed as being a segregating line (Figure 3-3). Nishida, et al., (2011) reported the use of SALK 054554C, stating that *AtIRT1* expression was, however, detected in this mutant line since it was predicted to produce a truncated *AtIRT1* lacking biological activity. They further state, however, that SALK 054554C displayed hypersensitivity towards Fe starvation as previously reported for other lines (Vert, et al., 2002). However, given that SALK 054554C was a segregating line, it was decided that SALK 097869C should be used for further phenotyping. Although the genotyping confirmed the homozygosity of SALK 097869C, no phenotypic differences in root or shoot biomass production could be observed between this line and wildtype plants, nor was any evidence of lethality obtained. This is in contrast to the results of Vert, et al., (2002) who demonstrated a severe lethality of *irt1* when grown in the absence of Fe, stating that their *irt1* failed to develop beyond the four-to-six-true-leaf stage and were subsequently sterile. Furthermore, (Nishida, et al., 2011) also observed a similar lethal phenotype for the two lines they used (SALK 024525 and SALK 054554C), implying that this particular phenotypic observation is a reliable one. Thus, while SALK-097869C was demonstrated to be homozygous, and the T-DNA is inserted in the first 100 bp upstream of the transcriptional start site of the *IRT1* gene, and thus may be expected to abolish expression, the phenotypic characterisation of this line provided no evidence that it was indeed behaving as an *IRT1* knock-out line would be expected. A potential problem in obtaining strong *irt1* phenotypes from stock centres such as SALK could be due to the fact that these mutants require higher than usual Fe to complete their life cycle and produce seeds.

Unfortunately, it was not possible to determine whether *IRT1* was indeed expressed in SALK-097869C or not. The RT-PCRs were unconvincing as although a cDNA product was noted for one wildtype individual grown in the absence of Fe, the intensity of the amplicon was very

faint. In addition, the actual PCR required a lot of optimisation to work, redesigning of primers, temperature and $MgCl_2$ gradients, as well as testing the efficiency of various thermostable polymerases and cycling parameters, to name a few. Independent repeats of the RT-PCR produced the same cDNA band for that one wildtype individual but failed to produce any other amplicons for any of the other individuals (wildtype and *irt1*). Further gene expression analysis e.g. by means of quantitative PCR, would need to be done to assess the true degree of expression of this gene. It may possibly be the case where on an agarose gel, where the least amount of DNA detectable is 10ng, the amount of DNA generated through end-point PCR is perhaps not sufficient to fully assess any real expression. These end-point RT-PCR approaches simply indicate a presence or absence and do not provide any quantitative information on true gene expression.

With the above suggesting that although SALK 097869C was a homozygous mutant it did not behave phenotypically as expected, *irt1* mutants were grown in the presence of Fe for two weeks then deprived of Fe for ten days and supplemented with 10 µM Ni. From this, no apparent differences in wild type or *irt1* staining was noted (data not shown), suggesting that this particular T-DNA insertion line was not a true knock-out and that it still had a potentially fully functional *IRT1.* In hindsight, it would have been fruitful to have also characterised SALK 054554C to determine whether it would have made a better candidate for complementation, but this would have required isolating homozygous plants from the segregating population.

In terms of the cloning work done, all four of the entry clones and one of the two expression clones were successfully generated. Optimisation of various aspects involved in the cloning were essential in generating the successful clones. The initial *att*B-containing PCR reactions required some optimisation owing to the larger length of the *attB*-containing primers themselves and their higher than usual guanine content. The amplification of *A. thaliana* promoters proved the most difficult to optimise as these promoters tend to be either AT-rich or have many tandem repeats or both. The use of a two-step PCR, which entails combining the primer annealing and thermostable polymerase extension steps together, resulted in full amplification of these promoters. It was then further established in the laboratory that the two-step PCR was not only effective at amplifying promoters used in herein but was just as effective at amplifying DNA which under standard conditions failed to amplify. When the initial BP-Clonase® II reactions were done, low transformation success was observed compared to the Multi-site Gateway® Pro manual. The manual suggests that when the entire 11 µL reaction is transformed and plated, one could expect >1500 colonies if the reaction had been efficient. From the experience gained here, this was not the case. Although the total amount used for the transformation never exceeded 5 µL of the reaction, very few colonies (<20 CFUs on 90 µL BP Clonase® II reaction experimental plate) grew, not all of which were successful clones.

Furthermore, it became apparent that the *ccdB* selection of the pDONR221 vectors was not working at 100% efficiency as many false positives were found which emphasises the unreliability of colony-based PCR using only gene-specific primers. PCR fragments which had not been successfully recombined into the pDONR221 vector can result in many false positives when only using gene specific primers in colony PCR. In hindsight, following the initial screening of the first BP-Clonase® II colonies and the subsequent discovery that the *ccdB* was not providing adequate selection, a pDONR221 vector-specific primer could have been used in combination with a gene specific primer to screen colonies. A lot of time spent on troubleshooting the reactions could have been saved. Nevertheless, all the entry clones were eventually produced, and sequence verified bar one.

The BP-Clonase® II for *ScIREG2* failed to work on numerous occasions for reasons unknown. Initially it was thought that the PCR primers might have been incorrectly made. As such, new primers were ordered, and the process repeated. One of those attempts resulted in less than 10 CFUs on the 90 µL BP Clonase® II experimental plate, one of which gave a positive colony PCR result and was sent for sequencing. However, the sequencing data revealed that point mutations were introduced into the PCR fragment during initial amplification. These point mutations resulted in amino acid changes which could result in altered functioning of the transport protein which would have made any conclusion drawn from complementation potentially erroneous or invalid. However, when the process was subsequently repeated (including initial PCR amplification from cDNA) the same point mutations were found in a successful clone. It is extremely unlikely that a high-fidelity *Taq* with an error rate of 1 per 3.6 x 10^6 nucleotides is able to introduce three point mutation in a 1561 bp amplicon when the same *Taq* had failed to incorporate a single mutation in a multitude of other amplicons. It is therefore likely that the apparent point mutations observed are in fact the result of genetic variation between the *S. coronatus* plants used to generate cDNA here, and that sequenced for the *de novo* transcriptome assembly. It is unfortunate that the two-fragment clone for *IREG2* could not be generated since the assays developed to test complementation would have shown whether *S. coronatus IREG2* would be able to complement the *A. thaliana ireg2* mutant.

Given that the destination vector had to be Gateway compatible, this limited the available choices to three, of which only one proved successful. The suitability of the destination vector was based on five criteria: 1) it had to be a binary plant expression vector; 2) it had to have no 35S promoter capable of driving gene expression in plants; 3) it had to be Gateway compatible; 4) it had to have an antibiotic selection other than kanamycin; and 5) it had to give few to no false positives i.e. an effective *ccdB* gene. The initial destination vectors met many of the above criteria, however, pB7WG yielded many CFUs when transformed into *E.coli*

DH5α, suggesting its *ccdB* selection was not functioning optimally and given the low transformation rate of the BP Clonase® II reactions it would have made it unlikely to identify any two-fragment recombinant clones with such a high level of background. pEARLYGATE302, however, did not have the same problem as observed for pB7WG i.e. the *ccdB* gene was effective, but it had kanamycin resistance. To attempt to overcome the identical antibiotic resistance issue the entry clones were linearized using *PvuI* (results for these experiments not shown). However, when those linearized pDONR221 clones where transformed into *E.coli* DH5α cells, approximately 80 CFUs were counted on plates, indicating that digestion had not been 100% effective. As such, the final solution was to generate a vector that would meet all five criteria. The pB2WG7.0 vector (normally used to drive strong expression of a transgene due to the presence of the CaMV 35S promoter) was used as starting material and the 35S promoter was successfully removed by RE digest and the vector was successfully recircularised, as verified by diagnostic digests using *PstI* and *NcoI*. Those vectors where digests showed that recirculation was successful, were sent for DNA sequencing to validate that the CaMV 35S promoter had indeed been removed. Once this was verified, the newly generated pB2NADA vector was then transformed into *E. coli* DH5α cells to assess that its native *ccdB* was still intact and fully functional. Once all this was done, the next problem to solve was how to screen for a successful two-fragment clone as there are no positive controls that one could use, bar having to synthesise one. A combination of PCRs and restriction enzymes digests were used to identify potential two-fragment clones. Restriction enzymes were carefully chosen such that they had recognition sites in *pAtIRT1* and *ScIRT1* and would produce a banding pattern different to that of the empty pB2NADA destination vector (Figure 3-18). A combination of *attB*-containing *pAtIRT1* forward and *ScIRT1* reverse primers yielded an amplicon of the expected size that could only be generated if the two-fragment recombination reaction had worked. The results from the digests and the PCR reactions identified one potential two-fragment recombinant clone which was then sent for DNA sequencing which revealed the presence of the *Arabidopsis* promoter. Additional sequencing is required to confirm the presence of the *ScIRT1* ORF. It is unfortunate though that the two-fragment clone for the complementation had been generated but the homozygous line into which it was meant to be inserted did not appear to be an actual *IRT1* knock-out line.

The Gateway cloning strategy used here is very effective when one only has a single fragment that is shuttled from donor to destination vector. However, it becomes considerably more problematic when multiple fragments are introduced. Gateway® cloning technology relies on the homologous recombination between modified *att* sites derived from *E. coli* bacteriophage λ, resulting in some cases in a single base pair change, facilitated by engineered enzymes, λ integrase and *E. coli* integration host factor proteins (BP Clonase® II enzyme mix) and λ

integrase and excisionase proteins (LR Clonase® II+ enzyme mix). Although it is not a peer-reviewed source for information, *ResearchGate.net* provides a suite of troubleshooting queries and advice. A common theme encountered whilst troubleshooting the Gateway® cloning was that the initial BP Clonase ® II for multiple fragment recombination and multiple fragment LR Clonase® II reactions was not working as efficiently as they should. Researchers report on a number of problems encountered, such as no recombinant clones after BP Clonase® II reactions, i.e. empty entry vectors; zero to low transformant numbers on selection plates following either BP or LR Clonase® II+ transformations; and multiple recombinant clones with missing fragments for multiple fragment recombination LR Clonase® II+ reactions, to name a few. Under optimal conditions (25°C incubation for 16 hours at 10 fmol per entry clone and 20 fmol for the destination vector in a total reaction volume of 8 µL using 2 µL of LR Clonase® II+ enzyme), the Multi Site Gateway® Pro manual and ThermoFisher Scientific's webpage state that successful reactions (2-4 µL of LR Clonase® II+ reaction transformed into 50µL *E.coli* One Shot® MACH1 cells) can produce "hundreds" of antibiotic resistant colonies "with a cloning efficiency of 90%" for two-fragment reactions, 70-90% for three-fragment reactions, and 30-90% for four-fragment reactions. It should, however, be noted that the amount that is transformed for three- and four-fragment reactions is considerably more than for two-fragment reactions. The Multisite Gateway® Pro manual goes so far as to state that under optimal conditions 'experienced' persons can simultaneously perform the BP- and LR-Clonase® II+ reaction in one tube and whilst this has been done, by Liang, et al., (2018) for example, for single fragments, multiple fragment screening would be too time-consuming. Nevertheless, this hypothetical efficiency was never observed, not even for the BP Clonase® II reactions. Whilst there is no argument that the technology works and remains one of the leading cloning technologies for shuttling entry clones, its effectiveness, however, is under question. There is a plethora of studies where Gateway technology was used to generate multiple fragment clones, but very little on its efficiency. One study, for instance, only considered the effects of the residual *att* sites after an expression clone had been made (Perehinec, et al., 2007) but did not discuss if they had encountered any problems throughout the use of Gateway®. Several other studies such as (Chee & Chin, 2015; Wang, et al., 2014) only consider the laboriousness of having to do two bacterial transformations before having an expression clone and the costs involved in using Gateway technology, but do not discuss any other associated drawbacks or issues with using the technology. In essence, what was expected was seamless cloning, however, what was observed was technology that required too much optimisation.

There are many other cloning technologies that could have been employed other than Gateway®. Traditional restriction enzyme and ligase could have been used, but this technology relies on having unique enzyme recognition sites on your inserts and vectors which is not

always easily accomplished and can become very expensive when dealing with the assembly of multiple fragments. Gibson assembly is another popular cloning technology that relies on three steps to assemble DNA fragments under isothermal conditions; 1) a 5' exonuclease overhangs, 2) polymerase fills and anneals single strand regions, and 3) DNA ligase seals the newly annealed regions (Gibson, et al., 2009). This technology relies on generating complementary regions between fragments, similar to the *att* sites of Gateway®. GoldenBraid is another system which is based on type IIS restriction enzymes whose position on the destination vector introduces a double loop, also referred to as a braid, which ultimately allows for the fragment assembly (Sarrion-Perdigones, et al., 2011). The rationale for using the Gateway® technology was that the vectors and enzymes required to carry out the cloning were already present and by the time that it was realised that there were flaws in this technology, it was already too late to adopt a different cloning strategy. A key point to note here is that no one technology is superior to another as it is highly dependent on the application

CONCLUDING REMARKS

The dissertation as whole accomplished many of the aims as set out but suffered many unforeseen hindrances. Although the homozygosity of several T-DNA was confirmed using conventional PCR-based genotyping, genotype does not necessarily translate into phenotype, as proved to be the case for *irt1*. In contrast, an *ireg2* mutant (SALK_074442C) appears to be a knock-out line, and displays a robust phenotype (reduced root biomass production in the presence of Ni under Fe-deficient conditions) but the expression clone could not be generated, whereas the homozygous *irt1* T-DNA line that was identified did not present the expected phenotypes (and it is unclear that it is indeed a knock-out line), but its expression clone had been generated. It is unfortunate that no complementation was performed due to time constraints. A lot of time had been spent on optimising the cloning, which proved to be much less effective as expected. An important aspect to consider is that the foundations for future complementation work has been laid. Future work should focus on validating *irt1* phenotypes and homing in on elucidating whether the homozygous *irt1* line identified here has zero expression for *IRT1*. Alternatively, other potential homozygous lines could be investigated. Quantitative RT-PCR can be used to assess the degree to which *IRT1* and *IREG2* are being expressed in response to exogenous Fe levels as the RT-PCR experiments used herein failed to provide any conclusive gene expression results. Localisation by means of fluorescent tagging of the *S. coronatus* protein coding regions could shed light on whether *ScIRT1* and *ScIREG2* localise to the same subcellular locations as their *Arabidopsis* homologs do. These localisation experiments are also useful in elucidating function. The genetic variability of *S. coronatus IREG2* needs to be further characterised, as it is impossible to ascertain if the ‘point

mutations' which were observed are definitively *Taq* errors or just genetic variation, as seems likely.

REFERENCES

Alloway, B J. 2013. *Heavy Metals and Metalloids as Micronutrients for Plants and Animals.* Vol. 22, in *Heavy Metals in Soils, Environmental Pollution*, edited by B Alloway . Dordrecht: Springer.

Amir, H, N Perrier, F Rigault, and T Jaffre´. 2007. "Relationship between Ni-hyperaccumulation and mycorrhizal status of different endemic plant species from New Caledonian ultramafic soils." Plant Soil 293: 23-35. DOI: 10.1007/s11104-007-9238-0

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

Baker, A JM, S P McGrath, R D Reeves, and J AC Smith. 1999. "A review of the biological resource for possible exploration in the phytoremediation of metal-polluted soils." In *Phytoremediation of Contaminated Soil and Water*, edited by N Terry and G S Banuelos, 85-107. CRC Press LLC.

Bauer, P, and M L Guerinot. 2007. "FIT, the FER-LIKE IRON DEFICIENCY INDUCED TRANSCRIPTION FACTOR in *Arabidopsis*." *Plant Physiology and Biochemistry* 45 (5): 260-261.DOI: *10.1016/j.plaphy.2007.03.006*

Benes, I, K Schreiber, H Ripperger, and A Kircheiss. 1983. "Metal complex formation by nicotianamine, a possible phytosiderophore." *Experientia* 39 (3): 261-262.DOI: https://doi.org/10.1007/BF01955293

Borhldi, A. 2001. "Phylogenetic trends in Ni-accumulating plants." *South African Journal of Science* 97. <http://www.scielo.org.za/scielo.php?script=sci_arttext&pid=S0038-23532008000600022&lng=en&nrm=iso>. ISSN 1996-7489.

Boyd, R S. 2004. "Ecology of metal hyperaccumulation." New Phytologist 162: 563-567.DOI: https://doi.org/10.1111/j.1469-8137.2004.01079.x

Boyd, R S, and S N Martens. 1992. "The raison d'etre for metal hyperaccumulation." In *The vegetation of ultramafic (Sepentine) soils*, edited by A JM Baker, J Proctor and R D Reeves, 279-289. Hampshire: Andover.

Boyd, R S, and S N Martens. 1998. "The significance of metal hyperaccumulation for biotic interactions." *Chemoecology* 8: 1-7. DOI: 10.1007/s000490050002

Boyd, R S, M A Davis, and K Balkwill. 2008. "Elemental patterns in Ni hyperaccumulating and non-hyperaccumulating ultramafic soil populations of *Senecio coronatus*." *South African Journal of Botany* 74 (1): 158-162.DOI: https://doi.org/10.1016/j.sajb.2007.08.013

Brooks, R R, J Lee, R D Reeves, and T Jaffre. 1977. "Detection of nickeliferous rocks by analysis of herbarium speciments of indicator plants." *J Geochem Explor* 7: 49-57.DOI: https://doi.org/10.1016/0375-6742(77)90074-7

Callahan, D L, A J.M Baker, S D Kolev, and A G Wedd. a 2006. “Metal ion ligands in hyperaccumulating plants.” Biological Inorganic Chemistry 11: 2-12.DOI: 10.1007/s00775-005-0056-7

Callahan, D L, S D Kolev, R A O'Hair, D E Salt, A J Baker, M Lebrun, and P Czernic. b 2006. “Relationships of nicotianamine and other amino acids with nickel, zinc and iron in *Thlaspi* hyperaccumulators.” Journal of Experimental Botany 57: 4111-4122. DOI: https://doi.org/10.1111/j.1469-8137.2007.02216.x

Carter, E L, N Flugga, J L Boer, S B Mulrooney, and R P Hausinger. 2009. “Interplay of metal ions and urease.” *Metallomics* 1 (3): 207-221. doi: 10.1039/b903311d.

Campbell, M K, and S O Farrell. 2012. “Photosynthesis.” In *Biohemistry*, 638. Brooks/Cole Cengage Learning.

Chee, J Y, and C F Chin. 2015. “Gateway Cloning Technology: Advantages and Drawbacks.” *Clon Transgen* 4 (138). DOI: 10.4172/2168-9849.1000138

de la Torre, V S G, C Majorel-Loulergue, D A Gonzalez, L Soubigou-Taconnat, G J Rigail, Y Pillon, L Barreau, et al. 2018. “Wide cross-species RNA-Seq comparison reveals a highly conserved role for Ferroportins in nickel hyperaccumulation in plants.” *bioRxiv.* DOI: https://doi.org/10.1101/420729

Deng, Teng-Hao-Bo, A ven der Ent, Ye-Tao Tang, T Sterckeman, G Echevarria, Jean-Louis Mortel, and Rong-Liang Qui. 2017. “Nickel hyperaccumulation mechanisms: a review on the current state of knowledge.” *Plant Soil*. DOI: 10.1007/s11104-017-3539-8

Dhatterwal, P, S Mehrotra , and R Mehrotra. 2017. “Optimization of PCR conditions for amplifying an AT-rich amino acid transporter promoter sequence with high number of tandem repeats from Arabidopsis thaliana.” BMC Research Notes 10: 638.DOI: 10.1186/s13104-017-2982-1

Dilwroth, M J, A D Robson, and D L Chatel. 1979. “Cobalt and nitrogen fixation in Lupinus angustifolius L. II. Nodule formation and function.” New Phytologist 83: 63-79. DOI: https://doi.org/10.1111/j.1469-8137.1979.tb00727.x

Dutta, A, M Mitra, P Agarwal, K Mahapatra, S De, U Sett, and S Roy. 2018. “Oxidative and genotoxic damages in plants in response to heavy metal stress and maintenance of genome.” *Plant Signaling & Behavior* 13 (8). DOI: 10.1080/15592324.2018.1460048

Edwards, K, C Johnstone, and C Thompson. 1991. “A simple and rapid method for the preparation of plant genomic DNA for PCR analysis.” *Nucleic Acids Research* 19 (6): 1349. DOI: 10.1093/nar/19.6.1349

Eide, D, M Broderius, J Fett, and M L Geurinot. 1996. “A novel iron-regulated metal transporter from plants identified by functional expression in yeast.” *PNAS* 93 (11): 5624-5628. DOI: 10.1073/pnas.93.11.5624

El-Shintinawy, F, and A El-Ansary. 2000. “Differential effect of Cd2+ and Ni2+ on amino acid metabolism of soybean seedlings.” *Biol. Plant* 43: 79-84. DOI: https://doi.org/10.1023/A:1026507114354

Epstein, E. 1965. “Typical concentrations sufficient for plant growth.” In *Plant Biochemistry*, by J Bonner and J E Varner, 438-466. London: Academic Press.

Erdei, S, A Hegedus, A Hauptmann, J Szali, and G Horvath. 2002. “Heavy metal induced physiological changes in the antioxidative response system.” *Acta Biologica Szegediensis* 46: 89-90.: DOI: http://www.sci.u-szeged.hu/ABS

Ghasemi, R, S M Ghaderian, and U Krämer. 2009. “Accumulation of Nickel in Trichomes of a Nickel Hyperaccumulator Plant, *Alyssum inflatum*.” Northeastern Naturalist 16. DOI:10.1656/045.016.0507.

Gibson, D G, L Young, RY Chaung, J C Venter, C A Hutchinson, and H O Smith. 2009. “Enzymatic assembly of DNA molecules up to several hundred kilobases.” *Nature Methods* 6: 343-345. doi:https://doi.org/10.1038/nmeth.1318

Graham, L E, J M Graham, and L W Wilcox. 2014. “Photosynthesis and Respiration.” In *Plant Biology*, 116. Essex: Pearson Education Limited.

Halliwell, B, and J MC Gutteridge. 1984. “Oxygen toxicity, oxygen radicals, transition metals and disease.” *Biochemical Journal* 219: 1-14. DOI: 10.1042/bj2190001

Hong, C L, Y B Jia, X E Ynag, Z L He, and P J Stofella. 2008. “Assessing lead thresholds for phytotoxicity and potential dietary toxicity in selected vegetable crops.” *Bull. Environ. Contam. Toxicol.* 80 (3): 56-61. DOI: 10.1007/s00128-008-9375-z.

Hopkins, W G, and N P.A. Hünter. 2009. “Introduction to Plant Physiology.” Chap. 4, 61-76. John Wiley & Sons.

Ingle, R A, S T Mugford, J D Rees, M M Campbell, and J AC Smith. 2005. “Constitutively high expression of the histidine biosynthetic pathway contributes to nickel tolerance in hyperaccumulator plants.” *Plant Cell* 17: 2089-2106. DOI: https://doi.org/10.1105/tpc.104.030577

Jaffré, T, R R Brooks, J Lee, and R D Reeves. 1976. “*Sebertia acuminata*: A hyperaccumulator of Nickel from New Caledonia.” *SCIENCE* 193: 579-580. DOI: 10.1126/science.193.4253.579

Jaffré , T, R D Reeves, A J Baker, H Schat, and A van der Ent. 2018. “The discovery of nickel hyperaccumulation in the New Caledonian tree *Pycnandra acuminata* 40 years on: An introduction to a Virtual Issue.” New Phytologist 218: 397-400. DOI: 10.1111/nph.15105.

Jhee, E M, R S Boyd, M D Eubanks, and M A Davis. 2006. “Nickel hyperaccumulation by *Streptanthus polygaloides* protects against the folivore *Plutella xylostella* (Lepidoptera: Plutellidae).” *Plant Ecol.* 183: 91-104. DOI: https://doi.org/10.1007/s11258-005-9009-z

Kabata-Pendias, A, and H Pendias. 2001. *Trace Elements in Soils and Plants.* 3rd. London: Boca Raton, Fla CRC Press.

Keilig, K, and J Ludwig-Müller. 2009. “Effect of flavonoids on heavy metal tolerance in Arabidopsis thaliana seedlings.” *Botanical Studies* 50: 311-318.

Kerkeb, L, and U Krämer. 2003. “The Role of Free Histidine in Xyelm Loading of Nickel in *Alyssum lesbiacum* and *Brassica juncea*.” *Plant Physiology* 131 (2). DOI: https://doi.org/10.1104/pp102.010686

Kevrešan, S, N Petrović, M Popović, and J Kandrač. 1998. “Effect of heavy metals on nitrate and protein metabolism in sugar beet.” *Biologi Plantarum* 41: 235-240.DOI: https://doi.org/10.1023/A:1001818714922

Krämer, U. 2010. “Metal hyperaccumulation in Plants.” *Annual Review of Plant Biology* 61: 517-534. DOI: 10.1146/annurev-arplant-042809-112156.

Krämer, U, J D Cotter-Howells, J M Charnock, A J Baker, and J A C Smith. 1996. “Free histidine as a metal chelator in plants that accumulate nickel.” *Nature* 379: 635-638.DOI: 10.1038/379635a0

Krupta, Z, and T Baszynski. 1995. “Some aspects of heavy metals toxicity towards photosynthetic apparatus- direct and indirect effects on light and dark reactions.” *Acta Physiol Plant* 17: 177-190. ISBN978-93-81141-00-7

Küpper , H, E Lombi, F J Zhao, G Wieshammer, and S P McGrath. 2011. “Cellular compartmentation of nickel in the hyperaccumulators *Alyssum lesbiacum*, *Alyssum bertolonii* and Thlaspi goesingense.” Experimental Botany 2291-2300. DOI:10.1093/jexbot/52.365.2291.

Lee, J, R D Reeves, R R Brooks, and T Jaffre. 1977. “Isolation and identification of a citrato-complex of nickel from nickel-accumulating plants.” *Phytochemistry* 16: 1503-1505. DOI: https://doi.org/10.1016/0031-9422(77)84010-7

Lee, J, R D Reeves, R R Brooks, and T Jaffre. 1978. “The relation between nickel and citric acid in some nickel-accumulating plants.” *Phytochemistry* 17: 1033-1035.DOI: https://doi.org/10.1016/S0031-9422(00)94274-2

Liang, X, L Peng, CH Baek, and F Katzen. 2018. “Single step BP/LR combined Gateway reactions.” *Biotechniques* 55 (5).DOI: 10.2144/000114101

Macnair, M R. 2003. “The hyperaccumulation of metals by plants.” *Advances in Botanical Research* 40: 63-105. DOI: https://doi.org/10.1016/S0065-2296(05)40002-6

Maksymiec, W. 1997. “Effect of copper on cellular processes in higher plants.” *Photosynthetica* 34: 321-342. DOI: 10.1023/A:1006818815528

Martel, A E, and R M Smith. 1974. *Crtitical stability constants.* New York: Plenum.

Martens, S N, and R S Boyd. 2002. “The defensive role of Ni hyperaccumulation by plants: a field experiment.” *American Journal of Botany* 89: 998-1003.DOI: 10.3732/ajb.89.6.998

Martens, S N, and R S Boyd. 1994. “The ecological significance of nickel hyperaccumulation: a plant chemical defense.” *Oecologia* 98: 379-384.DOI: www.jstor.org/stable/4220712

McKie, A T, P Marciani, A Rolfs, K Brennan, K Wehr, D Barrow, S Miret, et al. 2000. “A novel duodenal iron-regulated transporter, IREG1, implicated in the basolateral transfer of iron to the circulation.” *Molecular Cell* 5 (2): 299-309.DOI: 10.1016/s1097-2765(00)80425-6

Meier, S K, N Adams, M Wolf, K Balkwill, A M Muasya, C A Gehring, J M Bishop, and R A Ingle. 2018. “Comparative RNA-seq analysis of nickel hyperaccumulating and non-hyperaccumulating populations of *Senecio coronatus* (Asteraceae).” *The Plant Journal* 95: 1023-1038. DOI: 10.1111/tpj.14008

Meindi, G A, D J Bain, and T-L Ashman. 2014. “Nickel accumulation in leaves, floral organs and rewards varies by serpentine soil affinity.” The Annals of Botany. DOI:10.1093/aobpla/plu036.

Mengoni, A, A JM Baker, M Bazzicalupo, R D Reeves, N Adigüzel, E Chianni, F Galardi, R Gabbrielli, and C Gonnelli. 2003. “Evolutionary dynamics of nickel hyperaccumulation in Alyssum revealed by ITS nrDNA analysis.” New Phytologist 159: 691-699. DOI: https://doi.org/10.1046/j.1469-8137.2003.00837.x

Merlot, S, L Hannibal, S Martins, L Merinelli, H Amir, M Lebrun, and S Thomine. 2014. “The metal transporter PgIREG1 from the hyperaccumulator *Psychorita gabriellae* is a candidate gene for nickel tolerance and accumulation.” *Journal of Experimental Botany* 65 (6): 1551-1564. DOI:10.1093/jxb/eru025.

Mesjasz-Przybylowicz, J, and W J Przybylowicz. 2001. “Phytophagous insects associated with the Ni-hyperaccumulating plant *Berkheya coddii* (Asteraceae) in Mpumalanga, South Africa.” *South African Journal of Science* 97: 596-598. DOI: https://hdl.handle.net/10520/EJC97236

Michalak, A. 2006. “Phenolic Compounds and Their Antioxidant Activity in Plants Growing under Heavy Metal Stress.” *Polish Journal of Environmental Studies* 15 (4): 523-530.

Mincey, K A, and R S Boyd. 2018. “Elemental defense of nickel hyperaccumulator seeds against a generalist insect granivore.” *Ecological Research* 33: 561-570. DOI: https://doi.org/10.1007/s11284-018-1583-3

Mincey, K A, P A Cobine, and R S Boyd. 2018. “Nickel hyperaccumulation by S*treptanthus polygaloides* is associated with herbivory tolerance.” *Ecological Researcg* 33: 571-580. DOI: https://doi.org/10.1007/s11284-018-1569-1

Mohanty, N, I Vass, and S Demeter . 1989. “Impairment of photosystem 2 activity at the level of secondary electron acceptor in chloroplasts treated with cobalt, nickel and zinc ions.” Physiol. Plant 76: 386-390. DOI: https://doi.org/10.1111/j.1399-3054.1989.tb06208.x

Mohseni, R, S M Ghaderian, R Ghasemi, and H Schat. 2018. “Differential effects of iron starvation and iron excess on nickel uptake kinetics in two Iranian nickel hyperaccumulators, *Odontarrhena bracteata* and *Odontarrhena inflata.*” *Plant and Soil* 425: 153-162. DOI: https://doi.org/10.1007/s11104-018-3666-x

Montargès-Pelletier, E, J Mesjasz-Przybylowicz, A Barnabas, G Echevarria, V Briois, T P Sechogela, and W J Przybylowicz. 2011. “Do hyperaccumulators develop specific chelates for nickel transport and storage? The cases of *Senecio coronatus* and *Berkheya coddii*.” *In Proceedings of the 11th International Conference on Biogeochemistry of Trace Elements, Florence, Italy.* DOI: https://prodinra.inra.fr/record/345036

Montargès-Pelletier, E, V Chardot, G Echevarria, L J Michot, A Bauer, and J-L Mortel. 2008. “Identification of nickel chelators in three hyperaccumulating plants: an X-ray spectroscopic study.” *Phytochemistry* 69: 1695-1709.DOI: https://doi.org/10.1016/j.phytochem.2008.02.009

Morrey, D R, K Balkwill, and M J Balkwill. 1989. “Studies on serpentine flora : Preliminary analyses of soils and vegetation associated with serpentinite rock formations in the south-eastern Transvaal.” *South African Journal of Botany* 55 (2): 171-177. DOI: https://doi.org/10.1016/S0254-6299(16)31203-0

Morrissey, J, I R Baxter, J Lee, L Li, B Lahner, N Grotz, J Kaplan, D E Salt, and M L Geurinot. 2009. “The Ferroportin Metal Efflux Proteins Function in Iron and Cobalt Homeostasis in *Arabidopsis*.” The Plant Cell 21 (10): 3326-3338. DOI:10.1105/tpc.109.069401.

Nishida, S, C Tsuzuki, A Kato, A Aisu, J Yoshida, and T Mizuno. 2011. “AtIRT1, the Primary Iron Uptake Transporter in the Root, Mediates Excess Nickel Accumulation in *Arabidopsis thaliana*.” *Plant* and Cell Physiology 52 (8): 1433-1442.DOI: https://doi.org/10.1093/pcp/pcr089

Pandey, N, G C Pathak, D K Pandey, and R Pandey. 2009. “Heavy metals, Co, Ni, Cu, Zn, and Cd, produce oxidative damage and evoke differential antioxidant responses in spinach.” Brazilian Journal of Plant physiology 21 (2): 103-11. DOI: http://dx.doi.org/10.1590/S1677-04202009000200003

Pelosi, P, R Fiorentini, and C Galoppini. 1976. “On the nature of nickel compounds in *Alyssum bertolonii* Desv.-II.” *Agricultural and Biological Chemistry* 40: 1641-1642. DOI: https://doi.org/10.1080/00021369.1976.10862273

Perehinec, T M, S NA Qazi, S R Gaddipati, V Salisbury, C ED Rees, and P J Hill. 2007. “Construction and evaluation of multisite recombinatorial (Gateway) cloning vectors for Gram-positive bacteria.” BMC Molecular Biology 8 (80). DOI: 10.1186/1471-2199-8-80

Persans, M W, X Yan, J-M M Patnoe, U Krämer, and D E Salt. 1999. “Molecular dissection of the role of histidine in nickel hyperaccumulation in *Thlaspi goesingense* (Hálácsy).” *Plant Phytologist* 121: 1117-1126. DOI: 10.1104/pp.121.4.1117

Pollard, A J, R D Reeves , and A J Baker. 2014. “Facultative hyperaccumulation of heavy metals and metalloids.” *Plant Sci* 8 (17): 217-218. DOI: https://doi.org/10.1016/j.plantsci.2013.11.011

Rascio, Nicoletta, and Flavia Navari-Izzo. 2011. “Heavy metal hyperaccumulating plants: How and why do they do it? And what makes them so interesting?” *Plant Science* 180 (2): 169-181.DOI: https://doi.org/10.1016/j.plantsci.2010.08.016

Rees, J D, R A Ingle, and J AC Smith. 2009. “Relative contributions of nine genes in the pathway of histidine biosynthesis to the control of free histidine concentrations in *Arabidopsis thaliana*.” *Plant Biotechnology Journal* 7: 499-511. DOI: 10.1111/j.1467-7652.2009.00419.x

Reeves, R D, A J Baker, T Jaffré, P D Erskine, G Echevarria, and A van der Ent. 2018. “A global database for plants that hyperaccumulate metal and metalloid trace elements.” New Phytologist 218: 407-411. doi:10.1111/nph.14907

Robson, A D, M J Dilworth, and D L Chatel. 1979. “Cobalt and nitrogen fixation in *Lupinus angustifolius* L. I Growth, nitrogen concentrations and cobalt distribution.” *New Phytologist* 83: 53-62.DOI: 10.1111/j.1469-8137.1979.tb00726.x

Römheld, V. 2012. “Diagnosis of Deficiency and Toxicity of Nutrients.” Chap. 11 in *Mineral Nutrition of Higher Plants*, by H Marschner, edited by P Marschner, 303. Elsevier Ltd.

Ryan, K C, O E Johnson, D E Cabelli, T C Brunold, and M J Maroney. 2010. “Nickel Superoxide Dismutase: Structural and Functional Roles of Cys2 and Cys6.” *J Biol Inorg Chem* 15 (5): 795-807.DOI: 10.1007/s00775-010-0645-y

Sarrion-Perdigones, A, E E Falconi, S I Zandalinas, P Juárez, A Fernández-del-Carmen, A Granell, and D Orzaez. 2011. “GoldenBraid: An Iterative Cloning System for Standardized Assembly of Reusable Genetic Modules.” *PLOS ONE* 6 (7). DOI:: 10.1371/journal.pone.0021622.

Schaaf, G, A Honsbein, A R Meda, S Kirchner, D Wipf, and N von Wiren. 2006. “AtIREG2 Encodes a Tonoplast Transport Protein Involved in Iron-dependent Nickel Detoxification in *Arabidopsis thaliana* roots.” *The Journal of Biological Chemistry* 281 (3): 25532-25540.DOI: 10.1074/jbc.M601062200

Schwartz, M D, and M A Wall. 2001. “*Melanotrichus boydi*, a new species of plant bug (Heteroptera: Miridae: Orthotylini) restricted to the nickel hyperaccumulator *Streptanthus polygaloides* (Brassicaceae).” Pan-Pacific Entomologist 77: 39-44.

Seregin, I V, and A D Kozhevnikova. 2006. “Physiological role of nickel and its toxic effects on higher plants.” Russian Journal of Plant Physiology 53 (2): 257-277.DOI: https://doi.org/10.1134/S1021443706020178

Seregin, I V, and V B Ivanov. 2001. “Physiological aspects of cadmium and lead toxic effects on higher plants Fiziol Rast (Moscow).” Russian Journal of Plant Physiology 48: 606-630 Engl. Transl pp 523-544. DOI: https://doi.org/10.1023/A:1016719901147

Shoeran , I S, H R Singal, and R Singh. 1990. “Effect of cadmium and nickel on photosynthesis and the enzymes of the photosynthetic carbon reduction cycle in Pigeonpea (Cajanus cajan L.).” Photsynth. Res 23: 345-351. DOI: 10.1007/BF00034865.

Siebert, S J, N C Schutte, S P Bester, D M Komape, and N Rajakaruna. 2017. “Senecio conrathii N.E.Br. (Asteraceae), a new hyperaccumulator of nickel from serpentinite outcrops of the Barberton Greenstone Belt, South Africa.” Ecological Research 33: 651-658. DOI: https://doi.org/10.1007/s11284-017-1541-5

Smith, S E, and D J Read. 1997. Mycorrhizal symbiosis. Second. San Diego: Academic Press.

Taylor, S I, and M R Macnair. 2006. “Within and between population variation for zinc and nickel accumulation in two species of Thlaspi (Brassicaceae).” New Phytologist 169: 505-514. DOI: 10.1111/j.1469-8137.2005.01625.x

Turnau, K, and J Mesjasz-Przybylowicz. 2003. “Arbuscular mycorrhiza of Berkhey coddii and other Ni-hyperaccumulating members of the Asteraceae from ultramafic soils in South Africa.” Mycorrhiza 13: 217-221. DOI: 10.1007/s00572-002-0213-6

Ülker, B, E Pieter, D P Dixon, C Moffat, R Capper, N Bouché, R Edwards, D Sanders, H Knight, and M R Knight. 2008. “Getting the most out of publicly available T-DNA insertion lines.” The Plant Journal 56 (4).DOI: 10.1111/j.1365-313X.2008.03608.x

van de Mortel, J E, H Schat, P D Moerland, E Ver Loren van Themaat, S van Ent, H Blankestijn, A Ghandilyn, S Tsiatsiani, and M GM Aarts. 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 & Environment 31: 301-324. DOI: 10.1111/j.1365-3040.2007.01764.x

van de Mortel, J E, L A Villanueva, H Schat, J Kwekkeboom, S Coughlan, P D Moerland, E Ver Loren van Themaat, M Koornneef, and M GM Aarts. 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.DOI: 10.1104/pp.106.082073

Veeranjaneyulu, K, and V SR Das. 1982. "Interchloroplast localization of 65Zn and 63Ni in a Zn-tolerant plants, Ocimum basilicum Benth." J. Exp. Bot 33: 1161-1165.

Verbruggen, N, C Hermans, and H Schat. 2009. "Molecular mechanisms of metal hyperaccumulation." New Phytologist 181: 759-776. DOI: https://doi.org/10.1111/j.1469-8137.2008.02748.x

Vert, G, J-F Briat, and C Curie. 2001. "Arabidopsis IRT2 gene encodes a root-periphery iron transporter." The Plant Journal 26 (2): 181-189. DOI: https://doi.org/10.1046/j.1365-313x.2001.01018.x

Vert, G, N Grotz, F Dédaldéchamp, M L Geurinot, Jean-François Briat, and C Curie. 2002. "IRT1, an Arabidopsis Transporter Essential for Iron Uptake from the Soil and for Plant Growth." Plant Cell 14 (6): 1223-1233.DOI: 10.1105/tpc.001388

Wang, B, and Y L Qui. 2006. "Phylogenetic distribution and evolution of mycorrhizas in land plants." Mycorrhiza 16: 299-263. DOI: 10.1007/s00572-005-0033-6

Wang, J, R Xu, and A Liu. 2014. "IRDL Cloning: A One-Tube, Zero-Background, Easy-to-Use, Directional Cloning Method Improves Throughput in Recombinant DNA Preparation." PLOS ONE. DOI: https://doi.org/10.1371/journal.pone.0107907.

Wycisk, K, E J Kim, J I Schroeder, and U Krämer. 2004. "Enhancing the first enzymatic step in the histidine biosynthesis pathway increases the free histidine pool and nickel tolerance in Arabidopsis thaliana." FEBS letters 578: 128-134. DOI: 10.1016/j.febslet.2004.10.086

Zhang, Z, S Schwartz, L Wagner, and W Miller. 2000. "A greedy algorithm for aligning DNA sequences." J Comput Biol 7 (1-2): 203-14.DOI: 10.1089/10665270050081478

Zhao, Y, J Luo, W Wang, T Liu, C Han, Y Chen, and L Kong. 2016. "Selection of Reference Genes for GeneExpression Normalization inPeucedanumpraeruptorumDunn under Abiotic Stresses,Hormone Treatments and Different Tissues." *PLOS ONE* 11 (3). DOI:10.1371/journal.pone.0152356.

APPENDIX

Vector maps of genes of interest in respective cloning vectors

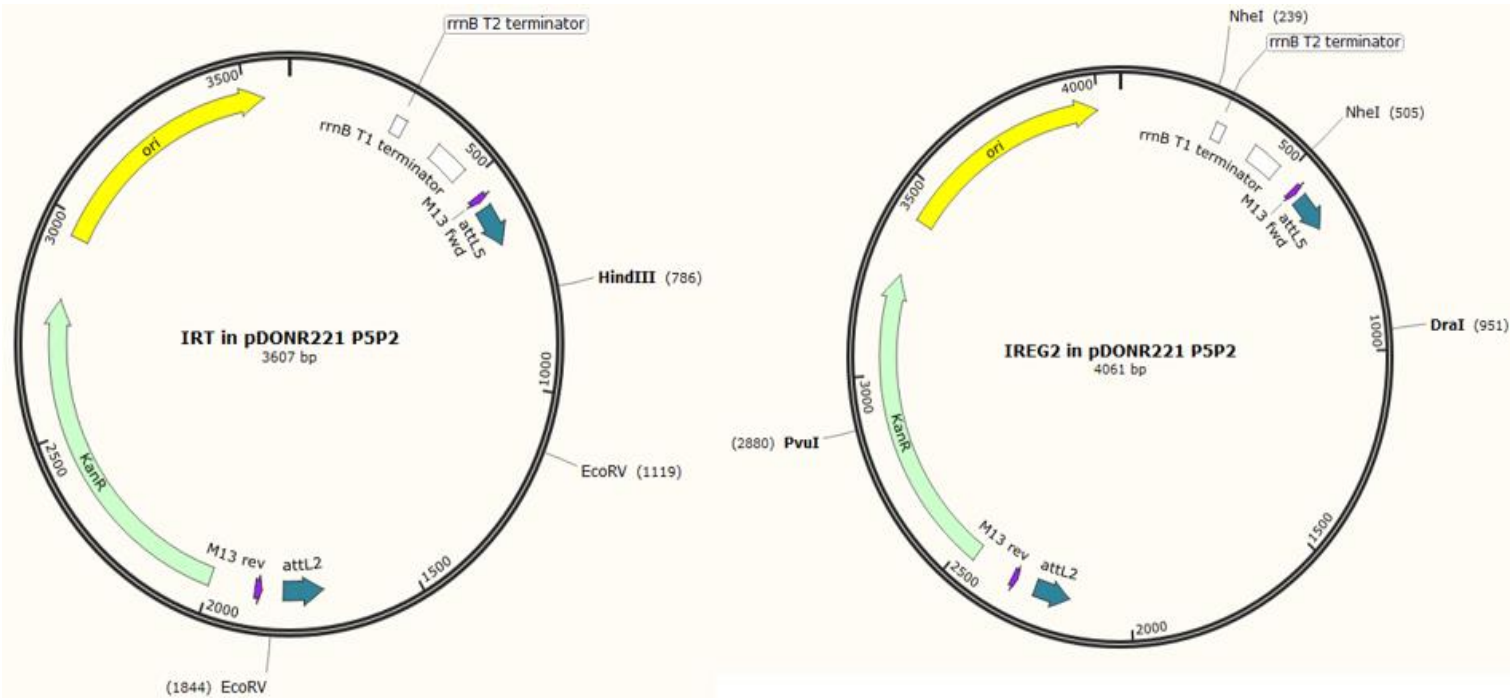


***Figure 7-1: S. coronatus BP-Clones**: Vector maps of simulated BP-Clonase ® reaction using the S. coronatus protein coding regions flank pDONR221 P5P2 donor vector. Simulation and maps generated using SnapGene*

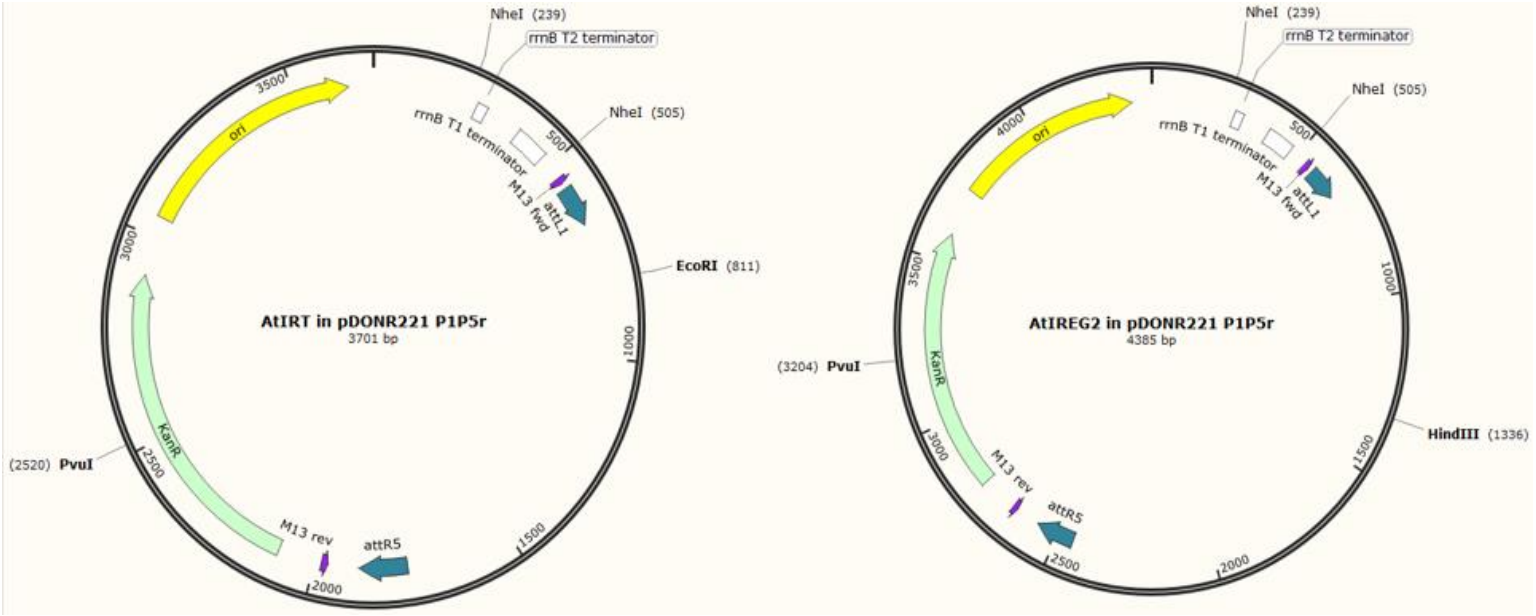


Figure 7-2: A. thaliana BP-Clones*: Vector maps of simulated BP-Clonase ® reaction using the A. thaliana promoters flanked by attB sites and p P1P5r donor vector. Simulation and maps generated using SnapGene*

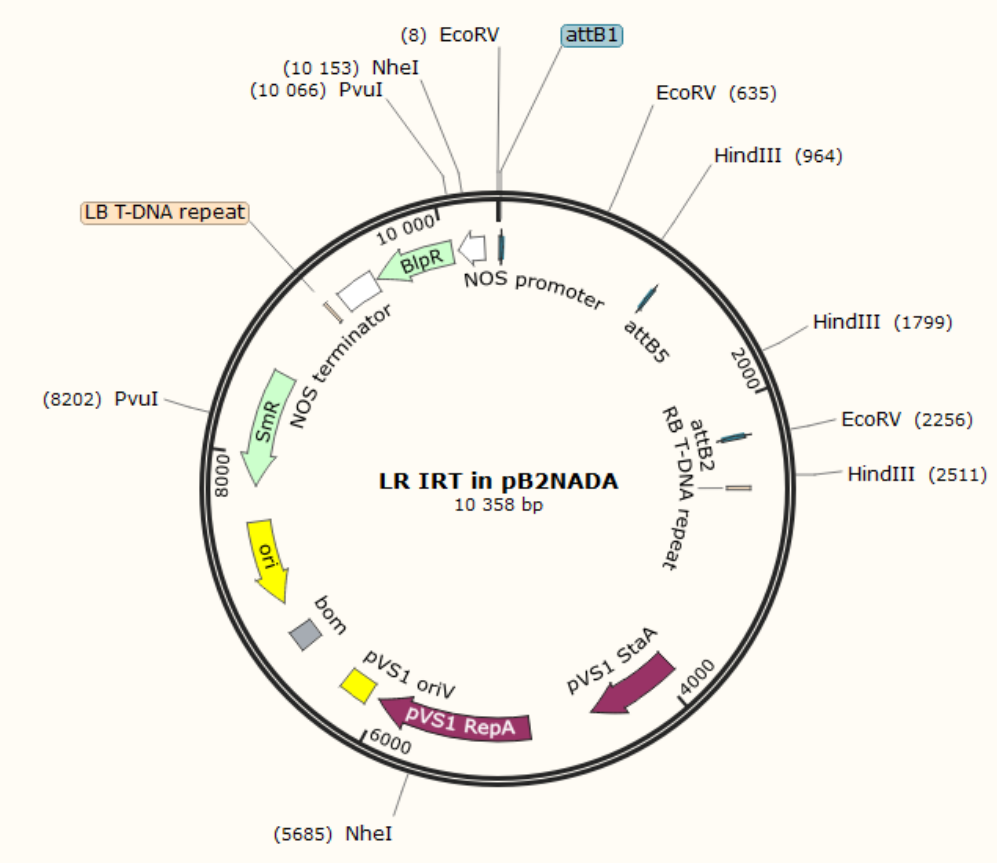


***Figure 7-3: LR-IRT Clone**: Simulated LR-Clonase® II+ using ScIRT1 in pDONR 221 P5P2 and AtIRT1 in pDONR221 P1P5r as donor vectors and pB2NADA as destination. Simulation and map generated using SnapGene*

Multiple sequence alignments for clones:

AtIRT1 in pDONR221 P1P5r

```
AtIRT1        .....GGGACAAGTTTGTACAAAAAAGCAGGCTTACATTA      36
AtIRT for   ttataatGCCAACTTTGTACAAAAAAGCAGGCTTACATTA    120
Consensus           g caa tttgtacaaaaaagcaggcttacatta

AtIRT1      AGTCAATAACCCCATAATCTTATAAATGAAGTCAGTGACT     76
AtIRT for   AGTCAATAACCCCATAATCTTATAAATGAAGTCAGTGACT    160
Consensus   agtcaataaccccataatcttataaatgaagtcagtgact

AtIRT1      ATATGTTCATGTAAAAATGCGTAAGTTCATAGCTAGATGT    116
AtIRT for   ATATGTTCATGTAAAAATGCGTAAGTTCATAGCTAGATGT    200
Consensus   atatgttcatgtaaaaatgcgtaagttcatagctagatgt

AtIRT1      CCCTTTTTTAGGAGCACATGGATTGACACATTAAACATTC    156
AtIRT for   CCCTTTTTTAGGAGCACATGGATTGACACATTAAACATTC    240
Consensus   ccctttttaggagcacatggattgacacattaaacattc

AtIRT1      ATACCCGATTTCATGGAATTCTATTTGTCATTATCCTTTC    196
AtIRT for   ATACCCGATTTCATGGAATTCTATTTGTCATTATCCTTTC    280
Consensus   atacccgatttcatggaattctatttgtcattatcctttc

AtIRT1      CCAAAAACTTAGTTGTTTCTTAGTTATTAAACCTTGAAAA    236
AtIRT for   CCAAAAACTTAGTTGTTTCTTAGTTATTAAACCTTGAAAA    320
```

```
Consensus  ccaaaaacttagttgtttcttagttattaaaccttgaaaa

AtIRT1     AAAATCAATTCAAAAATGAAAAATGACTTATATGTATGTC    276
AtIRT for  AAAATCAATTCAAAAATGAAAAATGACTTATATGTATGTC    360
AtIRT rev  AAAAATCATTCAAAAt..gAAAATGACTTATATGTATGTC     68
Consensus  aaaa   attcaaaa    aaaatgacttatatgtatgtc

AtIRT1     ATTTTTGAAAACTAGACATCAAGCTATGAAAAGAGGAATC    316
AtIRT for  ATTTTTGAAAACTAGACATCAAGCTATGAAAAGAGGAATC    400
AtIRT rev  ATTTTTGAAA.CTAGACATCAAGCTATGAAA.GAGGAATC    106
Consensus  atttttgaaa ctagacatcaagctatgaaa gaggaatc

AtIRT1     TATTTACCGAATAAAGGAGGAGGAATCTACATTCATATGT    356
AtIRT for  TATTTACCGAATAAAGGAGGAGGAATCTACATTCATATGT    440
AtIRT rev  TATTTACCGAATAA.GGAGGAGGAATCTACATTCATATGT    145
Consensus  tatttaccgaataa ggaggaggaatctacattcatatgt

AtIRT1     AGTGTTTTTGGTGACTGGATTGCAACCATACACTGAGCAC    396
AtIRT for  AGTGTTTTTGGTGACTGGATTGCAACCATACACTGAGCAC    480
AtIRT rev  AGTGTTTTTGGTGACTGGATTGCAACCATACACTGAGCAC    185
Consensus  agtgtttttggtgactggattgcaaccatacactgagcac

AtIRT1     AAGCTGTCACGTGTGATGTTGTGCTCGTAGCATAAACTAT    436
AtIRT for  AAGCTGTCACGTGTGATGTTGTGCTCGTAGCATAAACTAT    520
AtIRT rev  AAGCTGTCACGTGTGATGTTGTGCTCGTAGCATAAACTAT    225
Consensus  aagctgtcacgtgtgatgttgtgctcgtagcataaactat

AtIRT1     AAAGTATAAACCTAATTAAGCTTAAAGTTAAAAATAGTAC    476
AtIRT for  AAAGTATAAACCTAATTAAGCTTAAAGTTAAAAATAGTAC    560
AtIRT rev  AAAGTATAAACCTAATTAAGCTTAAAGTTAAAAATAGTAC    265
Consensus  aaagtataaacctaattaagcttaaagttaaaaatagtac

AtIRT1     AAGATTAAAAAGTAAACGTTTGTGAGAGTTGACTAACAGA    516
AtIRT for  AAGATTAAAAAGTAAACGTTTGTGAGAGTTGACTAACAGA    600
AtIRT rev  AAGATTAAAAAGTAAACGTTTGTGAGAGTTGACTAACAGA    305
Consensus  aagattaaaaagtaaacgtttgtgagagttgactaacaga

AtIRT1     ACTATAGAAGACGCATGGGGTTGACTTCAAGCGGTCTGCA    556
AtIRT for  ACTATAGAAGACGCATGGGGTTGACTTCAAGCGGTCTGCA    640
AtIRT rev  ACTATAGAAGACGCATGGGGTTGACTTCAAGCGGTCTGCA    345
Consensus  actatagaagacgcatggggttgacttcaagcggtctgca

AtIRT1     AACGGCATGAACAATGCATTCAAACGATAATATAAACCTT    596
AtIRT for  AACGGCATGAACAATGCATTCAAACGATAATATAAACCTT    680
AtIRT rev  AACGGCATGAACAATGCATTCAAACGATAATATAAACCTT    385
Consensus  aacggcatgaacaatgcattcaaacgataatataaaccttt

AtIRT1     AATTCCAATATAGCCTTTTTCAAATTTAAAAACTCGAATG    636
AtIRT for  AATTCCAATATAGCCTTTTTCAAATTTAAAAACTCGAATG    720
AtIRT rev  AATTCCAATATAGCCTTTTTCAAATTTAAAAACTCGAATG    425
Consensus  aattccaatatagccttttttcaaatttaaaaactcgaatg

AtIRT1     TTTACGAAACTATAACTTGATAAAACTGATTGAAATAAAA    676
AtIRT for  TTTACGAAACTATAACTTGATAAAACTGATTGAAATAAAA    760
AtIRT rev  TTTACGAAACTATAACTTGATAAAACTGATTGAAATAAAA    465
Consensus  tttacgaaactataacttgataaaactgattgaaataaaa

AtIRT1     TGTACTTACATTTTTTTAATAAAATATACTTAAATTTTTT    716
AtIRT for  TGTACTTACATTTTTTTAATAAAATATACTTAAATTTTTT    800
 tIRT rev  TGTACTTACATTTTTTTAATAAAATATACTTAAATTTTTT    505
Consensus  tgtacttacatttttttaataaaatatacttaaatttttt
```

```
AtIRT1      AAATAAAATATACTTTAAAATATACTTAATTTTTTAAAAT    756
AtIRT for   AAATAAAATATACTTTAAAATATACTTAATTTTTTAAAAT    840
AtIRT rev   AAATAAAATATACTTTAAAATATACTTAATTTTTTAAAAT    545
Consensus   aaataaaatatactttaaaatatacttaatttttaaaat

AtIRT1      CATACTATGGATTTTAAACAAAATAAATTGTCACTGAGTC    796
AtIRT for   CATACTATGGATTTTAAACAAAATAAATTGTCACTGAGTC    880
AtIRT rev   CATACTATGGATTTTAAACAAAATAAATTGTCACTGAGTC    585
Consensus   catactatggattttaaacaaaataaattgtcactgagtc

AtIRT1      ATATTTCTTTGGATAAAAGAACTTTACATAAGTGATTGGT    836
AtIRT for   ATATTTCTTTGGATAAAAGAACTTTACATAAGTGATTGGT    920
AtIRT rev   ATATTTCTTTGGATAAAAGAACTTTACATAAGTGATTGGT    625
Consensus   atatttctttggataaaagaactttacataagtgattggt

AtIRT1      ACAATATAATAATATGCAAATTTAAGACATTTTATTAAGT    876
AtIRT for   ACAATATAATAATATGCAAATTTAAGACATTTTATTAAGT    960
AtIRT rev   ACAATATAATAATATGCAAATTTAAGACATTTTATTAAGT    665
Consensus   acaatataataatatgcaaatttaagacattttattaagt

AtIRT1      TTATATTATCGTTTAAATGCATTTTCTATCAACTTAATAA    916
AtIRT for   TTATATTATCGTTTAAATGCATTTTCTATCAACTTAATAA    999
AtIRT rev   TTATATTATCGTTTAAATGCATTTTCTATCAACTTAATAA    705
Consensus   ttatattatcgtttaaatgcattttctatcaacttaataa

AtIRT1      AATCTAGCTAGGTACTTTCATAATTCTATCATTATTCCTT    956
AtIRT for   AATCTAGCTAGGTACTTTCATAATTCTATCATTATTCCTT   1039
AtIRT rev   AATCTAGCTAGGTACTTTCATAATTCTATCATTATTCCTT    745
Consensus   aatctagctaggtactttcataattctatcattattcctt

AtIRT1      TTGAACGATTAATTACTTTATTTTTCTCGTTTCACACTGG    995
AtIRT for   TTGAACGATTAATTACTTTATTTTtCTCGTTTC.......   1072
AtIRT rev   TTGAACGATTAATTACTTTATTTTTCTCGTTTCACACTGG    784
Consensus   ttgaacgattaattactttatttttctcgtttc

AtIRT1      AAGCAAACGAGTAATCAAATAAAATTTCGTCATCATCGAC   1035
AtIRT rev   AAGCAAACGAGTAATCAAATAAAATTTCGTCATCATCGAC    824
Consensus   aagcaaacgagtaatcaaataaaatttcgtcatcatcgac

AtIRT1      AATAAATGGCATGAAAACAACCAGCACACATTACTATATA   1075
AtIRT rev   AATAAATGGCATGAAAACAACCAGCACACATTACTATATA    864
Consensus   aataaatggcatgaaaacaaccagcacacattactatata

AtIRT1      AACCGATTCATCTAGTAAGTCCAAAGTCACACAAACATTA   1115
AtIRT rev   AACCGATTCATCTAGTAAGTCCAAAGTCACACAAACATTA    904
Consensus   aaccgattcatctagtaagtccaaagtcacacaaacatta

AtIRT1      AACAATCTACAACTTTGTATACAAAGTTGTCCC.......   1148
AtIRT rev   AACAATCTACAACTTTGTATACAAAAGTTGAACGAGAAAC    944
Consensus   aacaatctacaactttgtatacaaa gtt
```

ScIRT1 in pDONR221 P5P2

```
ScIRT     ......GGGACAACTTTGTATACAAAAGTTGTAAAAATGG     35
ScIRT for TTTATAATGCCAACTTTGTATACAAAAGTTGTAAAAATGG    120
Consensus         g caactttgtatacaaaagttgtaaaaatgg

ScIRT     CTTCAAGTTCAAAAAATGTCACAATCTTGGTCATCTGTAT     75
ScIRT for CTTCAAGTTCAAAAAATGTCACAATCTTGGTCATCTGTAT    160
```

```
Consensus cttcaagttcaaaaaatgtcacaatcttggtcatctgtat

ScIRT     AGCATTGCTATGCACCAAAGCCTTGTCACAATGTGAAGAT    115
ScIRT for AGCATTGCTATGCACCAAAGCCTTGTCACAATGTGAAGAT    200
ScIRT rev AGCATTGCTATGCACCAAAGCCTTGTCACAATGTGAAGAT     41
Consensus agcattgctatgcaccaaagccttgtcacaatgtgaagat

ScIRT     GAAACCAACAACCCCTGTAACAACAAATCGAAAGCTTTAT    155
ScIRT for GAAACCAACAACCCCTGTAACAACAAATCGAAAGCTTTAT    240
ScIRT rev GAACCACAC....CCTGTAC..ACAAATCGAAAGCTTTAT     75
Consensus gaa c  ac    cctgta   acaaatcgaaagctttat

ScIRT     CTTTAAAAATCATTGGTATTGCGGTGATCCTTTTAACAAG    195
ScIRT for CTTTAAAAATCATTGGTATTGCGGTGATCCTTTTAACAAG    280
ScIRT rev CTT..AAAATCATTGGTATTGCGGTGATCCTTTTAACAAG    111
Consensus ctt  aaaatcattggtattgcggtgatccttttaacaag

ScIRT     TCTGATCGGTGTATCTCTACCGCTTATCACTCGTTCGGTC    235
ScIRT for TCTGATCGGTGTATCTCTACCGCTTATCACTCGTTCGGTC    320
ScIRT rev TCTGATCGGTGTATCTCTACCGCTTATCACTCGTTCGGTC    151
Consensus tctgatcggtgtatctctaccgcttatcactcgttcggtc

ScIRT     CCTGCCCTTAGTCCAGACCGTAGCCTTTTTGTCATCGTTA    275
ScIRT for CCTGCCCTTAGTCCAGACCGTAGCCTTTTTGTCATCGTTA    360
ScIRT rev CCTGCCCTTAGTCCAGACCGTAGCCTTTTTGTCATCGTTA    189
Consensus cctgcccttagtccagaccgtagccttttgtcatcgtta

ScIRT     AAGCCTTTGCTTCAGGCATCATTTTGGCTACTGGATTCAT    315
ScIRT for AAGCCTTTGCTTCAGGCATCATTTTGGCTACTGGATTCAT    400
ScIRT rev AAGCCTTTGCTTCAGGCATCATTTTGGCTACTGGATTCAT    229
Consensus aagcctttgcttcaggcatcattttggctactggattcat

ScIRT     GCACGTGTTGCCAGACTCTTTTGACAAGTTGAGGTCTAGT    355
ScIRT for GCACGTGTTGCCAGACTCTTTTGACAAGTTGAGGTCTAGT    440
ScIRT rev GCACGTGTTGCCAGACTCTTTTGACAAGTTGAGGTCTAGT    269
Consensus gcacgtgttgccagactcttttgacaagttgaggtctagt

ScIRT     TGTTTGGCTGACAACCCGTGGCATAAGTTTCCATTTACGG    395
ScIRT for TGTTTGGCTGACAACCCGTGGCATAAGTTTCCATTTACGG    480
ScIRT rev TGTTTGGCTGACAACCCGTGGCATAAGTTTCCATTTACGG    309
Consensus tgtttggctgacaacccgtggcataagtttccatttacgg

ScIRT     GCTTTGTTGCTATGTTATCGGCTATTTTCACACTCATTGT    435
ScIRT for GCTTTGTTGCTATGTTATCGGCTATTTTCACACTCATTGT    520
ScIRT rev GCTTTGTTGCTATGTTATCGGCTATTTTCACACTCATTGT    349
Consensus gctttgttgctatgttatcggctattttcacactcattgt

ScIRT     TGACTCAATGGCTACAAGCATGTCTACCAAGAAGAATAAT    475
ScIRT for TGACTCAATGGCTACAAGCATGTCTACCAAGAAGAATAAT    560
ScIRT rev TGACTCAATGGCTACAAGCATGTCTACCAAGAAGAATAAT    389
Consensus tgactcaatggctacaagcatgtctaccaagaagaataat

ScIRT     GCGATATCAGCTGAAGGAGGTGAGGTGGTTGGCGGGGACC    515
ScIRT for GCGATATCAGCTGAAGGAGGTGAGGTGGTTGGCGGGGACC    600
ScIRT rev GCGATATCAGCTGAAGGAGGTGAGGTGGTTGGCGGGGACC    429
Consensus gcgatatcagctgaaggaggtgaggtggttggcggggacc

ScIRT     AGGAGATGGCGGGTGCGAGTAGTGGTGGTGCACATTTCCA    555
ScIRT for AGGAGATGGCGGGTGCGAGTAGTGGTGGTGCACATTTCCA    640
ScIRT rev AGGAGATGGCGGGTGCGAGTAGTGGTGGTGCACATTTCCA    469
Consensus aggagatggcgggtgcgagtagtggtggtgcacatttcca
```

```
ScIRT     TGGACACCATCATGTCCAAAAGGGAGCAGTAGGACCACAA    595
ScIRT for TGGACACCATCATGTCCAAAAGGGAGCAGTAGGACCACAA    680
ScIRT rev TGGACACCATCATGTCCAAAAGGGAGCAGTAGGACCACAA    509
Consensus tggacaccatcatgtccaaaagggagcagtaggaccacaa

ScIRT     CTTCTTCATTATCGGGTCGTAGCTATGGTGCTCGAACTTG    635
ScIRT for CTTCTTCATTATCGGGTCGTAGCTATGGTGCTCGAACTTG    720
ScIRT rev CTTCTTCATTATCGGGTCGTAGCTATGGTGCTCGAACTTG    549
Consensus cttcttcattatcgggtcgtagctatggtgctcgaacttg

ScIRT     GAATAATTGTTCACTCAATAGTTATCGGACTTGGAGTTGG    675
ScIRT for GAATAATTGTTCACTCAATAGTTATCGGACTTGGAGTTGG    760
ScIRT rev GAATAATTGTTCACTCAATAGTTATCGGACTTGGAGTTGG    589
Consensus gaataattgttcactcaatagttatcggacttggagttgg

ScIRT     GGCATCAAATGATGTATGCACTATTAAGCCATTGGTAGCA    715
ScIRT for GGCATCAAATGATGTATGCACTATTAAGCCATTGGTAGCA    800
ScIRT rev GGCATCAAATGATGTATGCACTATTAAGCCATTGGTAGCA    629
Consensus ggcatcaaatgatgtatgcactattaagccattggtagca

ScIRT     GTTCTTTGTTTCCATCAGATATTTGAAGGAATGGGTCTTG    755
ScIRT for GTTCTTTGTTTCCATCAGATATTTGAAGGAATGGGTCTTG    840
ScIRT rev GTTCTTTGTTTCCATCAGATATTTGAAGGAATGGGTCTTG    669
Consensus gttctttgtttccatcagatatttgaaggaatgggtcttg

ScIRT     GTGGTTGCATTCTTCAGGCCGAGTACAAAACCATGAAGAA    795
ScIRT for GTGGTTGCATTCTTCAGGCCGAGTACAAAACCATGAAGAA    879
ScIRT rev GTGGTTGCATTCTTCAGGCCGAGTACAAAACCATGAAGAA    709
Consensus gtggttgcattcttcaggccgagtacaaaaccatgaagaa

ScIRT     AGCAATAATGGTCCTTTTTTTCTCAATAACAATTCCATTT    835
ScIRT for AGCAATAATGGTCCTTTTTTTCTCAATAACAATTCCATTT    919
ScIRT rev AGCAATAATGGTCCTTTTTTTCTCAATAACAATTCCATTT    749
Consensus agcaataatggtccttttttttctcaataacaattccattt

ScIRT     GGGATTGCTCTTGGGATCGTGTTGTCAAAGACATACAAAG    875
ScIRT for GGGATTGCTCTTGGGATCGTGTTGTCAAAGACATACAAAG    959
ScIRT rev GGGATTGCTCTTGGGATCGTGTTGTCAAAGACATACAAAG    789
Consensus gggattgctcttgggatcgtgttgtcaaagacatacaaag

ScIRT     AAAATAGCCCGAGTGCGTTGATCACAGTAGGATTACTCAA    915
ScIRT for AAAATAGCCCGAGTGCGTTGATCACAGTAGGATTACTCAA    999
ScIRT rev AAAATAGCCCGAGTGCGTTGATCACAGTAGGATTACTCAA    829
Consensus aaaatagcccgagtgcgttgatcacagtaggattactcaa

ScIRT     TGCTTCATCAGCCGGACTTCTTATCTACATGGCGTTGGTT    955
ScIRT for TGCTTCATCAGCCGGACTTCTTATCTACATGGCGTTGGTT   1039
ScIRT rev TGCTTCATCAGCCGGACTTCTTATCTACATGGCGTTGGTT    869
Consensus tgcttcatcagccggacttcttatctacatggcgttggtt

ScIRT     GATCTACTTGCAGCTGATTTCATGGGTCCGAAACTTCAAG    995
ScIRT for GATCTACTTGCAGCTGATTTCATGGGTCCGAAACTT....   1074
ScIRT rev GATCTACTTGCAGCTGATTTCATGGGTCCGAAACTTCAAG    909
Consensus gatctacttgcagctgatttcatgggtccgaaactt

ScIRT     GAAGCATTAAGCTACAAATTAAATCTTATGTAGCAGTCTT   1035
ScIRT rev GAAGCATTAAGCTACAAATTAAATCTTATGTAGCAGTCTT    949
Consensus gaagcattaagctacaaattaaatcttatgtagcagtctt

ScIRT     GCTTGGCGCCGGTGGGATGTCTCTGATGGCAAAATGGGCT   1075
```

```
ScIRT rev GCTTGGCGCCGGTGGGATGTCTCTGATGGCAAAATGGGCT    989
Consensus gcttggcgccggtgggatgtctctgatggcaaaatgggct

ScIRT     TAAAACCCAGCTTTCTTGTACAAAGTGGTCCC........   1107
ScIRT rev TAAAACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAA   1029
Consensus taaaacccagctttcttgtacaaagt g c
```

AtIREG2 in pDONR221 P1P5r

Forward

```
AtIREG2      ..GGGACAAGTTTGTACAAAAAAGCA..GGCTTAGAG..A     35
AtIREG for   TAATGCCAACTTTGTACAAAAAAGCA..GGCTTAGAG..A    116
Consensus             a ttt t c a aa  c   g ctt ga

AtIREG2      GTAGTCGGAATGTTCTGGATAAATGACTTACGTAACCCGG     74
ATIREG for   GTAGTCGGAATGTTCTGGATAAATGACTTACGTAACCCGG    155
Consensus    gtagtcggaatgttctggataaataacttacgtaacccgg

AtIREG2      AAGAGTGGTTTAGAATACGCGCCGTTTACGCGTTTAACAC    114
ATIREG for   AAGAGTGGTTTAGAATACGCGCCGTTTACGCGTTTAACAC    195
Consensus    aagagtggtttagaatacgcgccgtttacgcgtttaacac

AtIREG2      ATCATGCCGTTTTCTTCATAAGAATATACAAAACGTGCCG    154
ATIREG for   ATCATGCCGTTTTCTTCATAAGAATATACAAAACGTGCCG   235
Consensus    atcatgccgttttcttcataagaatatacaaaacgtgccg

AtIREG2      TTTAATTTTCGGGTCTGCTTCGGCCCATATAACTAGCCC    191
ATIREG for   TTTAATTTTCGGGTCTGCTTCGGCCCATATAACTAGCCC    272
Consensus    tttaattttcgggtctgcttcggcccatataactagccc

AtIREG2      ATTTATAGATTGTTGGATACGTGGATTAATTAAAAAGAAT    231
ATIREG for   ATTTATAGATTGTTGGATACGTGGATTAATTAAAAAGAAT    312
Consensus    atttatagattgttggatacgtggattaattaaaaagaat

AtIREG2      ATTTTGTTGATGATCCAATAGATAGAGAAACAAACCCATT    271
ATIREG for   ATTTTGTTGATGATCCAATAGATAGAGAAACAAACCCATT  352
Consensus    attttgttgatgatccaatagatagagaaacaaacccatt

AtIREG2      AGACATCGTGGGATTTGATATTTTGTTGGACGTACATATT    308
ATIREG for   AGACATCGTGGGATTTGATATTTTGTTGGACGTACATATT    389
Consensus    agacatcgtgggatttgatattttgttggacgtacatatt

AtIREG2      TTTTTTTAACAAAATCACATCCCTAAACGTTATTCATTG    346
ATIREG for   TTTTTTTAACAAAATCACATCCCTAAACGTTATTCATTG    427
Consensus    ttttttaacaaaatcacatccctaaacgttattcattg

AtIREG2      TTTCTTACTTTCTTTTGTTGATGAAACATCGTTTGCCACC    386
ATIREG for   TTTCTTACTTTCTTTTGTTGATGAAACATCGTTTGCCACC    467
Consensus    tttcttactttctttttgttgatgaaacatcgtttgccacc

AtIREG2      ACCTATTTTGCTGTGAAAAGTTGCTAATTTTATGGAAAAG    426
ATIREG for   ACCTATTTTGCTGTGAAAAGTTGCTAATTTTATGGAAAAG    507
Consensus    acctattttgctgtgaaaagttgctaattttatggaaaag

AtIREG2      GTGTTTCAGACCAAAAACAATAAATCTACATAAATAATAT    466
ATIREG for   GTGTTTCAGACCAAAAACAATAAATCTACATAAATAATAT    547
Consensus    gtgtttcagaccaaaaacaataaatctacataaataatat

AtIREG2      TGAACTCAAAAGTCATGTATTAATTTCTTGTAATGTACAT    506
ATIREG for   TGAACTCAAAAGTCATGTATTAATTTCTTGTAATGTACAT    587
```

```
Consensus    tgaactcaaaagtcatgtattaatttcttgtaatgtacat

AtIREG2      ATTTTTTAATTGTGACAGCATAACGTGTCAATCAGTTTAG    546
ATIREG for   ATTTTTTAATTGTGACAGCATAACGTGTCAATCAGTTTAG    627
Consensus    atttttaattgtgacagcataacgtgtcaatcagtttag

AtIREG2      AATTAGGTACCAACACTTTGACTGTATTAAGACATGATGC    586
ATIREG for   AATTAGGTACCAACACTTTGACTGTATTAAGACATGATGC    667
Consensus    aattaggtaccaacactttgactgtattaagacatgatgc

AtIREG2      ATGCTACATAAAAGTTCATGTTTTAATATACAACAATGTT    626
ATIREG for   ATGCTACATAAAAGTTCATGTTTTAATATACAACAATGTT    707
Consensus    atgctacataaaagttcatgttttaatatacaacaatgtt

AtIREG2      TCTTATCCGTATGTGTATTTTATATATTTCATGATTTGAG    666
ATIREG for   TCTTATCCGTATGTGTATTTTATATATTTCATGATTTGAG    747
Consensus    tcttatccgtatgtgtattttatatatttcatgatttgag

AtIREG2      ATATAATGAGAAGATAATGTTGAAGCTTTTAATAGGACGA    706
AtIREG for   ATATAATGAGAAGATAATGTTGAAGCTTTTAATAGGACGA    787
Consensus    atataatgagaagataatgttgaagcttttaataggacga

AtIREG2      ATTGTTGTATTGTATGTAACAAGTTGTGTATATAAACTAT    745
AtIREG for   ATTGTTGTATTGTATGTAACAAGTTGTGTATATAAACTAT    826
Consensus    attgttgtattgtatgtaacaagttgtgtatataaactat

AtIREG2      GGCTGATGGCTCTCACAAGAAAACGTTGTCTTATTTACTC    785
AtIREG for   GGCTGATGGCTCTCACAAGAAAACGTTGTCTTATTTACTC    866
Consensus    ggctgatggctctcacaagaaaacgttgtcttatttactc

AtIREG2      AGTCGTTGTGATGTCCGAGGCGCTTTTTTTTTGTTTGTTT    825
AtIREG for   AGTCGTTGTGATGTCCGAGGCGCTTTTTTTTTGTTTGTTT    906
Consensus    agtcgttgtgatgtccgaggcgctttttttttgtttgttt

AtIREG2      ATTAGTATGATCACTGAGATATTCTAACGGACGGGACATA    865
AtIREG for   ATTAGTATGATCACTGAGATATTCTAACGGACGGGACATA    946
Consensus    attagtatgatcactgagatattctaacggacgggacata

AtIREG2      AACGACGATTAAAAAATTATGGTACCTGATCATCGTTGAC    905
AtIREG for   AACGACGATTAAAAAATTATGGTACCTGATCATCGTTGAC    986
Consensus    aacgacgattaaaaaattatggtacctgatcatcgttgac

AtIREG2      GCCAAAAATGAGGAAATATTCGGACATTTTT             936
AtIREG for   GCCAAAAATGAGGAAATATTCGGACATTTTT            1015
Consensus    gccaaaaatgaggaaatattcggacattttt
```

Middle region

```
AtIREG2      GTATTGTATGTAACAAGTTGTGTATATAAACTATGGCTGA    760
AtIREG mid   ...............AGTTGTGTATATAAACTATGGCTGA     25
Consensus                   agttgtgtatataaactatggctga

AtIREG2      TGGCTCTCACAAGAAAACGTTGTCTTATTTACTCAGTCGT    800
AtIREG mid   TGGCTCTCACAAGAAAACGTTGTCTTATTTACTCAGTCGT     65
Consensus    tggctctcacaagaaaacgttgtcttatttactcagtcgt

AtIREG2      TGTGATGTCCGAGGCGCTTTTTTTTTGTTTGTTTATTAGT    840
AtIREG mid   TGTGATGTCCGAGGCGCTTTTTTTTTGTTTGTTTATTAGT    105
Consensus    tgtgatgtccgaggcgctttttttttgtttgtttattagt
```

```
AtIREG2       ATGATCACTGAGATATTCTAACGGACGGGACATAAACGAC    880
AtIREG mid    ATGATCACTGAGATATTCTAACGGACGGGACATAAACGAC    145
Consensus     atgatcactgagatattctaacggacgggacataaacgac

AtIREG2       GATTAAAAAATTATGGTACCTGATCATCGTTGACGCCAAA    920
AtIREG mid    GATTAAAAAATTATGGTACCTGATCATCGTTGACGCCAAA    185
Consensus     gattaaaaaattatggtacctgatcatcgttgacgccaaa

AtIREG2       AATGAGGAAATATTCGGACATTTTTATCCCTAATCCCCAA    960
AtIREG mid    AATGAGGAAATATTCGGACATTTTTATCCCTAATCCCCAA    225
Consensus     aatgaggaaatattcggacatttttatccctaatccccaa

AtIREG2       AACAATTCCCATATAGTGCAACTAGAAATGTGTCTCAGTT   1000
AtIREG mid    AACAATTCCCATATAGTGCAACTAGAAATGTGTCTCAGTT    265
Consensus     aacaattcccatatagtgcaactagaaatgtgtctcagtt

AtIREG2       TCTGCGATAATTTCCTCGACTTCGATTTGGTGCTTCGACC   1040
AtIREG mid    TCTGCGATAATTTCCTCGACTTCGATTTGGTGCTTCGACC    305
Consensus     tctgcgataatttcctcgacttcgatttggtgcttcgacc

AtIREG2       ATTCATCCTTCAAACCCGTCCCTGTGCCAGGTGAGTTTTG   1080
AtIREG mid    ATTCATCCTTCAAACCCGTCCCTGTGCCAGGTGAGTTTTG    345
Consensus     attcatccttcaaacccgtccctgtgccaggtgagttttg

AtIREG2       TAGATATTCATGACGTTTTTTTTTTTTTTTTTTTTTGGT   1120
AtIREG mid    TAGATATTCATGACGTTTTTTTTTTTTTTTTTTTTTGGT    385
Consensus     tagatattcatgacgtttttttttttttttttttttggt

AtIREG2       CATCATATTCATGAAGTTTAGTTTTACGAAAAATTGAAAA   1160
AtIREG mid    CATCATATTCATGAAGTTTAGTTTTACGAAAAATTGAAAA    425
Consensus     catcatattcatgaagtttagttttacgaaaaattgaaaa

AtIREG2       TCGAACTCAAATTCTAGATTTCTTTGCGTCAAGAAATTTA   1200
AtIREG mid    TCGAACTCAAATTCTAGATTTCTTTGCGTCAAGAAATTTA    465
Consensus     tcgaactcaaattctagatttctttgcgtcaagaaattta

AtIREG2       GAAGTAATCAGAGACTTAATCAATTTTTGAACTAGACGTT   1240
AtIREG mid    GAAGTAATCAGAGACTTAATCAATTTTTGAACTAGACGTT    505
Consensus     gaagtaatcagagacttaatcaatttttgaactagacgtt

AtIREG2       CAAAACATGGAGAATATTTCACAAACGAAAATACCTTTTT   1280
AtIREG mid    CAAAACATGGAGAATATTTCACAAACGAAAATACCTTTTT    545
Consensus     caaaacatggagaatatttcacaaacgaaaatacctttttt

AtIREG2       TGGACGGATCAGAAATTGTCATATCTCAGATTTGTAATTT   1320
AtIREG mid    TGGACGGATCAGAAATTGTCATATCTCAGATTTGTAATTT    585
Consensus     tggacggatcagaaattgtcatatctcagatttgtaattt

AtIREG2       GTTTCAAGAAGATTAATTATCTTTTAAGGTTTTTCAAATT   1360
AtIREG mid    GTTTCAAGAAGATTAATTATCTTTTAAGGTTTTTCAAATT    625
Consensus     gtttcaagaagattaattatcttttaaggtttttcaaatt

AtIREG2       TTTAGTTTTTTCAAAAAATTTAAACATTAATTTAAGTTAC   1400
AtIREG mid    TTTAGTTTTTTCAAAAAATTTAAACATTAATTTAAGTTAC    665
Consensus     tttagtttttttcaaaaaatttaaacattaatttaagttac

AtIREG2       TAATAATTAGGGGAAAGAAAATATTATAATTAGGAAATAT   1440
AtIREG mid    TAATAATTAGGGGAAAGAAAATATTATAATTAGGAAATAT    705
Consensus     taataattaggggaaagaaaatattataattaggaaatat

AtIREG2       AATATTATAATTAGGAAATATATGTCAAAAAAAAAAAAAT   1480
```

```
AtIREG mid   AATATTATAATTAGGAAATATATGTCAAAAAAAAAAAAAT    745
Consensus    aatattataattaggaaatatatgtcaaaaaaaaaaaat

AtIREG2      CTAAAAATCAATTTTTATTTAGATAATAAGATAATAAATA   1520
AtIREG mid   CTAAAAATCAATTTTTATTTAGATAATAAGATAATAAATA    785
Consensus    ctaaaaatcaatttttatttagataataagataataaata

AtIREG2      AATTCTTGTAATATAAACGCAAAACCTCAGAAAATAAAAT   1560
AtIREG mid   AATTCTTG.............................      793
Consensus    aattcttg
```

Reverse

```
AtIREG2      CATCATATTCATGAAGTTTAGTTTTACGAAAAATTGAAAA   1160
AtIREG rev   CATCATATTCATGAAGTTTAGTTTTACGAAAAATTGAAAA    163
Consensus    catcatattcatgaagtttagttttacgaaaaattgaaaa

AtIREG2      TCGAACTCAAATTCTAGATTTCTTTGCGTCAAGAAATTTA   1200
AtIREG rev   TCGAACTCAAATTCTAGATTTCTTTGCGTCAAGAAATTTA    203
Consensus    tcgaactcaaattctagatttctttgcgtcaagaaattta

AtIREG2      GAAGTAATCAGAGACTTAATCAATTTTTGAACTAGACGTT   1240
AtIREG rev   GAAGTAATCAGAGACTTAATCAATTTTTGAACTAGACGTT    243
Consensus    gaagtaatcagagacttaatcaatttttgaactagacgtt

AtIREG2      CAAAACATGGAGAATATTTCACAAACGAAAATACCTTTTT   1280
AtIREG rev   CAAAACATGGAGAATATTTCACAAACGAAAATACCTTTTT    283
Consensus     caaaacatggagaatatttcacaaacgaaaatacctttttt

AtIREG2      TGGACGGATCAGAAATTGTCATATCTCAGATTTGTAATTT   1320
AtIREG rev   TGGACGGATCAGAAATTGTCATATCTCAGATTTGTAATTT    323
Consensus     tggacggatcagaaattgtcatatctcagatttgtaattt

AtIREG2      GTTTCAAGAAGATTAATTATCTTTTAAGGTTTTTCAAATT   1360
AtIREG rev   GTTTCAAGAAGATTAATTATCTTTTAAGGTTTTTCAAATT    363
Consensus    gtttcaagaagattaattatcttttaaggtttttcaaatt

AtIREG2      TTTAGTTTTTTCAAAAAATTTAAACATTAATTTAAGTTAC   1400
AtIREG rev   TTTAGTTTTTTCAAAAAATTTAAACATTAATTTAAGTTAC    403
Consensus    tttagtttttttcaaaaaatttaaacattaatttaagttac

AtIREG2      TAATAATTAGGGGAAAGAAAATATTATAATTAGGAAATAT   1440
AtIREG rev   TAATAATTAGGGGAAAGAAAATATTATAATTAGGAAATAT    443
Consensus    taataattaggggaaagaaaatattataattaggaaatat

AtIREG2      AATATTATAATTAGGAAATATATGTCAAAAAAAAAAAAAT   1480
AtIREG rev   AATATTATAATTAGGAAATATATGTCAAAAAAAAAAAAAT    483
Consensus    aatattataattaggaaatatatgtcaaaaaaaaaaaaat

AtIREG2      CTAAAAATCAATTTTTATTTAGATAATAAGATAATAAATA   1520
AtIREG rev   CTAAAAATCAATTTTTATTTAGATAATAAGATAATAAATA    523
Consensus    ctaaaaatcaatttttatttagataataagataataaata

AtIREG2      AATTCTTGTAATATAAACGCAAAACCTCAGAAAATAAAAT   1560
AtIREG rev   AATTCTTGTAATATAAACGCAAAACCTCAGAAAATAAAAT    563
Consensus    aattcttgtaatataaacgcaaaacctcagaaaataaaat

AtIREG2      TCTGCAAAATTTTATCATGTAATGCATACTCTCTAATTTT   1600
AtIREG rev   TCTGCAAAATTTTATCATGTAATGCATACTCTCTAATTTT    603
Consensus    tctgcaaaattttatcatgtaatgcatactctctaatttt
```

```
AtIREG2       CTTAATCCTTTTTTTGATACAGTTTATATGTTTTCTTCAT   1640
AtIREG rev    CTTAATCCTTTTTTTGATACAGTTTATATGTTTTCTTCAT    643
Consensus     cttaatccttttttgatacagtttatatgttttcttcat

AtIREG2       TATTCATCTCATCCATGTGATCATGTCTTACTAGCTGCTT   1680
AtIREG rev    TATTCATCTCATCCATGTGATCATGTCTTACTAGCTGCTT    683
Consensus     tattcatctcatccatgtgatcatgtcttactagctgctt

AtIREG2       GCACCCATTGAATTGCAAAATCTAGAGAAGCCAGGGATAT   1720
AtIREG rev    GCACCCATTGAATTGCAAAATCTAGAGAAGCCAGGGATAT    723
Consensus     gcacccattgaattgcaaaatctagagaagccagggatat

AtIREG2       GCAATGAACACTGAATTAGCTTTGATCGATGTTTCCGATT   1760
AtIREG rev    GCAATGAACACTGAATTAGCTTTGATCGATGTTTCCGATT    763
Consensus     gcaatgaacactgaattagctttgatcgatgtttccgatt

AtIREG2       TTCAGATTCAATTGATATTCAGAGGAAAGAATCAAAGTAG   1800
AtIREG rev    TTCAGATTCAATTGATATTCAGAGGAAAGAATCAAAGTAG    803
Consensus     ttcagattcaattgatattcagaggaaagaatcaaagtag

AtIREG2       TCAGAAGACAACTTTGTATACAAAAGTTGTCCC.......   1833
AtIREG rev    TCAGAAGACAACTTTGTATACAAAAGTTGaaCgagaaacg    843
Consensus     tcagaagacaactttgtatacaaaagttg  c
```

ScIREG2 in Pdonr221 P5P2

Forward

```
BP IREG FEB 2019          ACTTTGTATACAAAAGTTGTAAAAATGGAGGAGGGGTATT     66
BP IREG SEP 2018          ACTTTGTATACAAAAGTTGTAAAAATGGAGGAGGGGTATT     66
IREG LIGATION FEB 2019    AtTcaGTcgACtgAcGgatccAAAATGGAGGAGGGGTATT     80
IREG REF GENE             ....................................         0
Consensus

BP IREG FEB 2019          ATTATGCAATAACAGAAGATGGACACCATCATCATCATCA    106
BP IREG SEP 2018          ATTATGCAATAACAGAAGATGGACACCATCATCATCATCA    106
IREG LIGATION FEB 2019    ATTATGCAATAACAGAAGATGGACACCATCATCATCATCA    120
IREG REF GENE             ...ATGCAATAACAGAAGATGGACACCATCATCATCATCA     37
Consensus                    atgcaataacagaagatggacaccatcatcatcatca

BP IREG FEB 2019          AGATCCAATCCCTACATCTCTTCTAATTTCACTCTATGGT    146
BP IREG SEP 2018          AGATCCAATCCCTACATCTCTTCTAATTTCACTCTATGGT    146
IREG LIGATION FEB 2019    AGATCCAATCCCTACATCTCTTCTAATTTCACTCTATGGT    160
IREG REF GENE             AGATCCAATCCCTACATCTCTTCTAATTTCACTCTATGGT     77
Consensus                 agatccaatccctacatctcttctaatttcactctatggt

BP IREG FEB 2019          GGACATTTCTTAGCCAGATGGGGTGCCAGGATGTGGGAGT    186
BP IREG SEP 2018          GGACATTTCTTAGCCAGATGGGGTGCCAGGATGTGGGAGT    186
IREG LIGATION FEB 2019    GGACATTTCTTAGCCAGATGGGGTGCCAGGATGTGGGAGT    200
IREG REF GENE             GGACATTTCTTAGCCAGATGGGGTGCCAGGATGTGGGAGT    117
Consensus                 ggacatttcttagccagatggggtgccaggatgtgggagt

BP IREG FEB 2019          TTTCGGTCGGGTTGTATATGATAAACGTTTGGCCTAATTC    226
BP IREG SEP 2018          TTTCGGTCGGGTTGTATATGATAAACGTTTGGCCTAATTC    226
IREG LIGATION FEB 2019    TTTCGGTCGGGTTGTATATGATAAACGTTTGGCCTAATTC    240
IREG REF GENE             TTTCGGTCGGGTTGTATATGATAAACGTTTGGCCTAATTC    157
Consensus                 tttcggtcgggttgtatatgataaacgtttggcctaattc

BP IREG FEB 2019          GTTGCTTTTGGCGGCAACCTATGGTGTGGTCGAGTCAGCT    266
BP IREG SEP 2018          GTTGCTTTTGGCGGCAACCTATGGTGTGGTCGAGTCAGCT    266
IREG LIGATION FEB 2019    GTTGCTTTTGGCGGCAACCTATGGTGTGGTCGAGTCAGCT    280
```

```
IREG REF GENE            GTTGCTTTTGGCGGCAACCTATGGTGTGGTCGAGTCAGCT    197
Consensus                gttgcttttggcggcaacctatggtgtggtcgagtcagct

BP IREG FEB 2019         TCGACCACCTTATTTGGTCCTCTCGTAGGCCAATGGATCG    306
BP IREG SEP 2018         TCGACCACCTTATTTGGTCCTCTCGTAGGCCAATGGATCG    306
IREG LIGATION FEB 2019   TCGACCACCTTATTTGGTCCTCTCGTAGGCCAATGGATCG    320
IREG REF GENE            TCGACCACCTTATTTGGTCCTCTCGTAGGCCAATGGATCG    237
Consensus                tcgaccaccttatttggtcctctcgtaggccaatggatcg

BP IREG FEB 2019         ATAAGTCCACATATCCAAAGGTTTTAAAAGTTTGGCTGTT    346
BP IREG SEP 2018         ATAAGTCCACATATCCAAAGGTTTTAAAAGTTTGGCTGTT    346
IREG LIGATION FEB 2019   ATAAGTCCACATATCCAAAGGTTTTAAAAGTTTGGCTGTT    360
IREG REF GENE            ATAAGTCCACATATCCAAAGGTTTTAAAAGTTTGGCTGTT    277
Consensus                ataagtccacatatccaaaggttttaaaagtttggctgtt

BP IREG FEB 2019         GACACAAAACCTCTCGTTTATCGTTGCTGGAGTCACAGTG    386
BP IREG SEP 2018         GACACAAAACCTCTCGTTTATCGTTGCTGGAGTCACAGTG    386
IREG LIGATION FEB 2019   GACACAAAACCTCTCGTTTATCGTTGCTGGAGTCACAGTG    400
IREG REF GENE            GACACAAAACCTCTCGTTTATCGTTGCTGGAGTCACAGTG    317
Consensus                gacacaaaacctctcgtttatcgttgctggagtcacagtg

BP IREG FEB 2019         GTCGGACTTCTAATATCACCTGACTTGAGGATTAACAACC    426
BP IREG SEP 2018         GTCGGACTTCTAATATCACCTGACTTGAGGATTAACAACC    426
IREG LIGATION FEB 2019   GTCGGACTTCTAATATCACCTGACTTGAGGATTAACAACC    440
IREG REF GENE            GTCGGACTTCTAATATCACCTGACTTGAGGATTAACAACC    357
Consensus                gtcggacttctaatatcacctgacttgaggattaacaacc

BP IREG FEB 2019         GCGTTTTATTTGCTGTGCTCGTTGTGATGATAAATCTCTC    466
BP IREG SEP 2018         GCGTTTTATTTGCTGTGCTCGTTGTGATGATAAATCTCTC    466
IREG LIGATION FEB 2019   GCGTTTTATTTGCTGTGCTCGTTGTGATGATAAATCTCTC    480
IREG REF GENE            GCGTTTTATTTGCTGTGCTCGTTGTGATGATAAATCTCTC    397
Consensus                gcgttttatttgctgtgctcgttgtgatgataaatctctc

BP IREG FEB 2019         CGGAGCTCTTGCTGTTCTTTCGTCTCTTGCGGGCACCATA    506
BP IREG SEP 2018         CGGAGCTCTTGCTGTTCTTTCGTCTCTTGCGGGCACCATA    506
IREG LIGATION FEB 2019   CGGAGCTCTTGCTGTTCTTTCGTCTCTTGCGGGCACCATA    520
IREG REF GENE            CGGAGCTCTTGCTGTTCTTTCGTCTCTTGCGGGCACCATA    437
Consensus                cggagctcttgctgttctttcgtctcttgcgggcaccata

BP IREG FEB 2019         TTGATTGAAAGAGAATGGGTGGTGGTAATATCAGAAGGCC    546
BP IREG SEP 2018         TTGATTGAAAGAGAATGGGTGGTGGTAATATCAGAAGGCC    546
IREG LIGATION FEB 2019   TTGATTGAAAGAGAATGGGTGGTGGTAATATCAGAAGGCC    560
IREG REF GENE            TTGATTGAAAGAGAATGGGTGGTGGTAATATCAGAAGGCC    477
Consensus                ttgattgaaagagaatgggtggtggtaatatcagaaggcc

BP IREG FEB 2019         GCTCTTCAAACATACTAACAACGTTAAACTCCAGAATCCG    586
BP IREG SEP 2018         GCTCTTCAAACATACTAACAACGTTAAACTCCAGAATCCG    586
IREG LIGATION FEB 2019   GCTCTTCAAACATACTAACAACGTTAAACTCCAGAATCCG    600
IREG REF GENE            GCTCTTCAAACATACTAACAACGTTAAACTCCAGAATCCG    517
Consensus                gctcttcaaacatactaacaacgttaaactccagaatccg

BP IREG FEB 2019         TAGAATCGATTTAGTAAGCAAACTATTTGCGCCAGTTGCC    626
BP IREG SEP 2018         TAGAATCGATTTAGTAAGCAAACTATTTGCGCCAGTTGCC    626
IREG LIGATION FEB 2019   TAGAATCGATTTAGTAAGCAAACTATTTGCGCCAGcTGCC    640
IREG REF GENE            TAGAATCGATTTAGTAAGCAAACTATTTGCGCCAGTTGCC    557
Consensus                tagaatcgatttagtaagcaaactatttgcgccag tgcc

BP IREG FEB 2019         ACAGGATTCATCATAAGTTTCGTGTCAGAGATTGCATCTG    666
BP IREG SEP 2018         ACAGGATTCATCATAAGTTTCGTGTCAGAGATTGCATCTG    666
IREG LIGATION FEB 2019   ACAGGATTCATCATAAGTTTCGTGTCAGAGATTGCATCTG    680
IREG REF GENE            ACAGGATTCATCATAAGTTTCGTGTCAGAGATTGCATCTG    597
```

```
Consensus                acaggattcatcataagtttcgtgtcagagattgcatctg

BP IREG FEB 2019         CTGCCGGTTTAGCAATCTTTAATACATTATCAGTTTTCTT    706
BP IREG SEP 2018         CTGCCGGTTTAGCAATCTTTAATACATTATCAGTTTTCTT    706
IREG LIGATION FEB 2019   CTGCCaGcTTAGCAATCTTTAATAtATTATCAGTTTTCTT    720
IREG REF GENE            CTGCCaGTTTAGCAATCTTTAATAtATTATCAGTTTTCTT    637
Consensus                ctgcc g ttagcaatctttaata attatcagttttctt

BP IREG FEB 2019         GCAATATTGGCTCTTAAACTCTGTATACAAAGGAATCCCA    746
BP IREG SEP 2018         GCAATATTGGCTCTTAAACTCTGTATACAAAGGAATCCCA    746
IREG LIGATION FEB 2019   GCAATATTGGCTCcTAAACTCTGTATACAAAGGAATCCCc    760
IREG REF GENE            GCAATATTGGCTCcTAAACTCTGTATACAAAGGAATCCCA    677
Consensus                gcaatattggctc taaactctgtatacaaaggaatccc

BP IREG FEB 2019         TCTTTATCAGAAAGAAGCCGACAGCGTTCCCTAAGATTCG    786
BP IREG SEP 2018         TCTTTATCAGAAAGAAGCCGACAGCGTTCCCTAAGATTCG    786
IREG REF GENE            TCTTTATCAGAAAGAAGCCGACAGCGTTCCCTAAGATTCG    717
Consensus                tctttatcagaaagaagccgacagcgttccctaagattcg

BP IREG FEB 2019         CCACAGCAAATGACCAACAACAAACACAAACCTCTTCCAC    826
BP IREG SEP 2018         CCACAGCAAATGACCAACAACAAACACAAACCTCTTCCAC    826
IREG REF GENE            CCACAGCAAATGACCAACAACAAACACAAACCTCTTCCAC    757
Consensus                ccacagcaaatgaccaacaacaaacacaaacctcttccac

BP IREG FEB 2019         TCCTCAAGAACAAAATGATAATTCTA..............    852
BP IREG SEP 2018         TCCTCAAGAACAAAATGATAATTCTAaagacgtttattcg    866
IREG REF GENE            TCCTCAAGAACAAAATGATAATTCTgaagacgtttattcg    797
Consensus                tcctcaagaacaaaatgataattct
```

Reverse

```
SCIREG2       AATGGGTGGTGGTAATATCAGAAGGCCGCTCTTCAAACAT    540
IREG_REV      AATGGGTGGTGGTA.T..CAGAAGGGCGCTCTCCAACCAT     96
Consensus     aatgggtggtggta t  cagaagggcgctctccaa cat

SCIREG2       ACTAACAACGTTAAACTCCAGAATCCGTAGAATCGATTTA    580
IREG_REV      ACTACCA.CGTTAAACTCCAGAATCCGTAGAATCGATT.A    134
Consensus     actacca cgttaaactccaggatccgtagaatcgatt a

SCIREG2       GTAAGCAAACTATTTGCGCCAGTTGCCACAGGATTCATCA    620
IREG_REV      GTAAGCAAACTATTTGCGCCAGTTGCCACAGGAT.CATCA    173
Consensus     gtaagcaaactatttgcgccagttgccacaggat catca

SCIREG2       TAAGTTTCGTGTCAGAGATTGCATCTGCTGCCAGTTTAGC    660
IREG_REV      TAAGTTTCGTGTCAGAGATTGCATCTGCTGCCGGTTTAGC    213
Consensus     taagtttcgtgtcagagattgcatctgctgccgggttagc

SCIREG2       AATCTTTAATATTATTATCAGTTTTCTTGCAATATTGGCTC    700
IREG_REV      AATCTTTAATACATTATCAGTTTTCTTGCAATATTGGCTC    253
Consensus     aatctttaata attatcagttttcttgcaatattggctc

SCIREG2       CTAAACTCTGTATACAAAGGAATCCCATCTTTATCAGAAA    740
IREG_REV      TTAAACTCTGTATACAAAGGAATCCCATCTTTATCAGAAA    293
Consensus     ttaaactctgtatacaaaggaatcccatctttatcagaaa


ScIREG2       GAAGCCGACAGCGTTCCCTAAGATTCGCCACAGCAAATGA    778
IREG_REV      GAAGCCGACAGCGTTCCCTAAGATTCGCCACAGCAAATGA    331
Consensus     gaagccgacagcgttccctaagattcgccacagcaaatga


ScIREG2       CCAACAACAAACACAAACCTCTTCCACTCCTCAAGAACAA    814
```

```
IREG_REV      CCAACAACAAACACAAACCTCTTCCACTCCTCAAGAACAA    367
Consensus     ccaacaacaaacacaaacctcttccactcctcaagaacaa

ScIREG2       AATGATAATTCTGAAGACGTTTATTCGGGTAATTTGGTGG    854
IREG_REV      AATGATAATTCTAAAGACGTTTATTCGGGTAATTTGGTGG    407
Consensus     aatgataattctaaagacgtttattcgggtaatttggtgg

ScIREG2       TTAGAAAATTAGTCAAAAGGGTATCTGACAGTTCGTTTGT    894
IREG_REV      TTAGAAAATTAGTCAAAAGGGTATCTGACAGTTCGTTTGT    447
Consensus     ttagaaaattagtcaaaagggtatctgacagttcgtttgt

ScIREG2       TCGTGCTTGGCGAGTTTATCTGCAACAAGATGTGGTTTTA    934
IREG_REV      TCGTGCTTGGGGAGTTTATCTGCAACAAGATGTGGTTTTA    487
Consensus     tcgtgcttggggagtttatctgcaacaagatgtggtttta

ScIREG2       CCGGGATTATCTTTAGCGTTACTTTATTTTACGGTCCTCG    974
IREG_REV      CCGGGATTATCTTTAGCGTTACTTTATTTTACGGTCCTCA    527
Consensus     ccgggattatctttagcgttactttattttacggtcctc

ScIREG2       GTTTTGGGACTTTGATGACTGCAACATTGGAATGGGAAGG   1014
IREG_REV      GTTTTGGGACTTTGATGACTGCAACATTGGAATGGGAAGG    567
Consensus

ScIREG2       AATACCAGCATATATAATTGGAACAGGAAGAGGAATAAGT   1054
IREG_REV      AATACCAGCATATATAATTGGAACAGGAAGAGGAATAAGT    607
CONSENSUS     aaraccagcatatataattggaacaggaagaggaataagt

SCIREG2       GCGATTATTGGTATCTCTGCAACATTTTTGTATTCCTTTA   1094
IREG_REV      GCGATTATTGGTATCTCTGCAACATTTTTGTATCCTTTTA    647
CONSENSUS     gcgattattggtatctctgcaacattttgtatcctttta

SCIREG2       TGGAGACTCGAATGTCAACACTTCGAACCGGGCTTTGGTC   1134
IREG_REV      TGGAGACTCGAATGTCAACACTTCGAACCGGGCTTTGGTC    687
CONSENSUS     tggagactcgaatgtcaacacttcgaaccgggctttggtc

SCIREG2       TATTTGGTCACAGTGGAGTTGCCTCTTGGTTTGCGTTGGC   1174
IREG_REV      TATTTGGTCACAGTGGAGTTGCCTCTTGGTTTGCGTTGGC    727
CONSENSUS     tatttggtcacagtggagttgcctcttggtttgcgttggc

SCIREG2       TCGATATGGGTTAAAAACAACTCCACATCAGCATACCTAT   1214
IREG_REV      TCGATATGGGTTAAAAACAACTCCACATCAGCATACCTAT    767
CONSENSUS     tcgatatgggttaaaaacaactccacatcagcatacctat

SCIREG2       TAATGGCGGGAGTCGCAGCTTCACGTTTAGGATTATGGGT   1254
IREG_REV      TAATGGTGGGAGTCGCAGCTTCACGTTTAGGATTATGGGT    807
CONSENSUS     taatggtgggagtcgcagcttcacgtttaggattatgggt

SCIREG2       GTTTGATCTATCTGTCATCCAACAAATGCAGGATCAAGTG   1294
IREG_REV      GTTTGATCTATCTGTCATCCAACAAATGCAGGATCAAGTG    847
CONSENSUS     gtttgatctatctgtcatccaacaaatgcaggatcaagtg

SCIREG2       TCTGAATCCAATCGAGCTGTCGTTGGAGGGGTTCAAAACT   1334
IREG_REV      TCTGAATCCAATCGAGCTGTCGTTGGAGGGGTTCAAAACT    887
CONSENSUS     tctgaatccaatcgagctgtcgttggaggggttcaaaact


SCIREG2       CAGTTCAGTCGTTTTGGGATTTGATGACGTATATTATGGG   1374
IREG_REV      CAGTTCAGTCGTTTTGGGATTTGATGACGTATATTATGGG    927
CONSENSUS     cagttcagtcgttttgggatttgatgacgtatattatggg

SCIREG2       GTTAATCATTTCCAATCCTAAGGATTTTTGGATATTGATC   1414
```

```
IREG_REV      GTTAATCATTTCCAATCCTAAGGATTTTTGGATATTGATC    967
CONSENSUS     gttaatcatttccaatcctaaggatttttggatattgatc

SCIREG2       TTGATGTCGTCTGGGTTGGTGACGCTTGCAGCGATTATGT   1454
IREG_REV      TTGATGTCGTCTGGGTTGGTGACGCTTGCAGCGATTATGT   1007
CONSENSUS     ttgatgtcgtctgggttggtgacgcttgcagcgattatgt

SCIREG2       ATAGTGTACATATATATCGAGTTCGGAAACACCTATTTCA   1494
IREG_REV      ATAGTGTACATATATATCGAGTTCGGAAACACCTATTTCA   1047
CONSENSUS     atagtgtacatatatatcgagttcggaaacacctatttca


SCIREG2       TTTCGACAAGTTGCTTGCCCTGTTTCTTTGAAAACCCAGC   1534
IREG_REV      TTTCGACAAGTTGCTTGCCCTGTTTCTTTAAAACCCGCTT   1087
CONSENSUS     tttcgacaagttgcttgccctgtttctttaaaacccgctt

SCIREG2       TTTCTTGTACAAAGTGGTCCC.............         1555
IREG_REV      TCTTGTACAAAGTTGGCATTATAAGAAAGCATTG         1121
CONSENSUS     t t  t  a a    g
```

AtIRT1:ScIRT1 LR Clone in pB2NADA

```
AtIRT1 attB  .....................GGGACAAGTTTGTACAA     18
LR ATIRT FOR atgatcgataattcgctagtgatatcACAAGTTTGTACAA     80
Consensus                                acaagtttgtacaa

AtIRT1 attB  AAAAGCAGGCTTACATTAAGTCAATAACCCCATAATCTTA     58
LR ATIRT FOR AAAAGCAGGCTTACATTAAGTCAATAACCCCATAATCTTA    120
Consensus    aaaagcaggcttacattaagtcaataaccccataatctta

AtIRT1 attB  TAAATGAAGTCAGTGACTATATGTTCATGTAAAAATGCGT     98
LR ATIRT FOR TAAATGAAGTCAGTGACTATATGTTCATGTAAAAATGCGT    160
Consensus    taaatgaagtcagtgactatatgttcatgtaaaaatgcgt

AtIRT1 attB  AAGTTCATAGCTAGATGTCCCTTTTTTAGGAGCACATGGA    138
LR ATIRT FOR AAGTTCATAGCTAGATGTCCCTTTTTTAGGAGCACATGGA    200
Consensus    aagttcatagctagatgtccctttttttaggagcacatgga

AtIRT1 attB  TTGACACATTAAACATTCATACCCGATTTCATGGAATTCT    178
LR ATIRT FOR TTGACACATTAAACATTCATACCCGATTTCATGGAATTCT    240
Consensus    ttgacacattaaacattcatacccgatttcatggaattct

AtIRT1 attB  ATTTGTCATTATCCTTTCCCAAAAACTTAGTTGTTTCTTA    218
LR ATIRT FOR ATTTGTCATTATCCTTTCCCAAAAACTTAGTTGTTTCTTA    280
Consensus    atttgtcattatcctttcccaaaaacttagttgtttctta

AtIRT1 attB  GTTATTAAACCTTGAAAAAAAAATCAATTCAAAAATGAAAA    258
LR ATIRT FOR GTTATTAAACCTTGAAAAAAAAATCAATTCAAAAATGAAAA    320
Consensus    gttattaaaccttgaaaaaaaaatcaattcaaaaatgaaaa

AtIRT1 attB  ATGACTTATATGTATGTCATTTTTGAAAACTAGACATCAA    298
LR ATIRT FOR ATGACTTATATGTATGTCATTTTTGAAAACTAGACATCAA    360
Consensus    atgacttatatgtatgtcatttttgaaaactagacatcaa

AtIRT1 attB  GCTATGAAAAGAGGAATCTATTTACCGAATAAAGGAGGAG    338
LR ATIRT FOR GCTATGAAAAGAGGAATCTATTTACCGAATAAAGGAGGAG    400
Consensus    gctatgaaaagaggaatctatttaccgaataaaggaggag

AtIRT1 attB  GAATCTACATTCATATGTAGTGTTTTTGGTGACTGGATTG    378
LR ATIRT FOR GAATCTACATTCATATGTAGTGTTTTTGGTGACTGGATTG    440
Consensus    gaatctacattcatatgtagtgttttggtgactggattg

AtIRT1 attB  CAACCATACACTGAGCACAAGCTGTCACGTGTGATGTTGT    418
LR ATIRT FOR CAACCATACACTGAGCACAAGCTGTCACGTGTGATGTTGT    480
Consensus    caaccatacactgagcacaagctgtcacgtgtgatgttgt

AtIRT1 attB  GCTCGTAGCATAAACTATAAAGTATAAACCTAATTAAGCT    458
LR ATIRT FOR GCTCGTAGCATAAACTATAAAGTATAAACCTAATTAAGCT    520
Consensus    gctcgtagcataaactataaagtataaacctaattaagct

AtIRT1 attB  TAAAGTTAAAAATAGTACAAGATTAAAAAGTAAACGTTTG    498
LR ATIRT FOR TAAAGTTAAAAATAGTACAAGATTAAAAAGTAAACGTTTG    560
Consensus    taaagttaaaaatagtacaagattaaaaagtaaacgtttg

AtIRT1 attB  TGAGAGTTGACTAACAGAACTATAGAAGACGCATGGGGTT    538
LR ATIRT FOR TGAGAGTTGACTAACAGAACTATAGAAGACGCATGGGGTT    600
Consensus    tgagagttgactaacagaactatagaagacgcatggggtt

AtIRT1 attB  GACTTCAAGCGGTCTGCAAACGGCATGAACAATGCATTCA    578
LR ATIRT FOR GACTTCAAGCGGTCTGCAAACGGCATGAACAATGCATTCA    640
Consensus    gacttcaagcggtctgcaaacggcatgaacaatgcattca
```

```
AtIRT1 attB  AACGATAATATAAACCTTAATTCCAATATAGCCTTTTTCA    618
LR ATIRT FOR AACGATAATATAAACCTTAATTCCAATATAGCCTTTTTCA    680
Consensus    aacgataatataaaccttaattccaatatagccttttttca

AtIRT1 attB  AATTTAAAAACTCGAATGTTTACGAAACTATAACTTGATA    658
LR ATIRT FOR AATTTAAAAACTCGAATGTTTACGAAACTATAACTTGATA    720
Consensus    aatttaaaaactcgaatgtttacgaaactataacttgata

AtIRT1 attB  AAACTGATTGAAATAAAATGTACTTACATTTTTTTAATAA    698
LR ATIRT FOR AAACTGATTGAAATAAAATGTACTTACATTTTTTTAATAA    760
Consensus    aaactgattgaaataaaatgtacttacatttttttaataa

AtIRT1 attB  AATATACTTAAATTTTTTAAATAAAATATACTTTAAAATA    738
LR ATIRT FOR AATATACTTAAATTTTTTAAATAAAATATACTTTAAAATA    800
Consensus    aatatacttaaatttttttaaataaaatatactttaaaata

AtIRT1 attB  TACTTAATTTTTTAAAATCATACTATGGATTTTAAACAAA    778
LR ATIRT FOR TACTTAATTTTTTAAAATCATACTATGGATTTTAAACAAA    840
Consensus    tacttaatttttttaaaatcatactatggattttaaacaaa

AtIRT1 attB  ATAAATTGTCACTGAGTCATATTTCTTTGGATAAAAGAAC    818
LR ATIRT FOR ATAAATTGTCACTGAGTCATATTTCTTTGGATAAAAGAAC    880
Consensus    ataaattgtcactgagtcatatttctttggataaaagaac

AtIRT1 attB  TTTACATAAGTGATTGGTACAATATAATAATATGCAAATT    858
LR ATIRT FOR TTTACATAAGTGATTGGTACAATATAATAATATGCAAATT    920
Consensus    tttacataagtgattggtacaatataataatatgcaaatt

AtIRT1 attB  TAAGACATTTTATTAAGTTTATATTATCGTTTAAATGCAT    898
LR ATIRT FOR TAAGACATTTTATTAAGTTTATATTATCGTTTAAATGCAT    960
Consensus    taagacattttattaagtttatattatcgtttaaatgcat

AtIRT1 attB  TTTCTATCAACTTAATAAAATCTAGCTAGGTACTTTCATA    938
LR ATIRT FOR TTTCTATCAACTTAATAAAATCTAGCTAGGTACTTTCATA   1000
Consensus    tttctatcaacttaataaaatctagctaggtactttcata

AtIRT1 attB  ATTCTATCATTATTCCTTTTGAACGATTAATTACTTTATT    978
LR ATIRT FOR ATTCTATCATTATTCCTTTTGAACGATTAATTACTTTATT   1040
Consensus    attctatcattattccttttgaacgattaattactttatt
```

```
AtIRT1 attB  TTTCTATCAACTTAATAAAATCTAGCTAGGTACTTTCATA    938
LR ATIRT FOR TTTCTATCAACTTAATAAAATCTAGCTAGGTACTTTCATA   1000
Consensus    tttctatcaacttaataaaatctagctaggtactttcata

AtIRT1 attB  ATTCTATCATTATTCCTTTTGAACGATTAATTACTTTATT    978
LR ATIRT FOR ATTCTATCATTATTCCTTTTGAACGATTAATTACTTTATT   1040
Consensus    attctatcattattccttttgaacgattaattactttatt

AtIRT1 attB  TTCTCGTTTCACACTGGAAGCAAACGAGTAATCAAATAAA   1018
LR ATIRT FOR TTCTCGTTTCACACTGGAAGCAAACGAGTAATCAAATAAA   1080
Consensus    ttctcgtttcacactggaagcaaacgagtaatcaaataaa

AtIRT1 attB  ATTTCGTCATCATCGACAATAAATGGCATGAAAACAACCA   1058
LR ATIRT FOR ATTTCGTCATCATCGACAt..AATGGCATGAAAACA.CCA   1117
Consensus    atttcgtcatcatcgaca   aatggcatgaaaaca cca

AtIRT1 attB  GCACACATTACTATATAAACCGATTCATCTAGTAAGTCCA   1098
LR ATIRT FOR GCACACAT.ACTATATAA..CGATTCATCTAGTAg..tCA   1152
Consensus    gcacacat actatataa  cgattcatctagta    ca

AtIRT1 attB  AAGTCACACAAACATTAAACAATCTACAACTTTGTATACA   1138
LR ATIRT FOR AAGTC.................................   1157
Consensus    aagtc
```