IDENTIFYING GENES RELATED TO SYMBIOTIC NITROGEN FIXATION IN *Medicago truncatula Tnt1* INSERTION MUTANTS USING A FORWARD GENETICS APPROACH

by

K. Nicole Holland

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master of Science in Plant and Soil Sciences

Spring 2020

IDENTIFYING GENES RELATED TO SYMBIOTIC NITROGEN FIXATION IN *Medicago truncatula Tnt1* INSERTION MUTANTS USING A FORWARD GENETICS APPROACH

by

K. Nicole Holland

Approved: _____
D. Janine Sherrier, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____
Randall J. Wisser, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____
Rebecca Dickstein, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____
Erik Ervin, Ph.D.
Chair of the Department of Plant and Soil Sciences

Approved: _____

Mark Rieger, Ph.D.
Dean of the College of Agriculture and Natural Resources

Approved: _____

Douglas J. Doren, Ph.D.
Interim Vice Provost for Graduate and Professional Education and
Dean of the Graduate College

ACKNOWLEDGMENTS

I am so grateful for everyone who has helped support me through my educational journey. When I made the decision to return to college after spending nearly ten years out of the classroom, I never thought that I would have made it this far. I am thankful for all the knowledge and experience that I have gained. Most importantly, I am thankful for all the amazing people that I have met and worked with along the way. I have had the pleasure of working with some great mentors and fellow students that have taught me a lot and furthered my love for science.

I am very thankful for my family, who have given me nothing but love and support. My parents, my mother and father in-law, Nancy and Gary, and especially my husband, Mike, have been sources of encouragement, and I would not have been able to do this without them. I am also grateful for my two daughters, Anna and Lily, who continue to make me smile every day and motivate me through even the toughest days.

I would like to especially thank my primary advisor, Dr. Janine Sherrier. She has been such a great mentor to me and has continued to support and guide me throughout my career. She has shared her great knowledge of legumes and plant physiology in addition to helping me be a better scientist. I would also like to thank my advisor, Dr. Randy Wisser, who welcomed me into his laboratory and has always been quick to help me in any way he can. I am grateful to have worked with him and have furthered my understanding of genetics and the use of powerful bioinformatic tools to further my research. Last but not least, I would like to thank my advisor, Dr. Rebecca Dickstein. I am so happy that I had the chance to meet and work with her. I

learned so much in the short time that we worked together in the laboratory. She has always been quick to respond to any of my questions and has been a great source of support and knowledge since the day that I met her.

I would also like to thank all the other scientists who have supported me and my research. Everyone at the Noble Research Institute, especially Raj Nandety, have been extremely helpful and have been a pleasure to work with. At the Delaware Biotechnology Institute, I would like to thank Shannon Modla in the Bioimaging Center for sharing her knowledge and expertise in microscopy as well as my good friends Tim Chaya and Victoria Gundlah-Mooney for all of their feedback and support throughout my project. I would also like to thank Xiaolan Rao at the University of North Texas for her help with analyzing sequencing data in this research. Finally, I would like to thank all of the Sherrier lab members, Usha Bishnoi, Cherish Warner, Ryan Delpercio, and Angel Huang who have always supported and encouraged me along the way.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

atms	atmospheres
BLAST	Basic Local Alignment Search Tool
bp	base pairs
C	Celsius
ch	chromosome
DNA	deoxyribonucleic acid
dpi	days post inoculation
EMS	ethylmethane sulfonate
FNB	fast-neutron bombardment
IN	integrase
Kb	kilobase
LTR	long-terminal repeats
Mbp	megabase pairs
N_2	nitrogen
PCR	polymerase chain reaction
PIPES	Piperazine-1, 4-bis (2-ethanesulfonic acid)
PR	protease
rcf	relative centrifugal force
rpm	revolutions per minute
RH	ribonuclease H
RT	reverse transcriptase

T-DNA	transfer DNA
TAE	tris acetate EDTA
Tnt1	tobacco (*Nicotiana tabacum*) type I retrotransposon
T_m	annealing temperature
TY	tryptone yeast
V	volts
w/v	weight by volume
WGS	whole genome sequencing
WT	wild type
X-gal	5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside

ABSTRACT

Nitrogen is an essential nutrient required by all plants. Although the earth's atmosphere consists of approximately 78% nitrogen, it exists in an unusable form for plants. Current agricultural practices rely on industrial nitrogen fixation to produce fertilizers that improve crop productivity. However, this method requires the use of precious natural resources and can cause a variety of negative impacts on the environment and human health. Biologically fixed nitrogen is a favorable alternative accessible to leguminous plants through their symbiotic relationship with soil bacteria called rhizobia. Genetic studies aimed at identifying plant genes related to successful nodulation and nitrogen fixation are underway to help better understand this infection process and improve the performance of legumes in agriculture. To date, over 21,000 *Tnt1* mutant lines of *Medicago truncatula*, a model legume used for genetic research, have been created. Several lines displaying defects in symbiotic nitrogen fixation (SNF) have been identified and are referred to as Fix- mutants. In this research, segregation analysis of the Fix- mutant NF18598 indicated loss of SNF function could be governed by a dominant mutation. A total of 114 *Tnt1* insertions were identified in the NF18598 genome using three different sequencing methods. *Tnt1*-capture sequencing (SC) found the largest number of insertions (70) while the other two methods, thermal asymmetric interlaced PCR (TAIL-PCR) and whole genome sequencing (WGS), found a similar number of insertions (40 and 41, respectively). It was found that 50% of the insertions were located in coding regions. One of the *Tnt1* insertions located on chromosome 5 is found within the exon of a phosphatase 2C

family protein gene (Medtr5g009370) and is upregulated in nodules. While there was a significant SNF association with Medtr5g009370 in the R_1 generation, this association was not detected in the R_2 generation. Light microscopy analysis of nodules displayed clear differences between mutant samples identified as either Fix+ or Fix-. The Fix- mutant nodules displayed some cellular organization but distinct developmental zones could not be identified. Infection threads were observed in Fix- nodules that appeared to contain bacteria that had not been released; however, further microscopy analysis must be conducted to verify the presence of rhizobia. In addition, further research involving a larger sample population would help verify if the insertion located in Medtr5g009370 is responsible for the defective nodule phenotype observed and if the phosphatase gene is involved in symbiotic nitrogen fixation.

Chapter 1

INTRODUCTION

1.1 Importance of the Legume-Rhizobia Symbiosis

Legumes are a diverse group of plants that can be found growing in all parts of the world. The Leguminosae family is the third largest group of flowering plants and includes soybeans, peanuts, fresh peas and beans, and pulses such as chickpeas and lentils (Bruneau et al., 2013). Other legumes, such as alfalfa and clover, are used as forage crops for livestock. Legumes are an important staple of human diets throughout the globe and are an excellent source of plant protein. The success of legumes in diverse environments and their value as protein-rich food and fodder are underpinned by the ability of legumes to obtain nitrogen to support growth and development even when growing in soils with insufficient levels of this key element.

Nitrogen is an essential element for all forms of life on Earth and often is a growth limiting factor for life in terrestrial and marine environments. Although approximately 78% of the Earth's atmosphere is composed of di-nitrogen (N_2), this form is largely unavailable to living organisms. A large amount of energy is required to break the strong triple bond between the nitrogen molecules, and high-energy natural processes such lightning and photochemical reactions lead to the conversion of nitrogen into chemically-reduced forms within the atmosphere. Together, this accounts for only 10% of naturally occurring nitrogen fixation on Earth. A small number Bacteria and Archaea species can break this bond enzymatically accounting for the remaining 90% of naturally occurring nitrogen fixation (Taiz & Zeiger, 2002;

Galloway et al., 2004). Biological nitrogen fixation requires a specialized environment and is also an energy-intensive process. Despite the energy advantage conferred through the bacterial enzyme nitrogenase, only a subset of plants have evolved the ability to form these beneficial endo-symbioses.

Legume plants are well known for their ability to form symbiotic relationships with free-living soil bacteria collectively referred to as rhizobia to support nitrogen fixation (Graham & Vance, 2003). The plant provides the micro-aerobic environment required for nitrogen fixation by the bacterially-encoded nitrogenase enzyme and carbohydrates to support the high energy requirements of this process. In return, the bacteria provide the host plant with bioavailable nitrogen which can be incorporated into amino acids, nucleic acids, and proteins. While nitrogen-fixing symbioses occur between other microbes and hosts such as the symbiosis between alder trees and water ferns with bacteria in the genus *Frankia* and cyanobacteria, respectively, the most commonly studied nitrogen-fixing symbiotic relationship between legumes and rhizobia. Symbioses involving diverse bacteria genera such as *Azorhizobium*, *Bradyrhizobium*, *Photorhizobium*, *Rhizobium*, and *Sinorhizobium* have the potential to contribute to a sustainable source of a nutritious diet for the growing human population.

Until the early 1900s, the primary method of producing crops was the cultivation of legumes and the use animal waste and wood ash to provide nitrogen (Pankievicz, Irving, Maia, & Ané, 2019). Later, the invention of the Haber-Bosch process allowed for nitrogen fixation at an industrial scale and a rapid increase in food production. However, the Haber-Bosch process utilizes a high temperatures over 200°C and pressure exceeding 200 atms to convert atmospheric nitrogen into

ammonia which is the initial product of many agricultural fertilizers (Taiz & Zeiger, 2002). Therefore, this process has contributed to the increased carbon footprint of food production.

Research by Galloway et al. (2004) examined the rates of both naturally occurring and industrial nitrogen fixation over time and found that between 1860 and 1995 the population increased by a factor of 4.5, yet the industrially-produced nitrogen increased by a factor of 10. Much of the reactive nitrogen is used for agricultural purposes while some is used in the creation of biofuel or other industrial products. However, nearly half of the industrially produced nitrogen used in agriculture is not taken up by crops and enters the atmosphere or hydrosphere as pollution. It is estimated that 14% is released into the air leading to poor air quality and climate change, while approximately 13% leaches into surface and groundwater leading to eutrophication and associated human health problems (Suddick, Whitney, Townsend, & Davidson, 2013). Other research by Graham & Vance (2003), estimated that 30-50% of nitrogen fertilizer applied to legume crops is not assimilated by plants and instead becomes an environmental pollutant. The growing negative impact on both humans and the environment coupled with an increasing global population makes clear the need for more sustainable ways to increase agricultural production and efficiency. Exploiting more natural systems, such as the legume-rhizobia symbiosis, offers an opportunity to increase agricultural efficiency while minimizing negative consequences.

Despite the ecological significance of legume crops, cereals reign as the dominant crops of most agricultural systems and often receive significantly larger amounts of funding for scientific research when compared to legumes (Magrini et al.,

2018). However, a better understanding of the innate molecular mechanisms which allow the complex yet efficient relationships between legumes and rhizobia could be used to help alleviate our dependence on synthetic fertilizers and to transition to more sustainable agricultural systems.

1.2 Nodule Development

The beneficial relationship between legume plants and rhizobia bacteria is not an obligatory symbiosis. Legumes can be cultivated without rhizobia, just as the bacteria are able to live freely within soil without any plant host. However, when nitrogen concentrations in the soil are not sufficient, both of the organisms are stressed, and they engage in a complex exchange of signals which initiates symbiosis. The infection process and nodule initiation ultimately results in a newly developed plant structure referred to as a root nodule (Taiz & Zeiger, 2002).

The development of a root nodule accommodates the need of the micro-aerobic conditions required by the key enzyme involved in nitrogen fixation, nitrogenase. The nitrogenase enzyme complex catalyzes a reaction that results in the chemical reduction of nitrogen into ammonia (Ferguson et al., 2010). The oxygen binding protein leghemoglobin is abundantly found in the cytoplasm of infected root nodule cells where it scavenges oxygen and gives the symbiotic organ its characteristic pink color that indicates the occurrence of successful nitrogen fixation (Figure 1) (Taiz & Zeiger, 2002). The initiation of the symbiosis process begins when available nitrogen is limited and the plant releases signaling compounds into the soil surrounding its roots, an area known as the rhizosphere. Plant-derived signaling molecules in the form of phenolic compounds called flavonoids trigger the expression of specific genes in the rhizobia bacteria living freely within the soil. The expression of nodulation genes

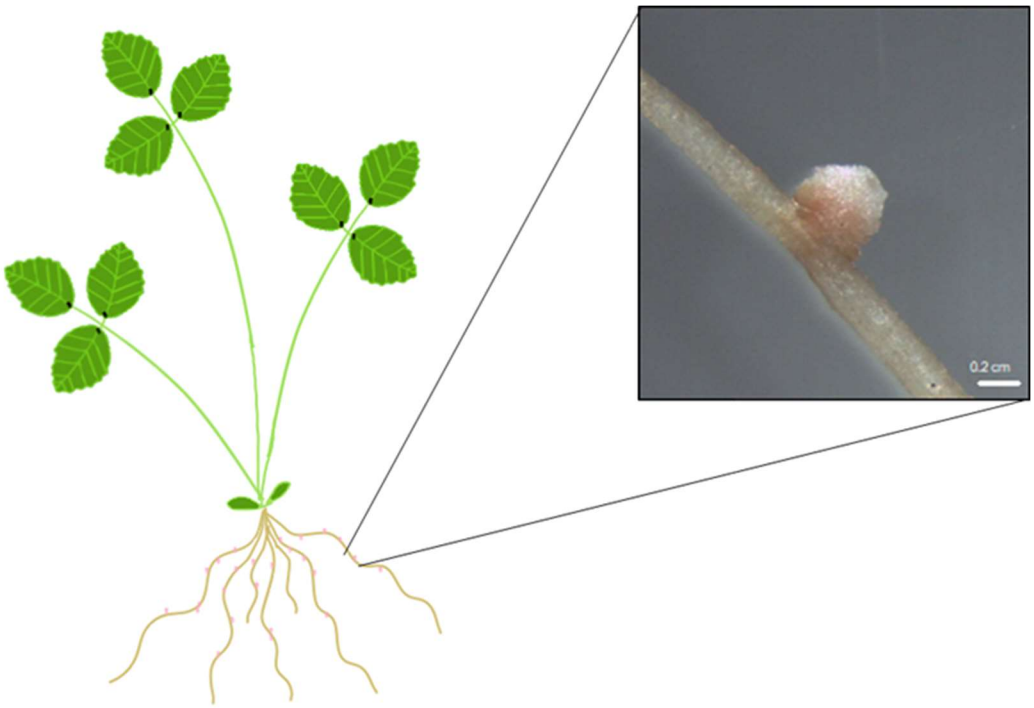

Figure 1 **Root nodules formed on model legume plant *Medicago truncatula*.** Root nodule shown is 15 dpi. The pink color indicates that rhizobia bacteria are actively fixing nitrogen within the nodule.

leads to the synthesis of Nod factors which are lipochitooligosaccharides that initiate the development of nodules on the plant roots (Via et al., 2016). Nod factors also play an important role in controlling the specificity of the symbiotic relationship. In most cases, each species of rhizobia bacteria are limited in the number of legume host species they are able to nodulate, with the exception of a few who are able to induce nodules on a wider range of hosts (Ferguson et al., 2010). The successful perception of Nod factor by the plant allows entry of the bacteria into root hair cells via the development of a special structure called an infection thread. After making contact with a root hair, the rhizobia initiate changes leading to the formation of infection

threads originating from the plant's own cell wall. The plant's plasma membrane also grows and surrounds the penetrating bacterial cells. Rhizobia continue to divide and multiply within the infection thread. After reaching cells of the nodule primordium, the plant-derived extracellular matrix surrounding the bacterial eventually thins, allowing entry of the bacteria into the plant cell cytoplasm via an endocytosis-like process (Catalano, 2004). A plant-derived membrane forms around each released bacterium leading to a specialized structure called a symbiosome. This important step in the infection process allows the bacteria to differentiate into mature bacteroids capable of fixing nitrogen within the symbiosome (Sergeevich, Alexandrovich, Yurievna, Yurievich, & Anatolievich, 2015). This complex process, which is outlined in Figure 2, allows for the mutually beneficial relationship between organisms from different domains of life to take place.

Root nodules on legume plants can develop into two different morphological types, referred to as either determinate or indeterminate. The type of nodule is determined by the legume species, and the major difference between them is the existence of a persistent meristem. Legumes such as soybean and lima bean produce determinate nodules that lack a persistent meristem and appear spherical in shape. In contrast, new cells at the apical meristem in indeterminate nodules continue to divide and become infected with bacteria leading to the cylindrical shape that is seen on the nodules of plants such as pea, alfalfa, clover, and the model legume *Medicago truncatula.* Indeterminate nodules contain a developmental gradient of cells that are categorized into different nodule zones. These developmental stages, which are outlined in Figure 3, are referred to as the meristematic zone, invasion (or infection) zone, interzone, nitrogen-fixation zone, and senescence zone (Ferguson et al., 2010).

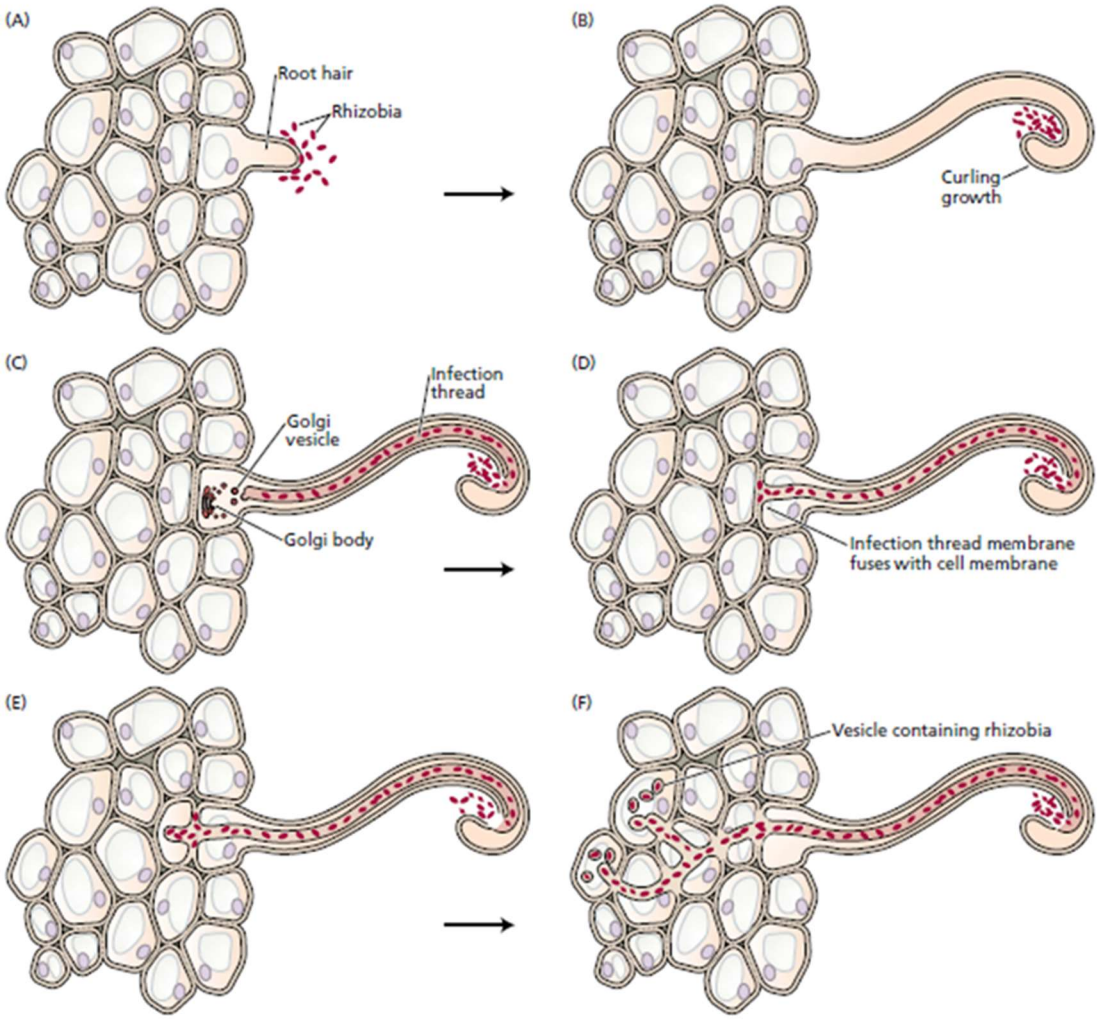

Figure 2 **The infection process during nodule formation.** (A) The rhizobia bacteria bind to a root hair in response to signal molecules exuded by the plant known as nod factors. (B) The plant responds by altering its root growth (curling). (C) The bacteria grow into the plant tissue between cells and develop an infection thread. (D) The membrane of the infection thread fuses with the cell membrane. (E) New infection threads are formed after the rhizobia are released into the apoplast. (F) The bacteria are released in membrane-bound bacteroids and begin to fix nitrogen. Figure used with permission (Taiz and Zeiger, 2002).

The meristematic zone contains the youngest plant cells and is devoid of bacteria. The invasion zone includes the area of both infection thread penetration and the formation and continued division of symbiosomes. The nitrogen-fixation zone is the stage at which rhizobia have differentiated into bacteroids and are able to begin converting atmospheric nitrogen. Finally, bacteroids in older nodules begin to degrade which leads to the developmental stage known as the senescence zone (Xiao et al., 2014). Various microscopy techniques have illustrated the different nodule zones and clear morphological differences can be seen when comparing nodules from wild-type (WT) plants versus plants containing mutations in genes related to nodule development (Figure 4).

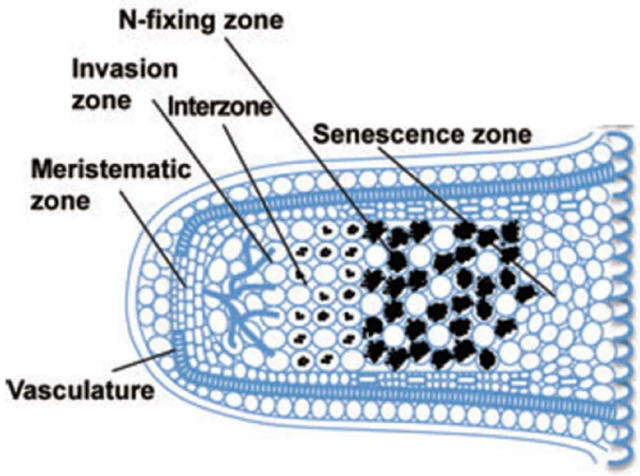

Figure 3 **Developmental zones of *M. truncatula* indeterminate root nodule.** The meristematic zone, invasion (or infection) zone, interzone, nitrogen-fixation zone, and senescence zone are indicated. Figure used with permission (Ferguson et al., 2010).

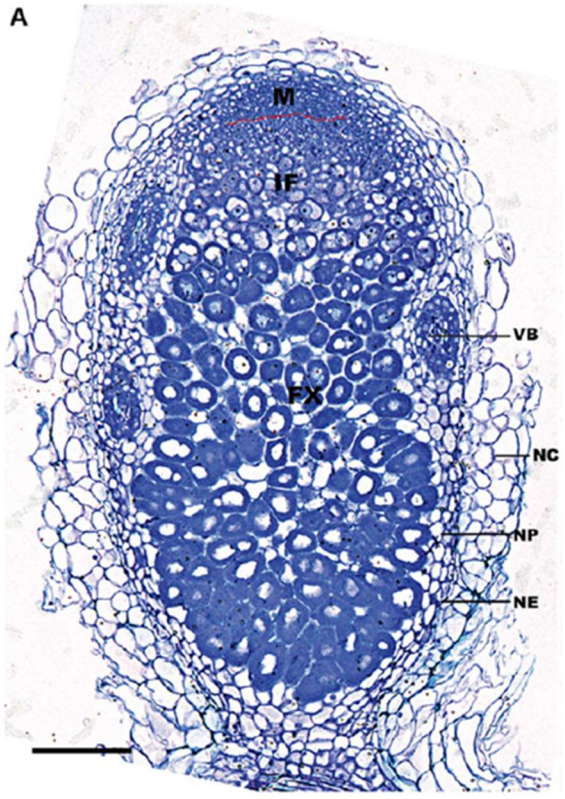

Figure 4 **Longitudinal section of a *M. truncatula* root nodule 14 dpi showing different nodule developmental zones.** M is the meristematic zone, IF is the infection or invasion zone, and FX is the fixation zone. The red line indicates the boundary between the meristematic and infection zones. Other cellular features highlighted include vascular bundles (VB), nodule parenchyma (NP), the nodule endodermis (NE), and the nodule cortex (NC). Scale bar is 75 μm. Figure used with permission (Xiao et al., 2014).

1.3 *Medicago truncatula*: A Model Legume

One of the legume species that produces indeterminate root nodules is *Medicago truncatula*, a forage crop native to the Mediterranean basin that is a close relative of pea, clover, and alfalfa. It recent decades, *M. truncatula* has emerged as an important model legume species that has aided researchers in genetic studies aimed at

better understanding the development of root nodules and other key components of the symbiotic relationship (Ané, Zhu, & Frugoli, 2007; Roy et al., 2020). Other legumes that are more agriculturally relevant, such as soybean and peanut, have large complex genomes that can pose challenges in the laboratory. *M. truncatula*, has a relatively small genome of around 500 Mbp, a short life cycle, is diploid, and is also self-fertilizing (Kang, Li, Sinharoy, & Verdier, 2016). Due to the ease of transformation, *M. truncatula* also provides a wealth of genetic resources created from various mutagenesis and transposon tagging projects making the species an attractive model to researchers (Roy et al., 2020).

A new and improved version of the *M. truncatula* genome (Mt4.0) was released and further annotated in 2014 with the majority (93%) of sequences being anchored to specific chromosomes (Tang et al., 2014). The JCVI *Medicago truncatula* website (http://blast.jcvi.org/Medicago-Blast/) contains a variety of genomic and bioinformatic tools relative to the model species including a BLAST search tool, an interactive genome browser, and the ability to search for specific genes or genetic coordinates. Transcriptomic studies utilizing microarray technologies such as the Affymetrix GeneChip® have also allowed for the creation of the *Medicago truncatula* Gene Expression Atlas (https://mtgea.noble.org) which enables researchers to view the expression of genes of interest in specific plant organs (Benedito et al., 2008). Continuing advances in plant genomic research and next-generation sequencing technologies will only further develop *M. truncatula* as a model species in the future.

1.4 *M. truncatula Tnt1* Mutant Population

With any model species, one of the keys to understanding genetic function relies on the creation of large mutant plant populations that can be used for

comparative research. A variety of methods exist that can induce genetic mutations such as chemical mutagenesis using ethylmethane sulfonate (EMS) which results in point mutations and fast-neutron bombardment (FNB) that causes deletion mutations within the genome. Other popular methods of mutagenesis involve the use of transfer DNA (T-DNA) and retrotransposons to introduce insertional mutations (Tadege, Wang, Wen, Ratet, & Mysore, 2009). While T-DNA insertional mutagenesis has been successful in other model species such as *Arabidopsis*, it is not a feasible method for creating a large mutant population in larger genomes such as *M. truncatula* (Tadege et al., 2008). Insertional mutagenesis using the tobacco type I retrotransposon (*Tnt1*) originally isolated from *Nicotiana tabacum* has been shown to be an excellent tool for creating a large *M. truncatula* mutant population that can be used to study gene function (Cheng, Wen, Tadege, Ratet, & Mysore, 2011). Retrotransposons are similar to retroviruses because they contain all the necessary genetic components to replicate and transpose themselves into host genomes via a copy and paste mechanism (d'Erfurth et al., 2003). Type I retrotransposons like *Tnt1* are characterized by their long-terminal repeats (LTR) and genes that encode for functions that are similar to the *gag* and *pol* genes typical of retroviruses (Finnegan, 2012). The structure of *Tnt1* is detailed in Figure 5 and includes genes that encode for a capsid-related protein (GAG), a protease (PR), an integrase (IN), a reverse transcriptase (RT), and a ribonuclease H (RH). Two LTR regions flank each end of *Tnt1* and contain sequences that are used for recovering the location of insertional mutations within the *M. truncatula* genome (d'Erfurth et al., 2003). Research by Tadege et al. (2008) and Sun et al. (2019) examined the genetic locations of insertions and found that *Tnt1* prefers gene-rich regions, with over 34% of the insertions being found in exons. This

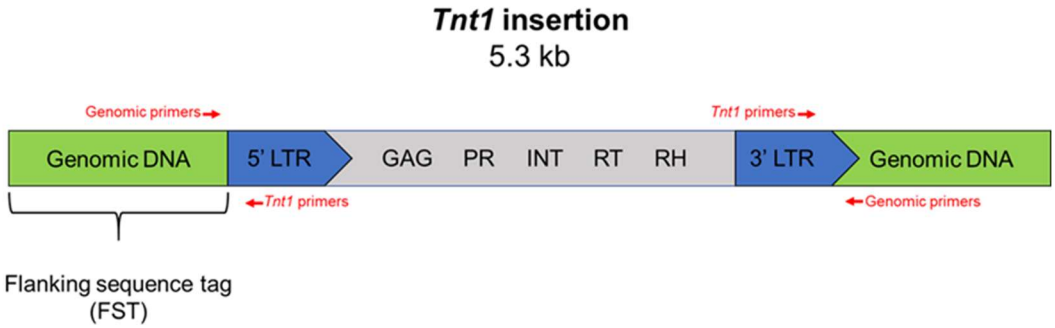

Figure 5 **Diagram of *Tnt1* genomic structure.** *Tnt1* is a 5.3 Kb retrotransposon element that encodes for a capsid protein (GAG), a protease (PR), an integrase (INT), a reverse transcriptase, and a ribonuclease H (RH). A 610-bp long-terminal repeat (LTR) region is located at each end of the *Tnt1* insertion. The genomic plant DNA that flanks each end of the *Tnt1* sequence is referred to as a flanking sequence tag (FST). A genomic primer located within the plant's genome is used in combination with a *Tnt1* primer sequence to genotype plants at different insertion locations.

favorable characteristic makes *Tnt1* an advantageous method of mutagenesis for functional genetic studies.

The creation of the *Medicago truncatula Tnt1* mutant population started with a single transgenic line (Tnk88-7-7) that contained five *Tnt1* insertions that were cloned into *M. truncatula* using *Agrobacterium*-mediated transformation. Transposition of *Tnt1* was activated via somatic embryogenesis from transgenic Tnk88-7-7 leaf explants (Tadege et al., 2008). This process is further detailed in Figure 6. Since the start of the project, over 21,000 independent mutant lines have been generated at the Noble Research Institute and are available to the research community (Cheng et al., 2017). Several preliminary phenotypic screenings have also been performed by the community of legume researchers. In research by Pislariu et al. (2012), approximately 30% of the 9,300 *Tnt1* lines that were screened displayed visible morphological or

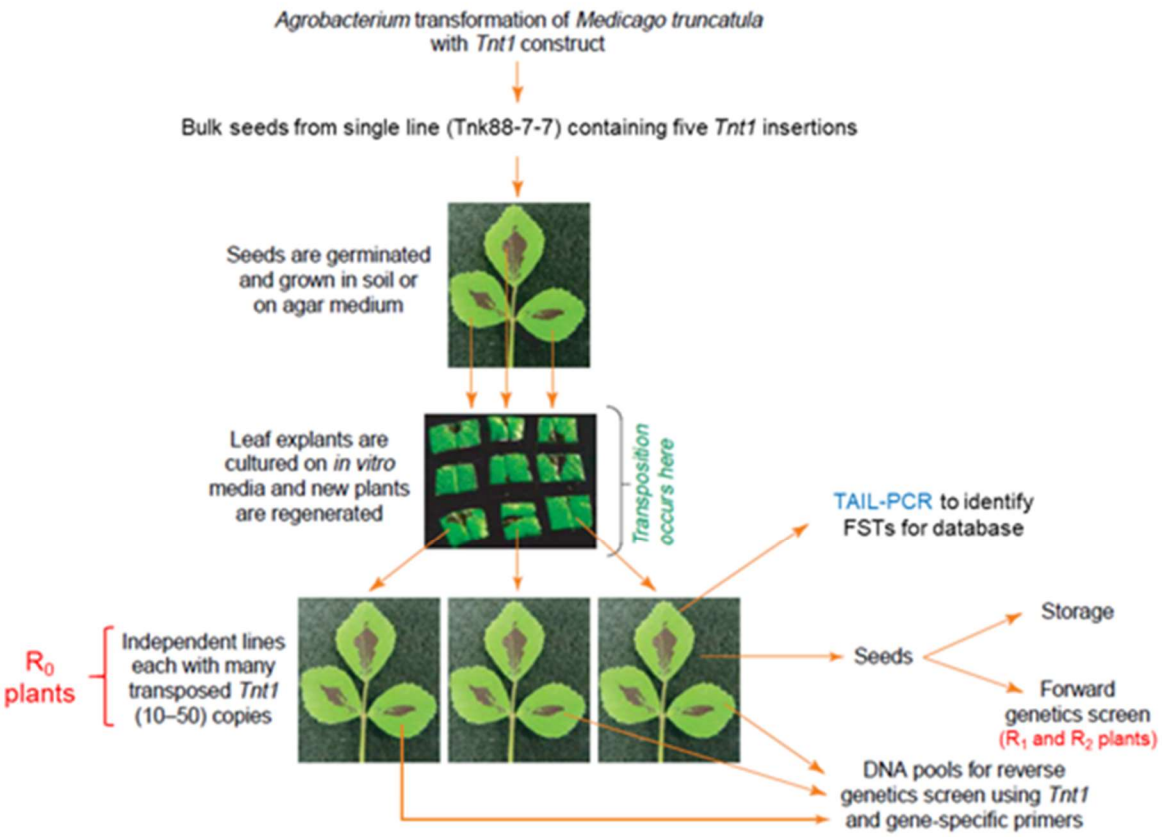

Figure 6 **Schematic representation of insertional mutagenesis using the *Tnt1* retrotransposon in *Medicago truncatula*.** *Tnt1* is introduced into *M. truncatula* R108 ecotype via *Agrobacterium*-mediated transformation. All mutant lines originate from one parental line (Tnk88-7-7) containing approximately five copies of *Tnt1*. Transposition is activated via somatic embryogenesis from leaf explants. Independent lines (R_0 plants) are generated from individual calli. Seeds from independent lines are stored and can be used to perform both forward and reverse genetic screens (R_1 and R_2 plants). Figure adapted with permission (Tadege et al., 2005).

nodule developmental phenotypes which are outlined in Table 1. Specifically, 51 mutants were identified that were able to form root nodules that were white in color indicating ineffective nitrogen fixation (Nod+ Fix-). Additional studies involving this group of mutant lines has the potential to identify novel genes required for SNF which will improve our understanding of the symbiotic relationship and help to increase the use of legumes in agriculture.

Phenotype	Phenotyping Criteria	
	Nodule Development, Appearance, and Infection Status	Overall Plant Development, Leaf Color
Wild type	Elongated, infected, and intense-pink nodules due to the accumulation of leghemoglobin, an indicator of efficient nitrogen fixation	Well-developed plants with healthy shoots and root systems, dark-green leaves, suggesting no nutrient deficiencies
Nod−	Absence of emerged nodules at 21 d post inoculation but may later develop a few small white bumps and a range of infection blocks; absence or reduction of root hair deformation; infection threads may form but terminate prematurely inside root hairs and do not penetrate past the epidermis or outer cortical cells into nodule primordia	Stunted plants with pale-green or yellowish leaves, due to severe nitrogen starvation
Nod+ Fix−	White, infected nodules, inefficient nitrogen fixation	Stunted plants with pale-green or yellowish leaves, due to severe nitrogen starvation
Nod+ Fix+/−	Infected nodules with pale-pink appearance due to less leghemoglobin, suggesting less efficient nitrogen fixation than in the wild type	Plant development is stunted but the yellowing of leaves is not as severe as in Nod− and Fix− mutants

Table 1 **Classification criteria for nodule phenotypes.** Table used with permission (Pislariu et al., 2012).

1.5 Research Objectives and Experimental Approach

The goal of this research was to identify *M. truncatula Tnt1* mutant lines that displayed the Nod+ Fix- (also referred to as Fix-) phenotype indicative of ineffective nitrogen fixation with a specific emphasis on defects in symbiosome formation or persistence. In addition, a goal of this research was to identify potential locations of insertions causing the defect using next-generation sequencing technologies and molecular marker mapping. As previously stated, *Tnt1* preferably targets exons within the genome. Therefore, it was hypothesized that the *Tnt1* insertion causing the Fix- phenotype would be located within a gene potentially related to successful SNF. Additionally, mutant nodules were examined at the cellular level using light microscopy to identify the developmental abnormalities related to altered gene function caused by the *Tnt1* insertional mutation.

Chapter 2

MATERIALS AND METHODS

2.1 Seed Germination and Plant Growth

Medicago truncatula ecotype R108 and *Tnt1* mutant seeds were chemically scarified for 6 minutes using concentrated sulfuric acid and rinsed thoroughly with sterile water. The seeds were then surface sterilized for 3 minutes using full strength sodium hypochlorite, rinsed thoroughly with sterile water, and allowed to imbibe overnight at 4°C. The following day, the seeds were rinsed with sterile water for 5 to 6 hours to remove any compounds that may inhibit germination. The fully imbibed seeds were than plated onto 0.8% (w/v) water agar plates and vernalized at 4°C in the dark for 3 to 7 days. The germinated seeds were sown into Ray Leach Cone-tainers™ (Stuewe and Sons, Tangent, OR) containing a 2:1 mixture of autoclaved Turface MVP® calcined clay and coarse vermiculite. Plants were grown in a Conviron® at 55% relative humidity on a day/night cycle of 16/8 hours and 22°C/18°C. At the time of planting, seedlings were fertilized with one-half strength Lullien nutrient solution (Lullien, Barker, de Lajudie, & Huguet, 1987) containing 1 mM of NH_4NO_3.

2.2 Bacterial Growth and Inoculation

A single colony of *Sinorhizobium meliloti* strain *Sm1021* carrying the *hemA:lacZ* reporter (Ardourel et al., 1994) was selected from a freshly grown TY agar plate (5 g/L tryptone peptone, 3 g/L yeast extract, 2.5 g/L Bacto™ Agar, and 0.9 g/L $CaCl_2$ $2H_2O$) prepared with 10 µg/ml tetracycline, 50 µg/ml streptomycin, and X-gal.

The bacterial colony was used to inoculate 50 ml conical tubes containing TY broth (5 g/L tryptone peptone, 3 g/L yeast extract, and 0.9 g/L $CaCl_2$ 2H2O). Cultures were grown at 28°C and 220 rpm for approximately 36 hours until the cultures reached mid-log phase. Cultures were then centrifuged at 3,220 rcf and 4°C for 10 minutes to pellet bacterial cells, supernatant discarded, and the pellets washed with a 10 mM $MgCl_2$ solution before being centrifuged again at 3,220 rcf and 4°C for 10 minutes. The resulting supernatant was discarded, and the bacterial pellets were resuspended using one-half strength nitrogen-free Lullien solution or sterile deionized water. Six days after planting, each seedling was inoculated with 1 ml of the resuspended bacterial culture, representing approximately $1.0x10^8$ colony forming cells per plant. Inoculated plants were watered regularly with nitrogen-free Lullien solution or sterile deionized water until phenotypic screening was performed.

2.3 Phenotypic Screening

Phenotypic screening of nodules was performed at 15 days post inoculation (dpi). Plants were gently removed from growth substrate and root systems were rinsed in sterile deionized water. Mutant plant nodules were compared to wild type (WT) R108 nodules. Nodule phenotypes were characterized as either Nod-, Nod+ Fix+, Nod+ Fix-, or Nod+ Fix+/- based on the criteria listed in Table 1. Plant phenotypes were also observed. Characteristics such as size of shoot and root systems and degree of redness in leaves were recorded. After phenotypic screening was performed, plants were transplanted to 4-inch square pots containing a 3:1 mixture autoclaved Cornell mix and Turface MVP® calcined clay. Plants were watered regularly with Lullien solution or sterile deionized water.

2.4 Preparation of Genomic DNA

Fresh leaf tissue weighing 50-100 mg was ground using a sterile microcentrifuge pestle and tube. Samples were homogenized using a Vortex-Genie® 2 mixer (Scientific Industries, Bohemia, NY) and a vortex adapter (Qiagen, Germantown, MD) at maximum speed for 10 minutes. Genomic DNA was extracted using a DNeasy® Plant Pro Kit (Qiagen, Germantown, MD) following the manufacturer's protocol and using a final elution volume of 50 µl. Genomic DNA was stored at -20°C.

2.5 Primer Design and PCR Genotyping

Custom DNA primers were designed using R108 genomic sequences downloaded from the *Medicago truncatula* Hapmap Project website (http://www.medicagohapmap.org/). Genotyping primers were used to target and amplify *Tnt1* insertion locations that were listed in the *Medicago truncatula* Mutant Database (https://medicago-mutant.noble.org/mutant/index.php) or identified through WGS or *Tnt1*-capture sequencing. Primers were synthesized by Integrated DNA Technologies (IDT) (Coralville, IA) at a 25 nmole scale and using a standard desalting procedure. Lyophilized primers were briefly centrifuged and resuspended in molecular biology grade water to make a stock concentration of 100 µM. Prior to performing PCR genotyping, primer stocks were diluted to a 10 µM working concentration. Primers were stored at -20°C. A full list of primer sequences can be found in Appendix C. PCR was performed in 12 µl reactions using DreamTaq Green Master Mix (2X) (ThermoFisher Scientific, Waltham, MA). PCR reactions were performed in either an Eppendorf Mastercycler® Pro S (Eppendorf, Hauppauge, NY) or a Biometra TAdvanced (Analytik Jena, Beverly, MA) Thermal Cycler. Thermal cyclers were

programmed as follows: 95°C for 2 minutes, 95°C for 30 seconds (30 cycles), 50°C to 56°C for 30 seconds (30 cycles), 72°C for 1 minute (30 cycles), 72°C for 5 minutes, and stored at 4°C. The annealing temperature varied from 50°C to 56°C depending on the primer pair used. The NEB Tm Calculator (New England BioLabs Inc., Ipswich, MA) was used to determine appropriate annealing temperatures for each reaction.

2.6 Gel Electrophoresis

PCR reaction products were visualized on a 1-1.8% (w/v) agarose gel prepared in 1X TAE buffer. Ethidium bromide was added to achieve a final concentration of 0.5 µg/ml. Approximately 8-10 µl of product was added to each well, and a 100 bp DNA ladder (GoldBio, St. Louis, MO) was used to determine size of DNA fragments. A Sub-Cell® GT Electrophoresis System (Bio-Rad, Hercules, CA) or an Owl Centipede™ Wide Gel System (ThermoFisher Scientific, Waltham, MA) was used, and electrophoresis was carried out at 100 V for 1 hour. A Gel Doc XR+ System (Bio-Rad, Hercules, CA) was used to visualize DNA and capture images.

2.7 Whole Genome Sequencing and Analysis

DNA library preparation and whole genome sequencing were performed by the University of Delaware Sequencing and Genotyping Center (Newark, DE). DNA concentration was determined using a Qubit Fluorometer and approximately 1 µg of high-quality genomic DNA from a Fix+ NF18598-2 R_2 plant and a Fix- NF18598-5 R_2 plant was provided to the sequencing center. An Illumina Nextera DNA Flex library preparation kit was used, and 150 bp paired-end reads were obtained from 500 bp insert libraries by sequencing on an Illumina NextSeq system. Bioinformatic analysis was performed by Xiaolan Rao at the University of North Texas (Denton,

TX). FASTX-Toolkit was used to filter reads to improve per base sequence quality. Putative *Tnt1* insertion locations were identified using the pipeline described in Veerappan et al. (2016). Briefly, paired-end reads were aligned to the *Tnt1* sequence and the Medicago truncatula A17 reference genome using a BLAST program (Altschul, Gish, Miller, Myers, & Lipman, 1990) installed on a Linux machine. A Perl script was used to parse the BLAST output and obtain hybrid reads containing both the *Tnt1* and genomic sequences. A VELVET program (Zerbino & Birney, 2008) was used to further develop reads into contigs or nodes. Hybrid reads and contigs were aligned to the *Medicago truncatula* A17 reference genome using BLAST to identify the genomic coordinates of *Tnt1* insertions. In addition, reads were grouped together based on their genomic coordinates. Each group, or putative insertion site, was classified as either high or low confidence. Groups consisting of four or more reads or nodes were classified as high confidence and all others were classified as low confidence.

2.8 *Tnt1*-Capture Sequencing and Analysis

Sequence capture analysis was performed at the Noble Research Institute (Ardmore, OK) with a sample from the same Fix- NF18598-5 R_2 plant that was used for WGS and was performed as described in Sun et al. (2019). Briefly, DNA probes were designed using sequence from the LTR region of Tnt1. Samples of DNA from different mutant lines were pooled using barcoded Illumina libraries prior to capture-based target enrichment of *Tnt1* fragments using reagents from IDT. Following hybridization capture, *Tnt1* fragments were sequenced using Illumina NextSeq. *Tnt1* insertions were identified using a bioinformatic tool called ITIS (Identification of Transposon Insertion Sites) which is described in the previous section.

2.9 Sample Preparation for Transmission Electron Microscopy

At 15 dpi nodules were harvested into fixative solution containing 4% formaldehyde, 1% glutaraldehyde, and 80 mM PIPES buffer (pH 7.0). Nodules were cut from roots and sliced longitudinally with double-edged razor blades. Samples were fixed overnight with rotation at 4°C. The following day, samples were rinsed thoroughly with deionized water and post-fixed in 1% osmium tetroxide solution with rotation for 2-3 hours. Samples were then rinsed thoroughly with deionized water and dehydrated through graded dilution with acetone. The dilutions of acetone were 10%, 30%, 50%, 70%, 90%, 100%, and 100% for 30 minutes each with rotation at room temperature. Samples were infiltrated with Epon 812-Araldite resin through a graded dilution with acetone. The dilutions of resin were 25%, 50%, 75%, 100%, and 100% for 1 hour each with rotation at room temperature. Samples were left to infiltrate overnight at room temperature. The following day, samples were infiltrated with fresh 100% resin for approximately 6 hours. Samples were embedded in fresh 100% resin in aluminum weigh boats and heat cured at 65°C for 48 hours.

2.10 Light Microscopy

Nodule samples were cut into 0.5 µm sections with a glass knife on a Reichert Ultracut E ultramicrotome. Sections were collected onto gelatin-coated slides and stained with 1% toluidine blue and 0.1% basic fuschin in 1% sodium metaborate tetrahydrate. Slides were digitally scanned using an Aperio AT2 brightfield scanner (Leica Biosystems, Buffalo Grove, IL). Images were captured using the Aperio eSlideManager, and scale bars were added using ImageJ (https://imagej.nih.gov/ij/).

Chapter 3

RESULTS AND DISCUSSION

3.1 Phenotypic Screening of *M. truncatula Tnt1* Insertion Mutants to Identify Defects in Symbiotic Nitrogen Fixation

The screened collection of *M. truncatula Tnt1* mutant lines were reviewed for preliminary phenotypes which were consistent with defective SNF. A total of eight independent lines were ordered from the Noble Research Institute for study. Multiple batches of seeds were ordered for each line with approximately 24 plants screened per line. The number of seeds received with each order varied depending on whether the seeds were from R_0 or R_1 plants, with an average of 15 and 23, respectively. Seeds were sterilized and vernalized before being sown in cone-tainers with a 2:1 mixture of Turface (calcined clay) and coarse vermiculite to allow for easier visualization of roots. After inoculation with *Sinorhizobium meliloti* strain *Sm1021*, seedlings were analyzed for nodule and plant phenotypes at 15 dpi. Visible nodule phenotypes and general plant phenotypes were documented, and seedlings were transplanted into four-inch square pots to allow plants to flower for genetic crossings and to produce seeds for further screenings of multiple generations. Prior to inoculation, plants were grown with low concentrations of NH_4NO_3 (1 mM) to ensure sufficient germination and growth of mutant seeds. Following inoculation, plants were watered with a nitrogen-free nutrient solution.

Of the eight mutant lines that were screened (NF11276, NF15135, NF15136, NF15747, NF18242, NF18598, NF18679, and NF19670), only one line (NF18598)

displayed a Fix- nodule phenotype. A few lines (NF15136, NF15747, and NF18242) displayed Fix+/- phenotypes that consisted of smaller and more spherical nodules with some visible pink color indicating that some nitrogen fixation was occurring. Another mutant line (NF15136) had an interesting plant phenotype of fused cotyledons. These plants did not survive long enough to be propagated. The mutant line NF18598 that displayed a clear Fix- phenotype became the focus of further investigation.

Several batches of NF18598 seeds were ordered that came from an R_0 plant. Seven of the 26 plants screened in the R_1 generation displayed a Fix- phenotype (26.9%) with the remaining 19 plants (73.1%) showing a Fix+ (WT) phenotype. Due to the way the *M. truncatula* mutant plants were propagated, the R_1 phenotypic ratios are not indicative of whether the mutation is recessive or dominant. Over the 15-year course of the project, greenhouse space became limited and many of the R_0 mutant plants were grown together in order to collect seeds for research. Therefore, the seeds screened in the R_1 generation are not necessarily a random sample of R_1 progeny. For this reason, further screening of the R_2 generation was necessary to determine the

		Phenotype		Phenotype Distribution	
Mutant ID	**Generation**	**Fix-**	**Fix+/WT**	**Fix-**	**Fix+/WT**
NF18598-1	R_2	28	9	75.7%	24.3%
NF18598-2	R_2	27	5	84.4%	15.6%
NF18598-5	R_2	42	0	100.0%	0.0%

Table 2 Screening results for R_2 generation of *M. truncatula Tnt1* mutant line NF18598.

inheritance pattern of the mutation. Several attempts were made to perform genetic crosses using a method described in Veerappan et al. (2014) in order to study the segregation pattern of the phenotype. This method involves making an incision on the keel petal which allows for easy access to the anthers and stigma inside the self-fertilizing *M. truncatula* flower bud. Additional attempts were also made by making an incision on the standard petal, but neither technique produced any successful genetic crosses in NF18598.

Seeds were collected from mature mutant plants in the R_1 generation, and the same methods were used as previously described to grow, inoculate, and screen the progeny, or R_2 generation. Of the seven Fix- mutants that were identified in the R_1 forward genetic screen, three (NF18598-1, NF18598-2, and NF18598-5) produced enough R_2 progeny that could be successfully screened. Seeds from the other four mutants identified in the R_1 generation had a very low germination rate or did not nodulate (Nod-). This could have possibly been due to other unknown and deleterious mutations carried by these particular mutants caused by *Tnt1*. The phenotypic ratios observed in the R_2 generation indicated that the defective phenotype could be governed by a dominant mutation (Table 2). Two of the mutants, NF18598-1 and NF18598-2, gave rise to an R_2 generation that displayed a Fix- phenotype in 75.7% and 84.4% of their progeny, respectively. One of the mutants identified in the R_1 generation, NF18598-5, produced an R_2 generation that consisted of 100% Fix- progeny. Plants classified as Fix+ displayed a plant phenotype that was similar to the wild-type control (R108), and the mutants that were determined to be Fix- displayed more reddening on the leaves (Figure 7).

Figure 7 **Phenotypic characterization of *M. truncatula* mutant line NF18598 R_2 plants.** Plants were grown in growth medium consisting of 2:1 mixture of Turface MVP and vermiculite in the presence of 1 mM of ammonium nitrate for 6 days and subsequently inoculated with *S. meliloti Sm1021 hemA:lacZ*. Screening was performed 15 dpi. Images of whole plant phenotypes are shown with R108 and Fix+ mutants displaying greener leaves and Fix- mutants having more red leaves indicating a possible nitrogen deficiency. (A) Wild-type *M. truncatula* R108 ecotype plant used as control. (B) Fix+ plant from mutant NF18598-1 R_2 generation. (C) Fix+ plant from mutant NF18598-2 R_2 generation. (D) Fix- plant from mutant NF18598-1 R_2 generation. (E) Fix- plant from mutant NF18598-2 R_2 generation. (F) Fix- plant from mutant NF18598-5 R_2 generation.

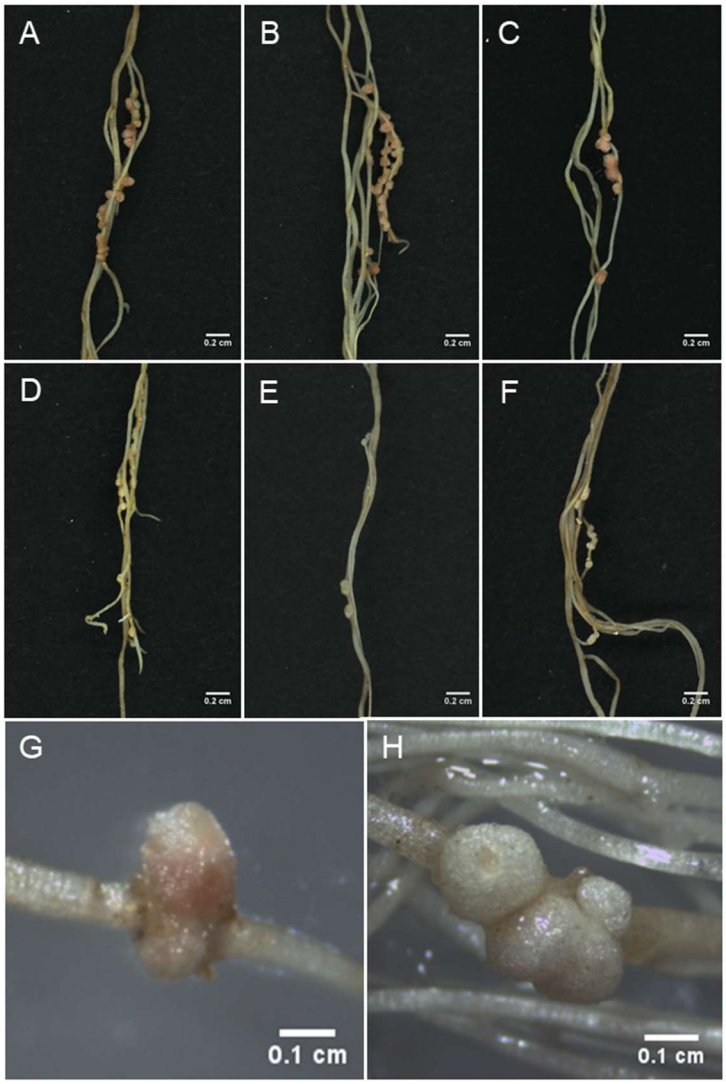

Figure 8 **Nodule phenotype of *M. truncatula* mutant line NF18598 R_2 plants.** Close-up images of visible nodule phenotypes from previous figure. R108 and fix+ plants display cylindrical shaped nodules that are pink in color indicating that rhizobia are fixing nitrogen. Fix- mutants display smaller, spherical nodules that are white in color indicating defects in symbiotic nitrogen fixation. (A) Wild-type *M. truncatula* R108 ecotype plant used as control. (B) Fix+ plant from mutant NF18598-1 R_2 generation. (C) Fix+ plant from mutant NF18598-2 R_2 generation. (D) Fix- plant from mutant NF18598-1 R_2 generation. (E) Fix- plant from mutant NF18598-2 R_2 generation. (F) Fix- plant from mutant NF18598-5 R_2 generation. (G) Image of R108 nodule obtained with dissecting scope. (H) Image of Fix- nodule from mutant line NF18598-1 obtained with dissecting scope.

Nodule phenotypes of the R_2 generation can be seen in Figure 8 and clearly show the difference in size, shape, and color that classifies Fix- mutants. The segregation ratios observed in R_2 families (3:1 and 4:0) derived from segregating R_1 plants were indicative of a dominant mutation, where NF18598-1 and NF18598-2 were presumably heterozygous and NF18598-5 homozygous for the mutant allele (chi-square test: NF18598-1: X^2=0.009, df=1, p=.92 and NF18598-2: X^2=1.5, df=1, p=.22). While this data supports the hypothesis that the defect in SNF is caused by a dominant mutation, the average sample size is low and further screenings involving a greater number of R_2 mutants would help to better support this theory.

3.2 Identification of *Tnt1* Insertions Using Different Sequencing Methods

Three separate sequencing methods were used to obtain *Tnt1* insertion locations: thermal asymmetric interlaced polymerase chain reaction (TAIL-PCR), whole genome sequencing (WGS), and *Tnt1*-caputure sequencing (SC). A total of 114 different *Tnt1* insertions were identified in NF18598 that were considered to be of high confidence. The TAIL-PCR method uses a nested PCR reaction with sets of known *Tnt1* primers and arbitrary degenerate (AD) primers each with different annealing temperatures (T_m) which allows for thermal control over product amplification (Liu & Chen, 2007). TAIL-PCR products were pooled and indexed using a two-dimensional strategy prior to being sequenced using an Illumina MiSeq platform. The *Tnt1* insertions were considered to be high confidence if they matched to one specific X and Y barcode or if they were sequenced using Sanger sequencing.

Forty insertions identified by TAIL-PCR were obtained from the *Medicago truncatula* Mutant Database (https://medicago-mutant.noble.org/mutant) maintained by the Noble Research Institute. Previous research by Veerappan et al. (2016)

demonstrated that TAIL-PCR is not as efficient at identifying *Tnt1* insertions when compared to other high-throughput methods like WGS. For that reason, both WGS and SC were also used to identify insertions in this line.

Insertions identified through WGS were considered to be high confidence if they had at least four supporting reads, and all of the insertions found by SC were considered high confidence. A total of 114 *Tnt1* insertions were identified: 41 were recovered by WGS and 70 were recovered by SC. The two high-throughput methods (WGS and SC) had the largest overlap of insertions recovered (20), and seven insertions were identified by all three methods (Figure 9). Both TAIL-PCR and WGS found a similar number of *Tnt1* insertions (40 and 41), and the SC method found the largest number of insertions (70). Additionally, 42, 12 and 30 insertions were identified by only SC, WGC and TAIL-PCR, respectively. These results are consistent with other recent research involving *M. truncatula Tnt1* mutants that has shown the average number of insertions per line to be greater than the 25 originally estimated by Tadege, Ratet, & Mysore (2005). Research by Veerappan et al. (2016) identified as many as 97 insertions in one *Tnt1* mutant studied, and other research by Sun et al. (2019) found an average of 80 insertions per line.

The *Tnt1* insertion sequences that were identified using all three methods were aligned to the *M. truncatula A17* reference genome (Mt4.0). The *M. truncatula* R108 ecotype was selected for the creation of the *Tnt1* mutants due to its ease of

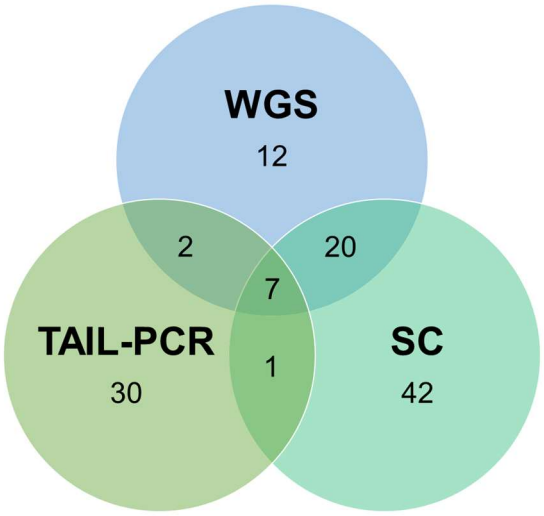

Figure 9 **Comparison of number of *Tnt1* insertions identified in mutant line NF18598 using TAIL-PCR, whole genome sequencing (WGS), and *Tnt1*-capture sequencing (SC) methods.**

transformation and ability to form symbiotic relationships with a range of well-studied strains of rhizobia bacteria (Hoffmann, Trinh, Leung, Kondorosi, & Kondorosi, 1997). Recent efforts have aimed to create a high quality R108 reference genome which would be an important resource to the legume research community; however, pseudomolecules of the genome are not yet available to the public (Moll et al., 2017). In contrast, the A17 genome has been extensively studied and annotated with an improved version released in 2014 (Mt4.0) and a variety of bioinformatic tools available (Tang et al., 2014). A more recent version of the genome (Mt5.0) was released in 2018, but the majority of online resources exist for the Mt4.0 genome making it a better reference genome for this project (Pecrix et al., 2018). Data, such as chromosome number, genetic location, and gene identification, was collected for each of the 114 insertions identified and is detailed in Appendix B.

When mapped against the A17 genome, only eight of the 114 insertions did not align to a chromosome with five of the insertions being mapped to A17 scaffolds and three matched to the R108 genome only. The other 106 *Tnt1* insertions were distributed among the eight *M. truncatula* chromosomes (Table 3).

The distribution of insertions in coding and non-coding regions was examined and it was found that 28.8% were located in exons, 21.6% were located in introns, and 49.5% were located in intergenic or non-coding regions (Table 4). Overall, the total percentage of insertions in coding regions was 50.5%. This evidence further supports previous findings that *Tnt1* preferentially targets gene-rich regions within the genome (Tadege et al., 2008). More recent studies by Sun et al. (2019) also obtained similar results with 56.1% and 43.9% of insertions being found in coding and non-coding regions, respectively. Efforts were made to verify the existence of the identified *Tnt1* insertions within NF18598 samples displaying both the Fix- and Fix+ phenotypes.

The availability of insertions found through TAIL-PCR that were listed on the *M. truncatula Tnt1* Mutant Database allowed for immediate verification while additional insertion locations were identified using WGS and SC. Custom DNA oligonucleotides were designed based upon R108 sequences obtained from the *Medicago truncatula* Hapmap Project website (http://www.medicagohapmap.org), and a combination of *Tnt1* and genome specific primers were used to verify the presence of the insertions in mutant samples. Of the 40 insertions that were listed in the *Tnt1* database, only 10 were verified through PCR. After using multiple sets of primers, the PCR was unable to amplify any fragments using a combination of different genomic and *Tnt1*-specific primers for the remaining 30 insertions, and the identified gene that

Genome Location	Insertions	Percentage
Chromosome 1	16	14.0%
Chromosome 2	11	9.6%
Chromosome 3	7	6.1%
Chromosome 4	15	13.2%
Chromosome 5	19	16.7%
Chromosome 6	3	2.6%
Chromosome 7	13	11.4%
Chromosome 8	22	19.3%
Scaffolds	5	4.4%

Table 3 **Genomic location of 114 *Tnt1* insertions identified for mutant line NF18598 in *M. truncatula* A17 reference genome.**

Genome Location	Insertions	Percentage
Exons	32	28.8%
Introns	24	21.6%
Intergenic	55	49.5%

Table 4 **Distribution of *Tnt1* insertions identified for mutant line NF18598 found in coding and non-coding regions of *M. truncatula* A17 reference genome.**

should have contained the insertion was also not disrupted for these samples. For these reasons, the unverified insertions were not considered to be the causes of the Fix- phenotype. The verified insertions were distributed between chromosomes 1, 5, 7, and 8 with another matching to scaffold 41.

The same custom primer sets were used to genotype each of the 26 mutant plant samples from the R_1 generation at each of the verified insertion locations, and logistic regression analysis was used to determine if there was any correlation between mutant phenotype and genotype under the assumption, based on results from analysis of R_2 families (described later), that the phenotype was conditioned by a dominant mutation (Table 5). The genotyping results for three of the 10 verified insertion loci were ambiguous with all 26 samples displaying a heterozygous genotype. These insertions were excluded from regression analysis and were not considered to be one of the causative insertions. A significant association (X^2=3.6, p=.0593) (was detected for the *Tnt1* insertion #51 on chromosome 5 (Table 5) which is located within the exon of Medtr5g009370, a gene that encodes a phosphatase 2C family protein. Phosphatase proteins are involved in plant signaling transduction and are involved with phosphorylation and dephosphorylation (Yang et al., 2018). Among all of the tested *Tnt1* insertions, only Medtr5g009370 was upregulated in nodules (Figure 10) .

3.3 Investigation into Medtr5g009370 as a Potential Causative Insertion

The *Tnt1* insertion located within the exon of Medtr5g009370 was further investigated to determine if it was a potential cause of the Fix- phenotype related to defective SNF. Three R_2 families derived by selfing NF18598-1, NF18598-2, and NF18598-5 were genotyped using the custom primer sets designed for the insertion located in Medtr5g009370. Consistent with a dominant mutation, NF18598-5 was

homozygous at the Medtr5g009370 locus and all of its progeny were Fix-. The remaining two parents were heterozygous and segregated for the Fix- phenotype However, logistic regression analysis showed a weak association (alpha=0.1) for only one of these parents (Table 6). Taken together, the evidence supports that either Medtr5g009370 or a linked mutation on chromosome 5 underlie the Fix- phenotype of NF18598.

Two other insertions located on the same arm of chromosome 5 were analyzed to see if the causative insertion could be in a genomic region linked to Medtr5g009370. These insertions were located in genes encoding a xyloglucan galactosyltransferase KATAMARI-like protein and a glyoxysomal malate dehydrogenase. However, these

Insertion #	Gene ID/Location	Chromosome/Scaffold	p-value
11	Intergenic	Chromosome 1	0.6908
51	Medtr5g009370	Chromosome 5	0.0593
76	Medtr7g012400	Chromosome 7	0.4255
79	Medtr7g105370	Chromosome 7	0.4765
88	Intergenic	Chromosome 8	0.6910
90	Medtr8g040410	Chromosome 8	0.4255
111	Medtr0041s0110	Scaffold 41	0.2458

Table 5 **Genotype and phenotype correlation of verified *Tnt1* insertions at their identified location within *M. truncatula* A17 reference genome in mutant line NF18598 R_1 generation (n=26).**

Mutant ID	R_2 Plants Genotyped	p-value
NF18598-1	27	0.0930
NF18598-2	22	0.3050

Table 6 Genotype and phenotype correlation of *Tnt1* insertion located in Medtr5g009370 in mutant line NF18598 R_2 generation.

6 RNA Seq Expressions		
Expression Level	Unit	Tissue
0.0	FPKM	Blade
0.430052	FPKM	Bud
64.3046	FPKM	Nodule
0.931685	FPKM	Flower
0.0	FPKM	Root
0.476	FPKM	SeedPod

Table 7 Medtr5g009370 expression in different organ tissues (RNA-Seq data obtained from the *Medicago truncatula* Genome Database).

Tnt1 insertions could not be verified via PCR, and the genes did not appear to be disrupted in the mutant samples.

Gene expression data was analyzed for all 114 insertion locations identified through the different sequencing methods, and no insertions were found in genes exclusively expressed in nodules. Medtr5g009370 was the only gene identified to be upregulated in nodules based on the data available in the *Medicago truncatula* Genome Database (Table 7). Based on this data, further genotyping experiments

involving a larger mutant population would be beneficial in determining if there is a genotype and phenotype correlation at the *Tnt1* insertion located in Medtr5g009370 .

One other possible explanation for the observation of a defective Fix- phenotype without the verification of a causative *Tnt1* insertion is that the mutation was caused by an active endogenous retroelement in the *M. truncatula* genome and not *Tnt1*. In previous research, a mutant was identified in a T-DNA collection that had defects in nodulation (Nod-). After the mutation could not be attributed to a T-DNA insertion, a copia-like retroelement called *Medicago RetroElement1-1(MERE1-1)* was identified as an insertion in a known symbiotic gene that was causing the defective phenotype. *MERE1-1* was shown to behave in a similar way as the *Tnt1* retrotransposon in that it could be activated in tissue culture and had a preference for gene-rich regions (Rakocevic et al., 2009). Therefore, *MERE1-1* could have been activated in tissue culture during the generation of *Tnt1* mutants from explants. This could have led to the insertion of *MERE1-1* in another known or unknown gene related to SNF. Another possible explanation is that a point mutation arose through somatic embryogenesis causing the defective phenotype observed in NF18598.

3.4 Analysis of Mutant Nodule Development Using Light Microscopy

Samples of mutant nodules harvested at 15 dpi were chemically fixed and embedded in resin before being thinly sectioned and placed onto slides. Samples of wild-type R108 nodules were also harvested and used as a control comparison. Images of the slides were captured, visualized, and analyzed to identify any developmental differences between NF18598 and wild-type R108 nodules. Images from control, Fix+ mutant, and Fix- mutant samples are displayed in Figure 10. Mutant samples that were classified as being a Fix+ phenotype appeared to be developmentally similar to wild-

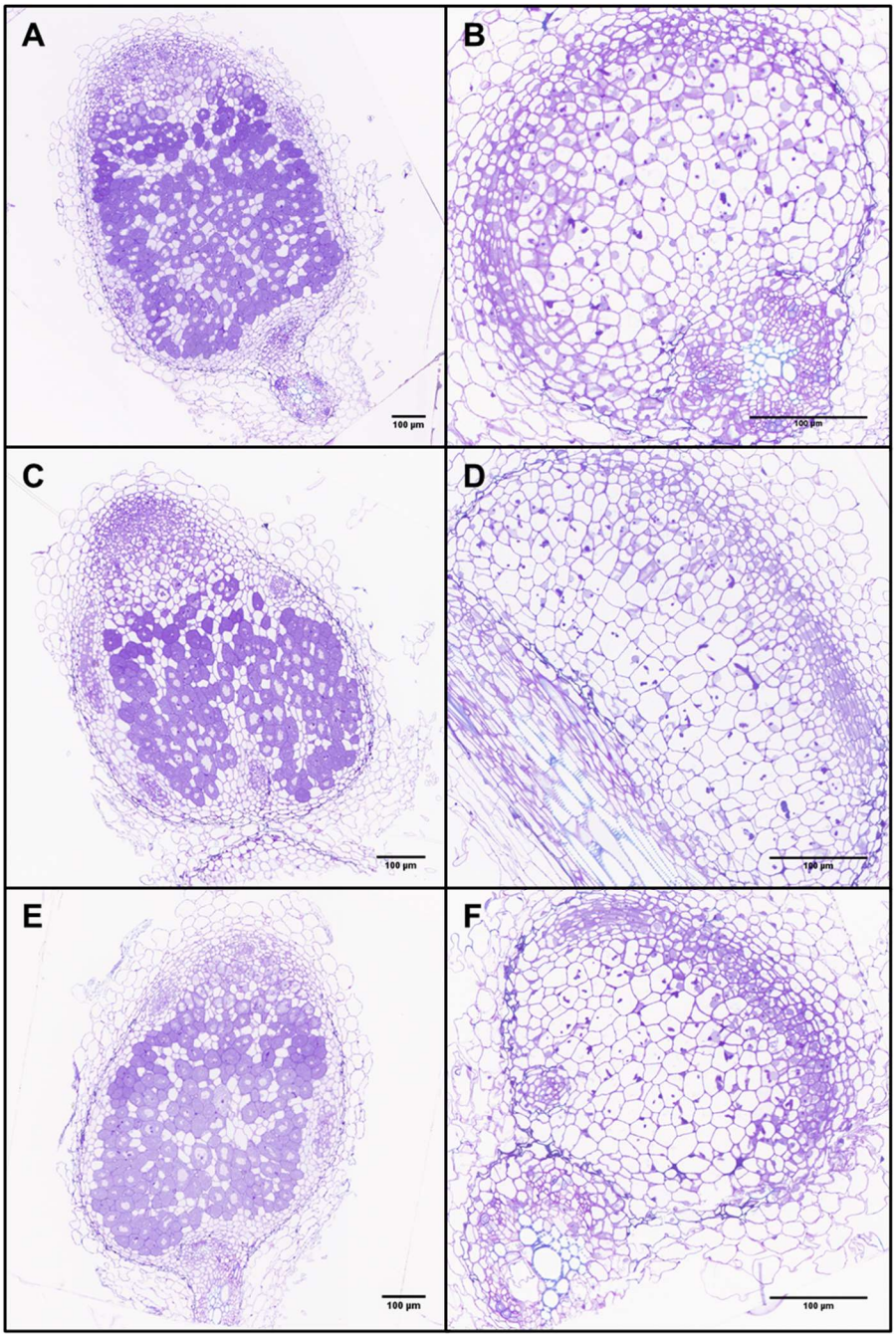

Figure 10 Light microscopy images of *M. truncatula Tnt1* mutant nodules and R108 wild-type control at 15 dpi. (A) R108 control. (B) NF18598-1 R_2 mutant with Fix- phenotype. (C) NF18598-1 R_2 mutant with Fix+ phenotype. (D) NF18598-2 R_2 mutant with defective Fix- phenotype. (E) NF18598-2 R_2 mutant with defective Fix+ phenotype. (F) NF18598-5 R_2 mutant with defective Fix- phenotype.

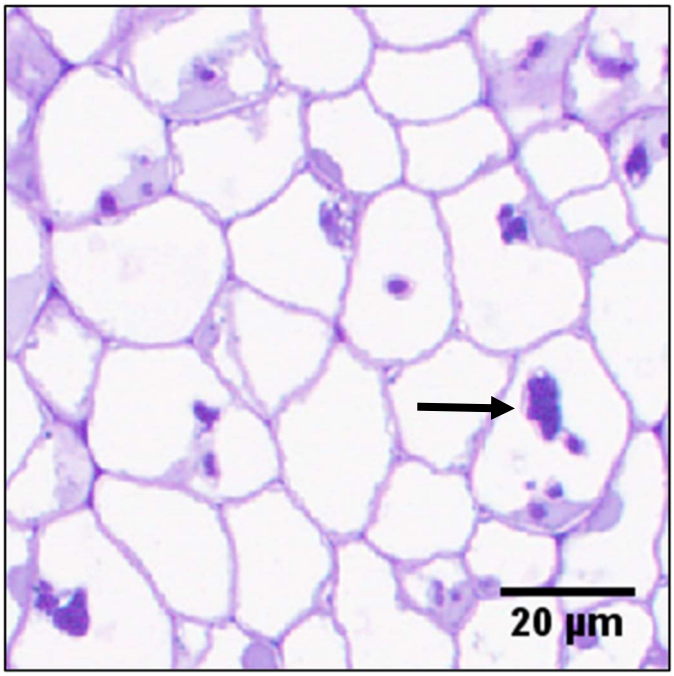

Figure 101 Light microscopy image showing the formation of infection threads in NF18598 Fix- mutant. Arrow indicates infection thread that appears to be filled with rhizobia that have not been released.

type nodules. The different nodule zones were visible in both the Fix+ mutant and wild-type samples including the meristematic, infection, and nitrogen-fixation zones. A mosaic of larger infected cells filled with bacteria among other smaller uninfected cells could be seen in Fix+ mutants which is also consistent with wild-type nodules.

The samples that were classified as being a Fix- phenotype displayed clear developmental differences when compared to wild-type nodules. While there was some cellular organization and differentiation, the nodule developmental zones could not be identified, and some samples lacked the formation of vasculature. Fix- mutant samples also displayed the formation of infection threads. These infection threads appeared to be filled with bacteria that had not been released (Figure 11). In addition,

the cells that contained infection threads did not display any evidence of degraded bacteria within their vacuoles. Further microscopy analysis including viewing images at a higher magnification must be performed to better visualize the rhizobia within infection threads.

3.5 Conclusions and Future Directions

Mutagenesis is an important tool in genetic studies aimed at better understanding plant function in order to improve crop production and agricultural practices. Mutant populations of plants allow researchers to perform both forward and reverse genetic studies that can improve our knowledge of known genes, as well as discover novel ones. The *Tnt1* retrotransposon has proven to be an effective method of mutagenesis in *M. truncatula* due to its preference for coding regions in the genome combined with a large mutant population that is available to researchers.

In this research, one mutant line (NF18598) was identified that displayed defects in SNF. The results of genotyping experiments involving both the R_1 and R_2 generations of NF18598 suggested that the defective Fix- phenotype could potentially be caused by a dominant mutation. Three different sequencing methods were used to locate *Tnt1* insertions in the genome of this mutant line. A total of 114 insertions were identified with *Tnt1*-capture sequencing finding the largest number of insertions when compared to TAIL-PCR and WGS. Logistic regression analysis of genotyping results from the R_1 generation identified one possible insertion as being a candidate for causing the Fix- phenotype. An insertion located on chromosome 5 in the exon of gene that encodes a phosphatase 2C family protein (Medtr5g009370) was investigated as being a potential causative insertion. However, further genotyping experiments involving the R_2 generation did not show a significant correlation between genotype

and phenotype at this location. Further genotyping experiments involving a larger mutant population are needed to determine if Medtr5g009370 is the causative gene responsible for the defective phenotype. Light microscopy analysis comparing both Fix- and Fix+ mutants to wild-type R108 nodules revealed clear differences. Fix+ nodules were developmentally similar to wild-type nodules. However, Fix- nodules lacked developmental zones and displayed the formation of infection threads that appeared to contain bacteria that had not been released. Further microscopy analysis must be performed to better visualize rhizobia within the infection threads of Fix- mutants. Examining samples under a higher magnification in the laboratory and further analysis using transmission electron microscopy could be helpful in better understanding the physiological defects in NF18598 related to SNF.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Ané, J. M., Zhu, H., & Frugoli, J. (2007). Recent advances in *Medicago truncatula* genomics. *International Journal of Plant Genomics*, *2008*, 1–11. https://doi.org/10.1155/2008/256597

Ardourel, M., Demont, N., Debelle, F., Maillet, F., de Billy, F., Prome, J.-C., … Truchet, G. (1994). Rhizobium meliloti Lipooligosaccharide Nodulation Factors: Different Structural Requirements for Bacterial Entry into Target Root Hair Cells and Induction of Plant Symbiotic Developmental Responses. *The Plant Cell*, *6*(10), 1357. https://doi.org/10.2307/3869974

Benedito, V. A., Torres-Jerez, I., Murray, J. D., Andriankaja, A., Allen, S., Kakar, K., … Udvardi, M. K. (2008). A gene expression atlas of the model legume *Medicago truncatula*. *The Plant Journal*, *55*, 504–513. https://doi.org/10.1111/j.1365-313X.2008.03519.x

Bruneau, A., Doyle, J. J., Hughes, C., Kenicer, G., Lewis, G., Mackinder, B., … Wojciechowski, M. F. (2013). Linked references are available on JSTOR for this article : Legume phylogeny and classification in the 21st century : Progress , prospects and lessons for other species-rich clades.

Cheng et al. (2017). Enabling reverse genetics in *Medicago truncatula* using high-throughput sequencing for *Tnt1* flanking sequence recovery. In W. Busch (Ed.), *Plant Genomics: Methods and Protocols* (pp. 25–37). Totowa, NJ: Humana Press. https://doi.org/10.1007/978-1-4939-7003-2_3

Cheng, X., Wen, J., Tadege, M., Ratet, P., & Mysore, K. S. (2011). Reverse Genetics in *Medicago truncatula* using *Tnt1* insertion mutants. In Andy Pereira (Ed.), *Plant Reverse Genetics: Methods and Protocols* (Vol. 678, pp. 179–190). Totowa, NJ: Humana Press. https://doi.org/10.1007/978-1-60761-682-5

d'Erfurth, I., Cosson, V., Eschstruth, A., Lucas, H., Kondorosi, A., & Ratet, P. (2003). Efficient transposition of the *Tnt1* tobacco retrotransposon in the model legume *Medicago truncatula*. *The Plant Journal*, *34*, 95–106.

Ferguson, B. J., Indrasumunar, A., Hayashi, S., Lin, M. H., Lin, Y. H., Reid, D. E., & Gresshoff, P. M. (2010). Molecular analysis of legume nodule development and autoregulation. *Journal of Integrative Plant Biology*, *52*(1), 61–76. https://doi.org/10.1111/j.1744-7909.2010.00899.x

Finnegan, D. J. (2012). Retrotransposons. *Current Biology*, *22*(11), R432–R437. https://doi.org/10.1201/9781351076432

Galloway, J. N., Dentener, F. J., Capone, D. G., Boyer, E. W., Howarth, R. W., Seitzinger, S. P., … Vo, C. J. (2004). *Nitrogen cycles : past , present , and future*.

Graham, P. H., & Vance, C. P. (2003). Update on Legume Utilization Legumes : Importance and Constraints to Greater Use. *Plant Physiology*, *131*(March), 872–877. https://doi.org/10.1104/pp.017004.872

Hoffmann, B., Trinh, T. H., Leung, J., Kondorosi, A., & Kondorosi, E. (1997). A new *Medicago truncatula* line with superior in vitro regeneration, transformation, and symbiotic properties isolated through cell culture selection. *Molecular Plant-Microbe Interactions*, *10*(3), 307–315. https://doi.org/10.1094/MPMI.1997.10.3.307

Kang, Y., Li, M., Sinharoy, S., & Verdier, J. (2016). A snapshot of functional genetic studies in *Medicago truncatula*. *Frontiers in Plant Science*, *7:1175*, 1–22. https://doi.org/10.3389/fpls.2016.01175

Liu, Y., & Chen, Y. (2007). Short Technical Reports High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences Short Technical Reports. https://doi.org/10.2144/000112601

Lullien, V., Barker, D. G., de Lajudie, P., & Huguet, T. (1987). Plant gene expression in effective and ineffective root nodules of alfalfa (Medicago sativa). *Plant Molecular Biology*, *9*(5), 469–478. https://doi.org/10.1007/BF00015878

Magrini, M. B., Anton, M., Chardigny, J. M., Duc, G., Duru, M., Jeuffroy, M. H., … Walrand, S. (2018). Pulses for Sustainability: Breaking Agriculture and Food Sectors Out of Lock-In. *Frontiers in Sustainable Food Systems*, *2*(October), 1–17. https://doi.org/10.3389/fsufs.2018.00064

Moll, K. M., Zhou, P., Ramaraj, T., Fajardo, D., Devitt, N. P., Sadowsky, M. J., … Mudge, J. (2017). Strategies for optimizing BioNano and Dovetail explored through a second reference quality assembly for the legume model , *Medicago truncatula*, 1–16. https://doi.org/10.1186/s12864-017-3971-4

Pankievicz, V. C. S., Irving, T. B., Maia, L. G. S., & Ané, J. (2019). Are we there yet ? The long walk towards the development of efficient symbiotic associations between nitrogen-fixing bacteria and non-leguminous crops, 1–17.

Pecrix, Y., Staton, S. E., Sallet, E., Lelandais-brière, C., Moreau, S., Carrère, S., … Benhamed, M. (2018). Whole-genome landscape of *Medicago truncatula*symbiotic genes. *Nature Plants*, *4*(December). https://doi.org/10.1038/s41477-018-0286-7

Pislariu, C. I., D. Murray, J., Wen, J., Cosson, V., Muni, R. R. D., Wang, M., … Udvardi, M. K. (2012). A *Medicago truncatula* tobacco retrotransposon insertion mutant collection with defects in nodule development and symbiotic nitrogen fixation. *Plant Physiology*, *159*, 1686–1699. https://doi.org/10.1104/pp.112.197061

Rakocevic, A., Mondy, S., Tirichine, L., Cosson, V., Brocard, L., Iantcheva, A., … Ratet, P. (2009). MERE1, a low-copy-number copia-type retroelement in *Medicago truncatula* active during tissue culture. *Plant Physiology*, *151*(3), 1250–1263. https://doi.org/10.1104/pp.109.138024

Roy, S., Liu, W., Nandety, R. S., Crook, A., Mysore, K. S., Pislariu, C. I., … Udvardi, M. K. (2020). Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation. *The Plant Cell*, *32*(1), 15–41. https://doi.org/10.1105/tpc.19.00279

Sergeevich, S. A., Alexandrovich, Z. V., Yurievna, S. O., Yurievich, B. A., & Anatolievich, T. I. (2015). Nod-Factor Signaling in Legume-Rhizobial Symbiosis. In H. El-Shemy (Ed.), *Plants for the Future*. IntechOpen. https://doi.org/10.5772/61165

Suddick, E. C., Whitney, P., Townsend, A. R., & Davidson, E. A. (2013). The role of nitrogen in climate change and the impacts of nitrogen – climate interactions in the United States : foreword to thematic issue, 1–10. https://doi.org/10.1007/s10533-012-9795-z

Sun, L., Gill, U. S., Nandety, R. S., Kwon, S., Mehta, P., Dickstein, R., … Wen, J. (2019). Genome-wide analysis of flanking sequences reveals that *Tnt1* insertion is positively correlated with gene methylation in *Medicago truncatula*. *Plant Journal*, *98*(6), 1106–1119. https://doi.org/10.1111/tpj.14291

Tadege, M., Ratet, P., & Mysore, K. S. (2005). Insertional mutagenesis: A Swiss Army knife for functional genomics of *Medicago truncatula*. *Trends in Plant Science*, *10*(5), 229–235. https://doi.org/10.1016/j.tplants.2005.03.009

Tadege, M., Wang, T. L., Wen, J., Ratet, P., & Mysore, K. S. (2009). Mutagenesis and beyond! Tools for understanding legume biology. *Plant Physiology*, *151*, 978–984. https://doi.org/10.1104/pp.109.144097

Tadege, M., Wen, J., He, J., Tu, H., Kwak, Y., Eschstruth, A., … Mysore, K. S. (2008). Large-scale insertional mutagenesis using the *Tnt1* retrotransposon in the model legume *Medicago truncatula*. *Plant Journal*, *54*, 335–347. https://doi.org/10.1111/j.1365-313X.2008.03418.x

Taiz, L., & Zeiger, E. (2002). *Plant Physiology* (Third). Sunderland, MA: Sinauer Associates, Inc.

Tang, H., Krishnakumar, V., Bidwell, S., Rosen, B., Chan, A., Zhou, S., … Town, C. D. (2014). An improved genome release (version Mt4.0) for the model legume *Medicago truncatula*. *BMC Genomics*, *15:312*, 1–14. https://doi.org/10.1186/1471-2164-15-312

Veerappan, V., Jani, M., Kadel, K., Troiani, T., Gale, R., Mayes, T., … Dickstein, R. (2016). Rapid identification of causative insertions underlying *Medicago truncatula Tnt1* mutants defective in symbiotic nitrogen fixation from a forward genetic screen by whole genome sequencing. *BMC Genomics*, *17:141*, 1–11. https://doi.org/10.1186/s12864-016-2452-5

Veerappan, V., Kadel, K., Alexis, N., Scott, A., Kryvoruchko, I., Sinharoy, S., … Dickstein, R. (2014). Keel petal incision: A simple and efficient method for genetic crossing in *Medicago truncatula*. *Plant Methods*, *10:11*, 1–10. https://doi.org/10.1186/1746-4811-10-11

Via, V. D., Zanetti, M. E., Blanco, F., Via, V. D., Zanetti, E., & Blanco, F. (2016). How legumes recognize rhizobia. *Plant Signaling & Behavior*, *11*(2), 1–6.

https://doi.org/10.1080/15592324.2015.1120396
Xiao, T. T., Schilderink, S., Moling, S., Deinum, E. E., Kondorosi, E., Franssen, H., … Bisseling, T. (2014). Fate map of *Medicago truncatula* root nodules. *Development*, *141*, 3517–3528. https://doi.org/10.1242/dev.110775
Yang, Q., Liu, K., Niu, X., Wang, Q., Wan, Y., Yang, F., … Wang, R. (2018). Genome-wide Identification of PP2C Genes and Their Expression Profiling in Response to Drought and Cold Stresses in *Medicago truncatula*. *Scientific Reports*, *8*(1), 1–14. https://doi.org/10.1038/s41598-018-29627-9
Zerbino, D. R., & Birney, E. (2008). Velvet : Algorithms for de novo short read assembly using de Bruijn graphs, 821–829. https://doi.org/10.1101/gr.074492.107

Appendix A

COPYRIGHT PERMISSIONS

Licence Terms

Licence Date:	18/02/2020
PLSclear Ref No:	32766

The Licensor

Company name:	Oxford Publishing Limited
Address:	Rights Department Great Clarendon Street Oxford OX2 6DP GB

The Licensee

Licensee Contact Name:	Kristen Holland
Licensee Address:	15 Innovation Way Newark 19711 United States

Licensed Material

title:	Plant Physiology
ISBN/ISSN:	9780878938230
publisher:	Oxford Publishing Limited

figure number & title / caption	Figure 12.11 The infection process during nodule organogenesis
page number	271
reproduction colour	Full Colour
reproduction size	Half page
positioning	inside or later pages
Are you requesting permission to reuse your own work?	No. I am NOT the author
Are you using the content as a prop?	content will NOT be used as a prop

For Use In Licensee's Publication(s)

usage type	Book, Journal, Magazine or Academic Paper...-Thesis
language	English
publication title	Identifying genes related to symbiotic nitrogen fixation in Medicago truncatula insertion mutants using a forward genetics approach
type of document	Thesis

Rights Granted

Exclusivity:	Non-Exclusive
Format:	THESIS/WHITEPAPER/CONULTATION DOCUMENT
Language:	English
Territory:	USA
Duration:	Lifetime of Licensee's edition
Maximum Circulation:	Total: 1 copies
Additional Terms:	If at some future date your thesis is published it will be necessary to re-clear this permission. Please also note that if the material to be used is acknowledged to any other source, you will need to clear permission with the rights holder and for any electronic version the © line must appear on the same page as the OUP material and the OUP material should not be included under a Creative Commons license, or any other open-access license allowing onward reuse.

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Jan 21, 2020

This Agreement between University of Delaware -- Kristen Holland ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4753671210826
License date	Jan 21, 2020
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Journal of Integrative Plant Biology
Licensed Content Title	Molecular Analysis of Legume Nodule Development and Autoregulation
Licensed Content Author	Brett J. Ferguson, Arief Indrasumunar, Satomi Hayashi, et al
Licensed Content Date	Jan 11, 2010
Licensed Content Volume	52
Licensed Content Issue	1
Licensed Content Pages	16
Type of use	Dissertation/Thesis
Requestor type	University/Academic

Format	Electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Figure 7 (A) on page 3522
Will you be translating?	No
Title of your thesis / dissertation	Identifying genes related to symbiotic nitrogen fixation in Medicago truncatula insertion mutants using a forward genetics approach
Expected completion date	Apr 2020
Expected size (number of pages)	1
Requestor Location	University of Delaware 15 Innovation Way NEWARK, DE 19711 United States Attn: University of Delaware
Publisher Tax ID	EU826007151

Copyright Clearance Center Marketplace™

The Company of Biologists Ltd. - License Terms and Conditions

Order Date	21-Jan-2020
Order license ID	1014613-1
ISSN	1477-9129
Type of Use	Republish in a thesis/dissertation
Publisher	COMPANY OF BIOLOGISTS,
Portion	Chart/graph/table/figure

LICENSED CONTENT

Publication Title	Development	Country	United Kingdom of Great Britain and Northern Ireland
Author/Editor	Company of Biologists.	Rightsholder	The Company of Biologists Ltd.
Date	01/01/1987	Publication Type	e-Journal
Language	English	URL	http://dev.biologists.org/

REQUEST DETAILS

Portion Type	Chart/graph/table/figure	Distribution	United States
Number of charts / graphs / tables / figures requested	1	Translation	Original language of publication
Format (select all that apply)	Electronic	Copies for the disabled?	No
Who will republish the content?	Academic institution	Minor editing privileges?	No
Duration of Use	Current edition and up to 15 years	Incidental promotional use?	No
Lifetime Unit Quantity	Up to 499	Currency	USD
Rights Requested	Main product		

NEW WORK DETAILS

Title	Identifying genes related to symbiotic nitrogen fixation in Medicago truncatula insertion mutants using a forward genetics approach	Institution name	University of Delaware
		Expected presentation date	2020-04-20
Instructor name	None		

ADDITIONAL DETAILS

Order reference number	N/A	The requesting person / organization to appear on the license	Kristen Holland, University of Delaware

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Figure 7 (A)	Title of the article/chapter the portion is from	Fate map of Medicago truncatula root nodules
Editor of portion(s)	n/a	Author of portion(s)	Company of Biologists.
Volume of serial or monograph	n/a	Issue, if republishing an article from a serial	N/A
Page or page range of portion	3520	Publication date of portion	1987-01-01

ELSEVIER LICENSE
TERMS AND CONDITIONS

Jan 21, 2020

This Agreement between University of Delaware -- Kristen Holland ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4753670564765
License date	Jan 21, 2020
Licensed Content Publisher	Elsevier
Licensed Content Publication	Trends in Plant Science
Licensed Content Title	Insertional mutagenesis: a Swiss Army knife for functional genomics of Medicago truncatula
Licensed Content Author	Million Tadege,Pascal Ratet,Kirankumar S. Mysore
Licensed Content Date	May 1, 2005
Licensed Content Volume	10
Licensed Content Issue	5
Licensed Content Pages	7
Start Page	229
End Page	235

Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables /illustrations	1
Format	electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title	Identifying genes related to symbiotic nitrogen fixation in Medicago truncatula insertion mutants using a forward genetics approach
Institution name	University of Delaware
Expected presentation date	Apr 2020
Portions	Figure 1 on page 233
Requestor Location	University of Delaware 15 Innovation Way NEWARK, DE 19711 United States Attn: University of Delaware
Publisher Tax ID	98-0397604

Copyright Clearance Center Marketplace™

American Society of Plant Biologists - License Terms and Conditions

Order Date	21-Jan-2020
Order license ID	1014665-1
ISSN	1532-2548
Type of Use	Republish in a thesis/dissertation
Publisher	AMERICAN SOCIETY OF PLANT BIOLOGISTS
Portion	Chart/graph/table/figure

LICENSED CONTENT

Publication Title	Plant physiology	Publication Type	e-Journal
Article Title	A Medicago truncatula tobacco retrotransposon insertion mutant collection with defects in nodule development and symbiotic nitrogen fixation.	Start Page	1686
		End Page	1699
		Issue	4
		Volume	159
Author/Editor	American Society of Plant Biologists., American Society of Plant Physiologists.	URL	http://www.jstor.org/journals/00320889.html
Date	01/01/1998		
Language	English		
Country	United States of America		
Rightsholder	American Society of Plant Biologists		

REQUEST DETAILS

Portion Type	Chart/graph/table/figure	Distribution	United States
Number of charts / graphs / tables / figures requested	1	Translation	Original language of publication
		Copies for the disabled?	No
Format (select all that apply)	Electronic	Minor editing privileges?	No
Who will republish the content?	Academic institution	Incidental promotional use?	No
Duration of Use	Life of current and all future editions	Currency	USD
Lifetime Unit Quantity	Up to 499		
Rights Requested	Main product		

NEW WORK DETAILS

Title	Identifying genes related to symbiotic nitrogen fixation in Medicago truncatula insertion mutants using a forward genetics approach	Institution name	University of Delaware
		Expected presentation date	2020-04-20
Instructor name	None		

ADDITIONAL DETAILS

Order reference number	N/A	The requesting person / organization to appear on the license	Kristen Holland, University of Delaware

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Table 1 on page 1690	Title of the article/chapter the portion is from	A Medicago truncatula tobacco retrotransposon insertion mutant collection with defects in nodule development and symbiotic nitrogen fixation.
Editor of portion(s)	Pislariu		
Volume of serial or monograph	159		
Page or page range of portion	1686-1699	Author of portion(s)	Pislariu et al.
		Issue, if republishing an article from a serial	4
		Publication date of portion	2012-08-01

Appendix B

LIST OF IDENTIFIED HIGH-CONFIDENCE *Tnt1* INSERTIONS FOR MUTANT LINE NF18598

							Identified by:		
	Insertion	**Ch. #**	**Subject Range**	**Location**	**Med ID**	**Annotation**	**TAIL-PCR**	**SC**	**WGS**
	1	1	24284237…24284653	Intergenic					x
	2	1	30689871...30690430	Intergenic			x		
	3	1	31329643...31330270	Intron	Medtr1g070515.1	F-box/LRR protein, putative	x		
	4	1	3453698…3454524	Intron	Medtr1g013360.1	histidine kinase cytokinin receptor	x	x	x
	5	1	34854777…34855711	Intergenic				x	x
	6	1	34905834...34906623	Intron	Medtr1g078050.1	nucleoporin autopeptidase, putative	x		
	7	1	36950058…36950063	Exon	Medtr1g083050.1	sucrose nonfermenting-like protein		x	
	8	1	38786494…38786499	Intergenic				x	
	9	1	46079393…46080380	Intergenic				x	x
	10	1	46138464…46139180	Exon	Medtr1g102160.1	gibberellin receptor GID1, putative		x	x
	11	1	4662544…4662545	Intergenic				x	
	12	1	5684421…5684426	Intergenic				x	
*	**13**	1	6606599…6606646	Intergenic					x
	14	1	7474911…7474916	Intergenic				x	
	15	1	7816240…7816245	Intergenic				x	
	16	1	7926225…7926230	Exon	Medtr1g024840.1	protein disulfide-isomerase LQY1		x	
	17	2	14659125…14659130	Intergenic				x	
	18	2	22409418…22410535	Exon	Medtr2g450510.1	flavin-binding monooxygenase-like protein			x
	19	2	32896240…32897115	Intergenic				x	x
	20	2	34607758…34608704	Exon	Medtr2g082195.1	LCR related	x	x	x
	21	2	35181046…35182069	Exon	Medtr2g083830.1	disease resistance protein (TIR-NBS-LRR class), putative		x	x
	22	2	35339557...35339839	Intergenic			x		
	23	2	41236727...41237262	Intron	Medtr2g096550.1	RAB GTPase-like protein A1D	x		
	24	2	42979341...42979799	Intergenic			x		
	25	2	44272576…44272581	Exon	Medtr2g102740.1	plastidic ATP/ADP-transporter-like protein		x	
	26	2	5125948…5125953	Exon	Medtr2g016630.1	cellulose synthase-like protein		x	
	27	2	649593...649767	Intergenic			x		

	28	3	32644542...32644737	Exon	Medtr3g072560.1	nodulin MtN21/EamA-like transporter family protein	x		
	29	3	34899367…34900400	Intergenic				x	x
	30	3	39081073…39081078	Intergenic				x	
	31	3	41071756...41072109	Intergenic			x		
	32	3	44284706...44285229	Intron	Medtr3g096770.1	transmembrane protein, putative	x		
	33	3	53297163...53297562	Exon	Medtr3g114120.1	GDSL-like lipase/acylhydrolase	x		
	34	3	9864129…9864134	Intergenic				x	
	35	4	18143979...18144305	Exon	Medtr4g051103.1	F0F1-type ATP synthase, subunit A	x		
	36	4	21625625…21625630	Intron	Medtr4g058740.1	lysine decarboxylase family protein		x	
	37	4	25989231…25989236	Intron	Medtr4g069140.1	phospholipid/glycerol acyltransferase family protein		x	
*	**38**	4	28712400…28713010	Intergenic				x	
*	**39**	4	31213182…31213848	Intergenic				x	
	40	4	40613046...40613333	Intron	Medtr4g098530.1	plant intracellular ras group-related LRR protein	x		
	41	4	44130536…44130541	Intergenic				x	
	42	4	45517481…45517486	Exon	Medtr4g109530.1	respiratory burst oxidase-like protein		x	
	43	4	46997968…46999002	Intergenic				x	x
	44	4	49755235…49755240	Intergenic				x	
	45	4	53866377...53866771	Intron	Medtr4g129310.1	Rho GTPase	x		
	46	4	55320666...55320883	Intergenic			x		
	47	4	55573947…55574839	Intron	Medtr4g132860.1	Serine/Threonine-kinase HT1-like protein		x	x
	48	4	6127191…6127196	Intergenic				x	
	49	4	8887791...8888012	Intron	Medtr4g025900.1	plant-specific transcription factor YABBY family protein	x		
	50	5	11902257…11902262	Exon	Medtr5g028420.1	LRR and NB-ARC domain disease resistance protein, putative		x	
	51	5	2223039…2223881	Exon	Medtr5g009370.1	phosphatase 2C family protein	x	x	x
*	**52**	5	2755377…2755975	Exon	Medtr5g010350.1	xyloglucan galactosyltransferase KATAMARI-like			x
	53	5	31417355…31417360	Intergenic				x	
	54	5	33814557...33814882	Intergenic			x		
	55	5	34307425…34307430	Intergenic				x	
	56	5	37341883…37342633	Intergenic				x	x

	57	5	37437520…37437525	Intergenic				x	
	58	5	37704058…37704977	Exon	Medtr5g087080.1	receptor-like protein		x	x
	59	5	40142515…40142520	Intron	Medtr5g091960.1	peroxisomal small heat shock protein		x	
	60	5	40310329...40310559	Exon	Medtr5g092310.1	disease resistance protein (TIR-NBS-LRR class), putative	x		
	61	5	43356427…43356432	Intergenic				x	
	62	5	4991693…4991698	Exon	Medtr5g014710.1	glyoxysomal malate dehydrogenase		x	
	63	5	5685650…5686603	Intron	Medtr5g016070.1	ribonuclease III domain protein		x	x
*	**64**	5	6145484…6146121	Intergenic					x
	65	5	732108...732387	Intergenic			x		
	66	5	8624300…8624305	Exon	Medtr5g022020.1	naringenin 3-dioxygenase (flavanone-3-hydroxylase)		x	
*	**67**	5	8714062…8714655	Intergenic					x
	68	5	885745…885750	Intergenic				x	
	69	6	29295866...29296023	Exon	Medtr6g077990.3	GRAM domain protein/ABA-responsive-like protein	x		
#	**70**	6	30138366…30139254	Intron	Medtr6g080170.1	triacylglycerol lipase SDP1			x
	71	6	5883590...5883827	Exon	Medtr6g016200.1	leucine-rich receptor-like kinase family protein	x		
	72	7	28512256…28512261	Exon	Medtr7g075885.1	CRS2-associated factor 1		x	
	73	7	30384311…30384316	Intergenic				x	
	74	7	35442629…35443511	Intron	Medtr7g090270.1	hypothetical protein		x	x
*	**75**	7	3588690…3589102	Intergenic					x
	76	7	3593254…3593983	Intron	Medtr7g012400.1	Serine/Threonine-kinase RIO1-like protein	x		x
	77	7	40870458…40870463	Exon	Medtr7g101275.1	auxin response factor 1		x	
	78	7	40878475…40878480	Exon	Medtr7g101280.1	auxin response factor		x	
	79	7	42723070...42723350	Intron	Medtr7g105370.1	B3 DNA-binding domain protein	x	x	x
	80	7	43488376…43489217	Intergenic				x	x
	81	7	43688702...43689235	Intron	Medtr7g107150.1	hypothetical protein	x		
	82	7	45628730…45629390	Intron	Medtr7g111210.1	hypothetical protein		x	x
	83	7	4623501...4623657	Intron	Medtr7g015220.1	methylesterase	x		
	84	7	47306144…47306149	Exon	Medtr7g114570.1	5'-3' exoribonuclease		x	
	85	8	11909031…11909036	Intergenic				x	

	86	8	1202270…1202883	Intron	Medtr8g007040.1	patatin-like phospholipase		x	x
#	**87**	8	1202977…1203413	Intergenic					x
	88	8	13112520...13112649	Intergenic			x		x
	89	8	15047232…15047979	Intergenic				x	x
	90	8	15047604...15047978	Exon	Medtr8g040410.1	sulfotransferase domain protein	x	x	
	91	8	16780353…16780358	Intergenic				x	
	92	8	18218830…18218835	Exon	Medtr8g446550.1	DYAD-like protein		x	
	93	8	24686508…24687227	Intergenic				x	x
	94	8	29601561…29601566	Intergenic				x	
	95	8	3056964…3056969	Exon	Medtr8g011410.1	cysteine-rich RLK (receptor-like kinase) protein		x	
	96	8	32956067...32956435	Exon	Medtr8g077510.1	GATA type zinc finger transcription factor family protein	x		
	97	8	36999930…36999974	Intron	Medtr8g089030.1	TRAF superfamily protein, putative			x
	98	8	37696447...37696735	Exon	Medtr8g090055.1	wall-associated receptor kinase galacturonan-binding protein C	x		
	99	8	40623838...40624074	Intron	Medtr8g096780.1	squamosa promoter-binding 13A-like protein	x		
	100	8	41734988…41734993	Intergenic				x	
	101	8	43887053...43887256	Intron	Medtr8g104210.2	hypothetical protein	x		
	102	8	44370099…44370749	Intergenic				x	x
	103	8	4641784…4642437	Intergenic				x	x
*	**104**	8	6512285…6513108	Intergenic					x
	105	8	6512431…6513092	Intergenic			x	x	x
	106	8	8778692…8779614	Intergenic				x	x
	107	scaffold0004*	3565917...3566359				x		
	108	scaffold0004*	4718606...4718298				x		
	109	scaffold0004*	5195472...5196012				x		
	110	scaffold0007	203072…203758	Intergenic			x	x	x
	111	scaffold0041	87370...87807	Exon	Medtr0041s0110.1	auxin response factor	x	x	x
	112	scaffold0041	95818…95823	Exon	Medtr0041s0120.1	auxin response factor, putative		x	
#	**113**	scaffold0568	324…588	Intergenic					x
	114	scaffold0568	75…70	Intergenic				x	

Note:

* indicate insertions identified through WGS in Fix+ NF18598-2 R_2 mutants only

indicate insertions identified through WGS in Fix- NF18598-5 R_2 mutants only

Appendix C

DESCRIPTION OF PRIMERS USED FOR PCR GENOTYPING

Primer Name	Primer Sequence (5' to 3')
Tnt1-F1	TCCTTGTTGGATTGGTAGCCAACTTTGTTG
Tnt1-R1	TGTAGCACCGAGATACGGTAATTAACAAGA
Medtr5g009370-F1	GGTAGGAGGAATCATGGATTGG
Medtr5g009370-F4	AAGGAGGGTAGGAGGAATCAT
Medtr5g009370-R4	GCCACCACTACAAACTCATCT
INS2-F1	ACTGCAGGTTGTATTCAATTATGG
INS2-F2	ACTGCAGGTTGTATTCAATTATGG
INS2-R1	CACAACACAAGTGACAAGTAAGG
INS3-F1	GGAGATCTCTGTCTTTGACACC
INS3-F2	GCAACCAATCGGATCAATGG
INS3-R1	AAATCGGCCTAGTTCATCCC
INS3-R2	CACCATCTTAGACAGAGAGTGC
INS4-F1	TGGATACGTTGTGATGTGTACG
INS4-F2	CTATCCAGAATCATGCGTTTAGC
INS4-R1	CAAGCTGTTGAAACCTTGTATGG
INS4-R2	GTGGTTGGAATTGGAAGAATCC
INS5-F1	CAGTATCAGTACATTCATAACCAAAGG

INS5-F2	CAGTGCAGTGCTTCAATAGC
INS5-R1	TGACTTGATGTGGGAACTAGC
INS5-R2	CATCATCAAATCCCAACCTTCC
INS6-F1	GCCAAATGCAGAAAGTGTTCC
INS6-F2	AATGATAGTCCTCCCATCTTAGC
INS6-R1	ATCAGTGTCAACGCCATTCC
INS8-F1	CAGTATCAGTACATTCATAACCAAAGG
INS8-F2	CCCTTCGTTCCTTTCTCTCC
INS8-R1	AACTAGTGCCATTACCACTCC
INS8-R2	ACCAATGTACGGTTCTCAGG
INS11-F1	ATGAGTCGTGTGTGGAGTAAC
INS11-R1	ACTGCAAGGGCTGCATATAG
INS11-R2	CTAGATTAAGTTGTTCCACCGAATACTCC
INS12-F1	GGGTTATTCACCAACAGTATAAGG
INS12-F2	TCCCATAGTCGTTGAATACTGC
INS12-R1	CCTTCCTTTCGCAAACTTTAGT
INS12-R2	CCACAATCTACTCCACGATCC
INS20-F1	AATGATATGGTTGGTCTGAGAGG
INS20-F2	GTTCTGCTGAACTATGGTGAGG
INS20-R1	TCTGCTTCTTCATCTTGCTAGG
INS20-R2	CATCAATGCTCGTTGTTTCTGG
INS22-F1	TTCACCCGGGTTTGTTAAGG
INS22-F2	CGAACTGAGCAGGGAAACTATAA
INS22-R1	TCTTATGCCAGGTTTCTAACGG

INS22-R2	ACGGTGAGATTGAACCTCTTG
INS23-F1	CCATAGCACCACGATAGTAAGC
INS23-F2	GTTGCTCAATGCTACGAATCC
INS23-R1	TCTCGACAATGTGACGAACC
INS23-R2	GGAAAGAGTTCCCTTCTTCTACG
INS24-F1	TGCAGGCAACCAAATTAAGAAC
INS24-R1	GCAAATCAAGAGTGTGCTCAAG
INS27-F1	TTTAGTCCAGCCACCTATGC
INS27-F1	AAAGTTCCAAGTTCTTGAGATGC
INS27-F2	TGATGATAGGAGTTTAGTGTGAAGG
INS27-R1	CACAAGGAAGTGGTTTGATGC
INS27-R1	TGGATTAGCAACAAATCATCACC
INS27-R2	TGGCCTTAGTTCTGTTACTGC
INS30-F1	GAGATTCACTAAATCGCCAATACC
INS30-F2	GCAATCACAGTGAGTAATGTGC
INS30-R1	CCCTTAGTTTCAGGTTATGTTTCC
INS30-R2	CGTACCAATTAGGTCCTCTAAGC
INS31-F1	GGACACTCTGCAATTTCATCC
INS31-R1	GTACTTTCGTTCGTTCAGTTGG
INS32-F1	TAAGGTGGAGGAAGCAAAGG
INS32-R1	GCCCTGTCAAGGGATAAGG
INS33-F1	CTGTACCACAACAACCTTTATCC
INS33-R1	AGACAACTACACTACGATGTTCC
INS34-F1	GGAACGGACAATTTGACTTAATGG

INS34-F2	GGGCGAAGTCAAGTAGTTAAGG
INS34-R1	TGATGGCAAAGGATAGAAGTGG
INS34-R2	TCATGTGCCTAGACAGATTAAGG
INS35-F1	TTCGCGTGTATACCGATGG
INS35-R1	ATGATATGCCACCTTGTAATTTCC
INS35-R2	GTGCATACTCAAGTGCTTATTACC
INS40-F1	AAGCACTTGTATGATTGTTCCTAAG
INS40-R1	TCTCAATGCTTCCATCTCCTAC
INS43-F1	TGGATCTTCATTGCATCATTGC
INS43-F2	TCAGGTGACAGCTAGAAATGG
INS43-R1	AGCCATTTCTAGCTGTCACC
INS43-R2	CAGCACACAAACGAACAGG
INS44-F1	GTTTCGTTACTTCCCTTGATGC
INS44-F2	GTAGGAGGTGGTAGCTTTGTATTG
INS44-R1	GTACCATGTGGTAATCCTCTCC
INS44-R2	GGAGTTCGAATTCTTACATATAACCTAGTC
INS45-F1	GTAAGTCCACGTTCATTCATACC
INS45-R1	TGAAATTCCAAGACTCTCAAACC
INS46-F1	TCTGATTCTTCCCACTATTTCCC
INS46-R1	GCTCCAACAATCATTCACAGG
INS47-F1	TTGTGTTGGTTTAGATTTCACAGG
INS47-F2	TGGGAACTTATCGTTGGATGG
INS47-R1	GTTCAGATACATCCTCAGAGTTCC
INS47-R2	ATTGAATGTAACCAGATGTGTTGG

INS49-F1	AAATGGTATGGATCACCTTTGAAC
INS49-R1	AAGCCTCTCTGTGGCTTATATC
INS54-F1	TGGAAAGAAAGCCTCCTATGG
INS54-F1	TCTAAAGTTGGTTGTCGATGGG
INS54-R1	CAGGTCATACGTCTCATTCTCC
INS54-R1	GGGACCAAGCTACTTCAATCC
INS58-F1	GCCAAGTTGAGTTGTTGAAGG
INS58-R1	CCCTTAGTTTCAGGTTATGTTTCC
INS58-R2	GATTCCAACTTTCGTGCTAATTACGTTC
INS60-F1	ACAAAGTTGGTTAAAGCAGTGG
INS60-R1	CTCATCGTCTATCCCTGAATGC
INS63-F1	TTCGTACTGTGTCTGATTGTCC
INS63-F2	TTTGGTCTTCCAGAAGTCTACC
INS63-R1	CCTCATCATCTGGATCCATTCC
INS63-R2	CAAACTCAGCAGATACATAATCCACATC
INS63-R3	TCCAAGTCCAAGCTCTATATTCC
INS63-R4	GATTTGTGCTTTAAATAGCAAGAAAGGGA
INS652-F1	TAGGCCTTTATACAGTTGTGTGG
INS652-R1	GGCACATTCTGATTTGATTAGAGC
INS65-F1	CCGTGATACATACAGAGTGTCC
INS65-R1	GGCTATTGTTGCTGCTTTCC
INS69-F1	TGTTCAAGAAGGTGGAAGAGG
INS69-R1	TGATCCAAACAGGTCTGAAGG
INS71-F1	ATCGACTTTGTGGACAACCG

INS71-F2	GCGATGTCACTTCTCCAACTA
INS71-R1	CGATTTCATGCACTATGCTCTCC
INS71-R2	TGATGTGATATAGTGAAAGCTTGTTCTTGG
INS74-F1	GATGGAGAAGGAACTTGTTATTAGG
INS74-R1	AGGAATCACCTTTGTTCTTTACC
INS74-R2	ACCTCTCTCTAATGATATGATACACTTCACT
INS76-F1	TCATTCTGCTAATACACACTGC
INS76-R1	CGGCTATTGGAATCGAACC
INS79-F1	GGGATGATGATGAGTAGGATAAGG
INS79-R1	AGGTTCCGGTATTCGTATTGG
INS81-F1	AGGGAGACAACGAGAATTTGG
INS81-R1	CCATCGGTCTTCTTGCTTAGG
INS83-F1	TGCACACACAACCAATCATAAC
INS83-R1	ATCTCTGCTCTCTTCTCAAACC
INS83-R2	GGTCGGTCTAGCTAATGTAGAATC
INS88-F2	CTCAATGTTTAAACCGACGATGG
INS88-R1	AACCTATACGTATCTGCTTCTATACC
INS88-R2	ACAGCAGGTGACAGTATGC
INS90-F1	CAACGTAAGGAACCAATTCAAGG
INS90-R1	TCTGTGTCCATCAACTCTTATCC
INS96-F1	CTGTGACTCCACTTCTACACC
INS96-R1	CTGAATCATCCATGAAACGTAGC
INS98-F1	CCCTACTGATCCTACTCATTGG
INS98-F2	GTGTACTAGTTGTGAGAGTAGTGG

INS98-R1	CCAATGAGTAGGATCAGTAGGGT
INS98-R2	ACCACCACTACTCTCACAACTAG
INS99-F1	AGATCAAAGGAGACACATCATCC
INS99-R1	TCATTCCAAGACTGCACAGG
INS99-R2	CAGTCAACACATATAGTTAGACATTCC
INS101-F1	GAACTTCTTGAAATCATCCATCCC
INS101-R1	TTCATCGGAACTCCACATCC
INS105-F1	TGATGTGGGAAGGATGAATGG
INS105-R1	GTGCTTATGTAGCTAATGTGTATTGG
INS107-F1	TTCCGGCGTCCTAATTACC
INS107-F2	GCTTACAGGCTCAATTCAAGG
INS107-R1	CCTTGAATTGAGCCTGTAAGC
INS107-R2	ATGCTCAGATGAGGAATGGG
INS108-F1	TGGGAAGGAATATGTGAAGTGG
INS108-R1	GATAACAGCCTTTGCTTGTTGG
INS109-F1	GACGGTCCAAACCATTATTGC
INS109-R1	TGAAGTGATATGGAAAGGAGTGG
INS110-F1	GGTTTCTGAGGTAGTCTCTAAGG
INS110-R1	TGATCCTCGATATCTGGATTGG
INS110-R2	GGATAGAGTGTCAATGATGTAGGGA
INS111-F1	TTGCCATCCCTTGTGTACTT
INS111-R1	CTACCAGATTTCCACCCTGAC

Appendix D

CO-SEGREGATION ANALYSIS OF NF18598 FIX- PHENOTYPE WITH *TNT1* INSERTION IN PHOSPHATASE 2C FAMILY PROTEIN GENE

Mutant	Phenotype	Genotype
NF18598-1	Fix-	Heterozygous
NF18598-2	Fix-	Heterozygous
NF18598-5	Fix-	Defective
NF18598-23	Fix-	Heterozygous
NF18598-24	Fix-	Defective
NF18598-25	Fix-	Defective
NF18598-26	Fix-	Heterozygous
NF18598-3	Fix+	Defective
NF18598-4	Fix+	WT
NF18598-6	Fix+	Heterozygous
NF18598-7	Fix+	Heterozygous
NF18598-8	Fix+	Heterozygous
NF18598-9	Fix+	Heterozygous
NF18598-10	Fix+	WT
NF18598-11	Fix+	Heterozygous
NF18598-12	Fix+	WT
NF18598-13	Fix+	Heterozygous
NF18598-14	Fix+	WT
NF18598-15	Fix+	Defective
NF18598-16	Fix+	Heterozygous
NF18598-17	Fix+	Heterozygous
NF18598-18	Fix+	Defective
NF18598-19	Fix+	Defective
NF18598-20	Fix+	Defective
NF18598-21	Fix+	Heterozygous
NF18598-22	Fix+	WT
NF18598-1-1	Fix-	Defective
NF18598-1-2	Fix-	Heterozygous
NF18598-1-3	Fix-	Heterozygous
NF18598-1-4	Fix-	Defective
NF18598-1-5	Fix-	WT

NF18598-1-6	Fix-	WT
NF18598-1-7	Fix+	WT
NF18598-1-8	Fix+	Defective
NF18598-1-9	Fix+	WT
NF18598-1-10	Fix-	Heterozygous
NF18598-1-11	Fix-	Heterozygous
NF18598-1-12	Fix-	Defective
NF18598-1-13	Fix-	Heterozygous
NF18598-1-14	Fix-	Defective
NF18598-1-15	Fix-	WT
NF18598-1-16	Fix-	Defective
NF18598-1-17	Fix-	WT
NF18598-1-18	Fix-	Heterozygous
NF18598-1-19	Fix-	Heterozygous
NF18598-1-20	Fix-	WT
NF18598-1-21	Fix-	WT
NF18598-1-22	Fix-	Heterozygous
NF18598-1-23	Fix-	WT
NF18598-1-24	Fix+	WT
NF18598-1-25	Fix+	Heterozygous
NF18598-1-26	Fix+	WT
NF18598-1-27	Fix+	WT
NF18598-2-9	Fix-	Heterozygous
NF18598-2-10	Fix-	Heterozygous
NF18598-2-11	Fix-	Heterozygous
NF18598-2-12	Fix-	Defective
NF18598-2-13	Fix-	WT
NF18598-2-14	Fix-	Defective
NF18598-2-15	Fix-	WT
NF18598-2-16	Fix-	Heterozygous
NF18598-2-17	Fix-	Heterozygous
NF18598-2-18	Fix-	WT
NF18598-2-19	Fix-	Heterozygous
NF18598-2-20	Fix-	Heterozygous
NF18598-2-21	Fix+	Heterozygous
NF18598-2-22	Fix+	Defective
NF18598-5-1	Fix-	Defective
NF18598-5-2	Fix-	Defective
NF18598-5-3	Fix-	Defective
NF18598-5-4	Fix-	Defective
NF18598-5-5	Fix-	Defective

NF18598-5-6	Fix-	Defective
NF18598-5-8	Fix-	Defective
NF18598-5-9	Fix-	Defective
NF18598-5-10	Fix-	Defective
NF18598-5-11	Fix-	Defective
NF18598-5-12	Fix-	Defective
NF18598-5-13	Fix-	Defective
NF18598-5-14	Fix-	Defective
NF18598-5-15	Fix-	Defective
NF18598-5-16	Fix-	Defective
NF18598-5-17	Fix-	Defective
NF18598-5-18	Fix-	Defective
NF18598-5-19	Fix-	Defective
NF18598-5-20	Fix-	Defective
NF18598-5-21	Fix-	Defective
NF18598-5-22	Fix-	Defective
NF18598-5-23	Fix-	Defective
NF18598-5-24	Fix-	Defective
NF18598-5-25	Fix-	Defective
NF18598-5-26	Fix-	Defective
NF18598-5-27	Fix-	Defective
NF18598-5-29	Fix-	Defective
NF18598-5-30	Fix-	Defective
NF18598-5-31	Fix-	Defective
NF18598-5-32	Fix-	Defective

Co-segregation analysis was performed on plants identified as either a Fix+ or Fix- phenotype in the R_1 and R_2 generations of mutant line NF18598. Phosphatase 2C family protein gene (Medtr5g009370) specific primers Medtr5g009370-F4 and Medtr5g009370-R4 were used to genotype the wild-type locus and Medtr5g009370-F1 and *Tnt1*-R1 primers were used to genotype the *Tnt1* insertion in the phosphatase gene. Genotyping primers can be found in Appendix C. Plants numbered NF18598-1

to NF18598-26 are in the R_1 generation. Plants NF18598-1-1 to NF18598-1-27 are selfed progeny of plant NF18598-1; plants NF18598-2-9 to plants NF18598-2-22 are selfed progeny of plant NF18598-2; and plants NF18598-5-1 to NF18598-5-32 are selfed progeny of plant NF18598-5, and are thus in the R_2 generation.