Institut für Nutzpflanzenwissenschaften und Ressourcenschutz
Lehrstuhl für Molekulare Phytomedizin

Genome-wide-association studies (GWAS) on the interaction between the cyst nematode *Heterodera schachtii* and *Arabidopsis thaliana*

Dissertation
zur
Erlangung des Grades
Doktor der Agrarwissenschaften
(Dr. agr.)

der

Landwirtschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

von

Muhammad Arslan Anwer

aus

Shakergarh, Pakistan

Bonn, 2018

Referent: Prof. Dr. Florian M.W. Grundler

Korreferent: Prof. Dr. Jens Léon

Tag der mündlichen Prüfung: 21.09.2018

Angefertigt mit Genehmigung der Landwirtschaftlichen Fakultät der Universität Bonn

“This work is dedicated to my parents”

Abstract

Plant-parasitic cyst nematodes are obligate sedentary parasites that infect the roots of a wide range of host plants. In research, the beet cyst nematode *Heterodera schachtii* infecting *Arabidopsis thaliana* is often used as a model. *H. schachtii* is sexually dimorphic and differentiation into male and female is strongly influenced by interactions with the environment. Under favorable conditions the unsexed infective juveniles (J2) tend to develop into females, whereas they rather form males under adverse conditions. After hatching from the egg, J2 enters the root tissue to find out a suitable infection site. Once a feeding site is selected, they become immobile and release specific proteinaceous effectors triggering the formation a syncytial nurse cell system that serves as the only nutrient supply throughout the nematode life cycle. To identify host loci underlying variation in susceptibility, we performed a genome-wide association study (GWAS) using different accessions of Arabidopsis that were exposed to *H. schachtii* infections. Three different susceptibility parameters were examined: average numbers of female per plant, average numbers of male per plant and their subsequent female to male sex ratio. Based on female- to -male sex ratio, we categorized all accessions into three groups: lowly susceptible, moderately susceptible and highly susceptible. GWAS revealed a novel QTL allele at chromosome four underlying potential candidates associated to sex ratio variations. Using the physical map of this locus, we selected five candidate genes from the potential QTL: *GDSL-lipase* (At4g18970), *PPR-protein* (pentatricopeptide repeat-protein; At4g18975), *AtS40-3* (senescence-associated protein; At4g18980), XTH-29 (xyloglucan endotransglucosylase 29; At4g18990) and *IWS-2* (interacts with SPT6; Atg19000). Finally, extreme bulks showing lowest and highest sex ratio were selected for detailed functional characterization of candidate genes. At first, we measured the expression of candidate genes among four extreme accessions showing lowest and highest sex ratio. This analysis revealed that out of the five, expression of two genes, *PPR* and *AtS40-3* was strongly down-regulated in lowly susceptible accessions compared with highly susceptible accessions. In comparison, expression of *GDSL-lipase*, *iWS2*, and *XTH29* was unchanged among all tested accessions. Together, these data suggest that changes in expression level of *PPR* and AtS40-3 might be correlated with variation in nematode sex ratio. *AtS40-3* and *PPR* are divergent genes that are organized head-to-head (in opposite orientations) in a non-overlapping manner. Because of their specific

regulation, we analyzed *AtS40-3* and *PPR* for sequence variations in their common promoter region. Notably, we observed a consistent deletion in the promoter in lowly susceptible accessions, whereas no such deletions were present in highly susceptible accessions. Based on these data, we hypothesized that the deletions in the promoter region might be one of the reasons for their strong down-regulation in root tissues in lowly susceptible accessions. To further investigate this hypothesis, we performed an *in silico* analysis of the commonly deleted 28 bp promoter region from susceptible line Col-0 for the occurrence of important *cis*-acting element. Our analysis revealed the presence of some prominent transcription factor binding sequences including a typical TATA box. To understand the mechanism that regulates the differential promoter activity for *AtS40-3* between lowly and highly susceptible lines, the 1500 bp promoter upstream of ATG for *AtS40-3* from the low susceptibility accession Xan-1 (carrying a deletion of 28 nt) and the susceptible accession Col-0 (without deletion) was cloned and used to drive the expression of GFP. The expression of GFP was analyzed via a transient expression system in *N. benthamiana* through real-time qPCR. Indeed, we found a significantly reduced expression of GFP driven by the promoter cloned from the Xan-1 accession (low susceptibility to nematodes) compared with that from Col-0 (high susceptibility to nematodes). We obtained the T-DNA-inserted loss of function mutant lines of *AtS40-3* and characterized these lines via nematode infection assays. Remarkably, a prominent increase in average number of males and decrease in average number of females was observed. Consequently, female-to-male sex ratio was also changed in Xan-1 compared with Col-0. Taken together, this study provides insights into genes and factors that influence the sexual variation in cyst nematode. Further exploitation of these factors may lead to development of novel resistance resources against nematode in crop plants.

Zusammenfassung

Pflanzenpararsitäre Zystennematoden sind obligat biotrophe Parasiten, die Wurzeln zahlreicher Pflanzen befallen. Zu Forschungszwecken wird oft der Rübenzystennematode Heterodera schachtii als Parasit von Arabidopsis thaliana als Modell verwendet. Diese Art weist einen deutlichen Sexualdimorphismus auf. Die Differenzierung der Geschlechter ist stark von Umweltinteraktionen beeinflusst. Unter günstigen Bedingungen entwickeln sich die undifferenzierten Infektionslarven (J2) eher zu weiblichen Tieren, während sich unter ungünstigen Bedingungen aus ihnen eher Männchen bilden. Nach dem Schlupf aus dem Ei dringen die J2 auf der Suche nach einer geeigneten Infektionsstelle in die Wurzel ein. Nachdem sie eine geeignete Stelle gefunden haben, werden sie immobil und geben Effektorproteine ab, die die Bildung eines syncytialen Nährzellensystems auslösen, das als einzige Nahrungsquelle der Nematoden dient. Um Wirtsloci zu identifizieren, die im Zusammenhang mit der Anfälligkeit gegenüber H. schachtii stehen, wurden im Rahmen dieser Arbeit so genannte Genomweite Assoziationsstudien (GWAS) auf der Basis verschiedener Arabidopsisherkünfte durchgeführt. Dabei wurden drei verschiedene Parameter untersucht: Die durchschnittliche Anzahl von Weibchen pro Pflanze, die durchschnittliche Anzahl von Männchen pro Pflanze und das daraus folgende Geschlechterverhältnis. Anhand des Geschlechterverhältnisses wurden alle Herkünfte in drei Gruppen eingeteilt: Gering anfällig, mäßig anfällig, und hoch anfällig. Die GWAS Analyse erbrachte neuartiges QTL Allel auf Chromosom vier das mit Variabilität des Geschlechterverhältnisses assoziiert war. Auf der Basis der Genkartierung dieses Lokus wurden fünf Kandidatengene identifiziert: GDSL-Lipase (At4g18970), PPR-Protein (pentatricopeptide repeat-protein; At4g18975), AtS40-3 (senescence-associated protein; At4g18980), XTH-29 (xyloglucan endotransglucosylase 29; At4g18990) und IWS-2 (interacts with SPT6; Atg19000). Außerdem wurden zwei Extremgruppen von Herkünften mit dem geringsten und dem höchsten Geschlechterverhältnis gebildet, um eine funktionelle Charakterisierung der Kandidatengene durchzuführen. Hierzu wurde die Expression der Kandidatengene in vier Extremherkünften mit hohem und niedrigem Geschlechterverhältnis untersucht. Diese Untersuchung ergab, dass die Gene PPR (At4g18975) und AtS40-3 (At4g18980) in gering anfälligen Herkünften stark abreguliert waren. Im Vergleich dazu war die Expression der Gene GDSL-Lipase (At4g18970), iWS2 (At4g19000),

and XTH29 (At4g18990) in diesen Herkünften unverändert. Diese Ergebnisse führten zu der Annahme, dass unterschiedliche Expressionslevel von PPR und AtS40-3 mit der Ausprägung des Geschlechterverhältnisses korrelieren könnten. Eine Sequenzanalyse der Promotorregion von PPR und AtS40-3 ergab, dass beide Gene Kopf and Kopf in gegensätzlicher Ausrichtung ohne Überlappung lokalisiert sind. Wegen ihrer spezifischen Regulation wurden die Promotersequenzen in den Extremherkünften untersucht. Es zeigte sich, dass die Promotorsequenzen von AtS40-3 und PPRder gering anfälligen Herkünfte Deletionen von 7-51 bp aufwiesen, während in den hoch anfälligen keine Deletionen zu finden waren. Infolge dessen verfolgten wir die Hypothese, dass die Deletionen im Zusammenhang mit der starken Abregulierung der Gene in gering anfälligen Herkünften stehen. Folglich führten wir eine in silico Analyse der zentralen 28/27 bp Promotordeletion durch, um wichtige Cis-Elemente zu identifizieren. Analyse ergab, dass die Deletion Transkriptionsfaktor-bindende Elemente, wie zum Beispiel eine TATA Box enthält. Um den Regulationsmechanismus für *AtS40-3* in gering und hoch anfälligen Linien zu untersuchen, wurde die 1500 bp umfassende Promotorregion von *AtS40-3* der gering anfälligen Linie *Xan-1* (mit einer 28 bp Deletion) und der anfälligen Linie Col-0 kloniert und zur transienten Expression von GFP in *N. benthamiana* verwendet. In Real-time qPCR-Analysen konnte eine signifikant geringere Expression mit dem Promotor von *N. benthamiana* Xan-1 gemessen werden. Infektionsversuche an einer T-DNA Mutationslinie von *AtS40-3* ergaben eine ähnliche Verschiebung der Geschlechterverhältnises bei der Entwicklung der Nematoden zu Gunsten von Männchen, wie sie bei Xan-1 beobachtet wurde. Die Ergebnisse zeigen, in welcher Weise die unterschiedliche Expression von Genen die Entwicklung von Nematoden beinflussen kann. Diese Erkenntnisse bilden eine wichtige Grundlage zu einer gezielten Züchtung von nematodenresistenten Pflanzen.

*D*ieses Ergebnis bestätigte daher die aufgestellte Hypothese, dass die unterschiedliche Ausprägung des Geschlechterverhältnisses auf die Unterschiede der Promotorregion von AtS40-3 und PPR zurückzuführen sind. Zusätzlich könnten hier noch Single-Nucleotide-Polymorphismen (SNPs) in der kodierenden Sequenz von AtS40-3 eine Rolle spielen.

Acronyms and abbreviations

PPN	Plant parasitic nematodes
ISC	Initial syncytial cell
J_{2s}	Second stage juveniles
QTL	Quantitative trait loci
GWAS	Genome wide association studies
LD	Linkage disequilbrium
AM	Association mapping
AFLP	Amplified fragment length polymorphism
RFLP	Restriction fragment length polymorphism
RAPD	Random amplified polymorphic DNA
SSR	Simple sequence repeat
GLM	Generalized mixed model
MLM	Mixed linear model
AMM	Accelerated mixed model
ANOVA	Analysis of variance
RIL	Recombinant inbred lines
IM	Interval mapping
CIM	Composit interval mapping
MS	Murashige and Skoog
DPI	Days post infection
UTR	Un-translated region
AtS40-3	Arabidopsis thaliana senescence 40-3
PCR	Polymerase chain reaction
cDNA	Complementary de-oxy-ribonucleic acid
CTAB	Cetyl tri-methyl ammonium bromide
Col-0	Columbia-0
MAF	Minor allele frequency
$K_3(Fe(CN)_6)$	Potassium Ferricyanide
$K_4(Fe(CN)_6)$	Potassium Ferrocyanide
kDa	Kilo Dalton
$MgCl_2$	Magnesium Chloride
mRNA	Messenger RNA
NaOCl	Sodium Hypochlorite
ng/µL	Nanogram/Microliter
Nu	Nucleus
qRT-PCR	Quantitative Reverse Transcription Polymerase Chain Reaction
RKN	Root Knot Nematode
RT-PCR	Reverse Transcription Polymerase Chain Reaction
S	Syncytium
SE	Standard error
St	Stylet
T-DNA	Transfer DNA

UTR	Untranslated Region
V	Vacuole
v/v	Volume/Volume
Vs	Versus
w/v	Weight/Volume
WRKY	Wrky DNA-Binding Protein
X	Xylem
µ	Micron
µM	Micro Molar

Table of Contents

Chapter 1

1.0 Introduction

Nematodes are multi-cellular organisms belonging to phylum Nematoda. They are simple invertebrate vermiforms with bilateral symmetrical and un-segmented body. Nematodes vary in size and are generally about 5-10 um in diameter and 2.5 mm long (Nyle and Weil, 2009). Their body is round in cross-section, tapering towards pointed anterior and posterior ends. They are transparent and colorless enclosed within an elastic cuticle that is considered one of the key characteristics used for identification of some species. This permeable cuticle is composed of collagens and with the passage of ions and water plays an imperative role in regulating the hydrostatic pressure in some species. They are structurally simple organisms but the term “tube within a tube”, alimentary canal and reproductive organs appeal to understand their fascinating internal structure.

Nematodes possess digestive, nervous, reproductive and excretory system; however, they lack a circulatory and respiratory systems. Their digestive system is similar to other organisms comprising of three parts; the stomodeum (mouth, pharynx and cavity between cheeks and jaws), intestine, and proctodeum which is posterior part of the alimentary canal. The mouth bears a stylet used to inject the secretions or suck liquids from the host (Barnes, 1980). Itopens into a buccal cavity that turns into pharynx.The esophagus glands are specialized producing effector proteins that are transported into the host through a stylet. The pharynx is divided into muscular bulbs containing enzymes to break down the food particles for further digestion in some species (Barnes, 1980). At the posterior end the pharynx leads to tube-like muscleless intestines equipped with valves at either side in order to control the movement of food. The intestine turns into a cloaca in females, whereas a rectum in males. However, in either situation rectum (male) or cloaca (females) leads to an opening “anus” at the posterior end of the body **(Fig. 1.1)**.

Nematodes possess a reasonably well developed nervous system that is composed of a circum-pharyngeal nerve ring made up from fournerve ganglia that run the length of the body. From these nerve ganglia, six longitudinal nerves run through the various parts of body. The sensory systems of nematodes are generally similar across the phylum. All nematodes possess bilaterally symmetrical chemosensory organs such as amphids (adjacent to mouth) and phasmids (tail). Phasmids sensilia are similar in

structure to amphids but smaller in size. Reproduction in nematodes is usually sexual, hermaphrodites are capable of self-fertilization. Females are larger than males because females carry large number of eggs. Female reproductive organs include two ovaries, two uteri, a single vegina and a genital pore, whereas male reproductive organs include testes, a seminal vesicle, vas deferens and a cloaca. The cloaca serves as a common channel for sperm and waste excretions. During copulation, male use spicules to open female genital pore to transfer sperm. Some nematode species are sexually pathenogenetic.

The secretory-excretory structure of nematodes has been a puzzle with its functional and structural variations. In general, there are two types of excretory systems in nematodes; i) glandular and ii) tubular. Parasitic nematodes are glandular type; while non-parasitic ones possess tubular excretory system.The excretory system removes nitrogenous wastes such as; ammonia, fatty acids, amino acids, peptides, amides and carbon dioxide through their body wall. It is comprised of a pore cell, a canal cell, a duct cell and a fused pair of gland cells. Presumably, it collects fluids and empties them outside via excretory duct and pore (Nelson and Riddle, 1984; Buechner et al., 1999).

The excretory gland cells are attached to the same duct and emit materials from large membrane-bound vesicles. The waste emitted material from the canal and gland cells passes through a duct placed just below the pharynx and is collected outside through the pore at ventral midline. However, these cells and their shapes are highly variable in nematodes (Chitwood and Chitwood, 1950). Some parasitic species do not possess excretory structure in general which indicates that functional and structural activities are solely based on species and might be multi-directional (Chitwood and Chitwood, 1950). However, in some species at least osmo-regulation is one function of excretory cells because the secretions are connected to the osmotic strength of the medium (Weinstein, 1952).

Nematodes are roundworms belong to one of the most abundant and diverse group of animals on earth (Decraemer and Hunt, 2006). It is challenging to identify the nematode species, however, more than 25000 species have been identified to date of which mostly is parasitic in nature (Hodda, 2011). A total of one million nematode species have been estimated (Lambshed, 1993) so far. It is difficult to find the soil without diverse community of soil invertebrates on earth and pre-requisites are to

evaluate initially these invertebrate groups during the colonization and soil formation in an agro-ecosystem (Walker and Moral 2003). In a study, four of every five multi-cellular animals on earth are nematodes (Ferris, 1998) and considered as potential bio-indicators. Soil nematodes are one of the main groups of soil invertebrates, which play a key role in the decomposition of soil organic matter, mineralization and nutrient cycling through their feeding interactions (de Goede and Bongers, 1994).

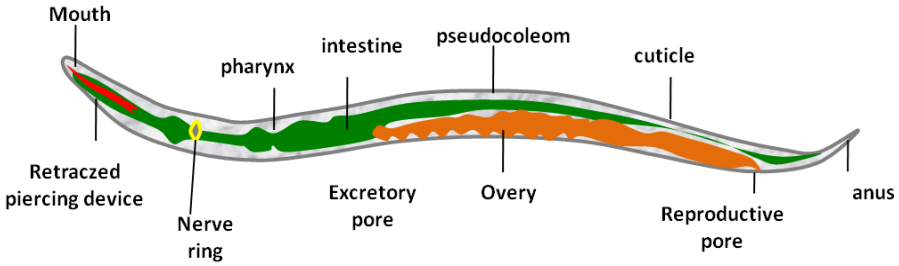


Figure 1.1: General morphology of nematode. Nematode with an opening "mouth" at the anterior end leading into the pharynx turns into intestine running across the body. The gut ends at the posterior end into an opening "anus" (Source: University of Illinois).

They are ubiquitous in a variety of climatic conditions (Hodda, 2011) and well adapted to every ecosystem such as; fresh water to marine, polar to tropics and from higher to lower elevations. Approximately, 50% nematodes is marine including littoral zones to several hundred meters depth, 15% animal parasitic, 25% free-livings and 10% are plant parasitic in nature (Maggenti, 1981). Moreover, diversity of nematodes can also be estimated based on their morphological organs, which are mostly associated with their specialized feeding habits. More recently, Yeates et al., 1993a reviewed different feeding groups of nematodes such as; bacterial feeding, plant feeding, fungal feeding, predation on soil animals and protozoa, substrate ingestion, unicellular eukaryote feeding and omnivores **(Fig. 1.2).**

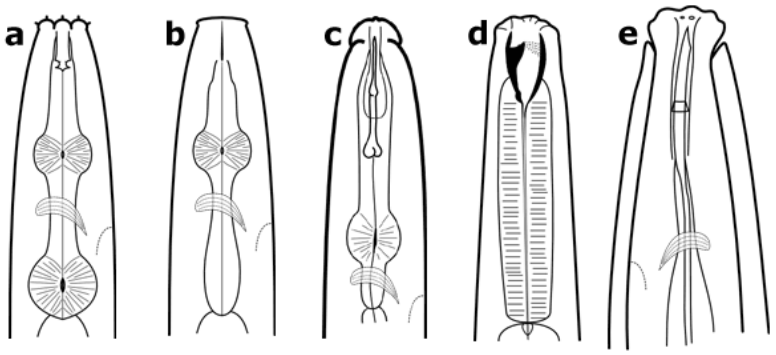


Figure 1.2: Nematode feeding groups based on the structure of their mouthparts. Bacterial feeder (a), fungal feeder (b), plant feeder (c), predator (d), and omnivore (e). (Source: University of Illinois).

1.1 Plant Parasitic Nematodes (PPNs)

Plant parasitic nematodes (PPNs) are microscopic round worms that develop an intimate parasitic relationship with host plants. They display a variety of adaptations and according to estimations; a handful cultivated soil contains thousands of nematodes (Johnson, 1982). There are more than 4100 plant parasitic species that have been described to date (Decraemer and Hunt, 2006). PPNs are recognized as a major threat to agricultural crops, and some studies have suggested an estimated loss of $US80 billion per annum (Nicol et al., 2011). Some economically important genera of PPNs include *Ditylenchus*, *Aphelenchoides*, *Heterodera*, *Globodera*, *Pratylenchus*, *Xiphinema* and *Trichodorus*. However, the majority of damage is caused by a small group of sedentary endo-parasites such as root-knot (*Meloidogyne spp*.) and cyst nematodes (*Globodera* and *Heterodera spp*.). They are biotrophic and induce complex feeding structures in their host plant's roots which supply the nematodes with a sufficient and enduring nutrients source. These genera cause 80% annual crop losses which makes them economically important throughout the world. In addition, both the cyst and root knot nematodes possess many similarities such as; modes of parasitism and host responses.

PPNs display a variety of interactions with their hosts. They possess a spearing device somewhat like a hypodermic needle "the stylet" to penetrate the host cell. Once their feeding sites are established, they transmit protein secretions leading to re-programming the parasitized host cells (Bird, 1992). These secretions play an important role in pathogenesis. One possibility is that they aid with the invasion and migration of juveniles (J2) in plant tissues (Wang et al., 1999), in addition to modifications of plant cells to re-structure the feeding site and digestion of cell contents to facilitate the nutrients availability. These stylet secretions are produced in three uni-cellular pharyngeal glands, one dorsal and two sub-ventral glands. Through cytoplasmic extensions these secretions are transported into the phyranx close to the base of stylet while sub-ventral glands blank their granules behind the pump chamber.

Extensive efforts have been made to characterize the molecular bases of nematodes invasions and manipulation of their feeding sites. However, it is accepted that pharyngeal secretions produced by pharyngeal glands account for the initiation of dramatic plant-cell modifications (Hussey, 1989a) and gene encoding secreted proteins can supposedly act as parasitism genes to endorse nematodes parasitism

spread in the host (Davis et al., 2000). Different species of PPNs are capable of feeding on various tissues such as; stems, leaves, flowers and seeds but the vast majority of them feed solely on plant roots. They exhibit a wide range of feeding habits; however, they are broadly categorized into either ecto-parasites or endo-parasites.

1.2 Types of plant parasitic nematodes (PPNs)

There are mainly two types of PPNs;

(i) Ecto-plant parasitic nematodes

(ii) Endo-plant parasitic nematodes

1.2.1 Ecto-plant parasitic nematodes

Ecto-parasitic nematodes do not penetratethe plant roots; rather they extract nutrients from root cells using their stylets. Most of the ecto-parasitic nematodes possess long stylets that facilitate their access to highly nutrient-enriched plant cells. They are either sedentary or migratory but capable to switch hosts within their lifecycle. Nevertheless, this movement makes them vulnerable to different dangers in soil such as predators etc. This group includes needle (*Longidorus*, *Paralongidorus*), dagger (*Xiphinema*), stubby root (*Paratrichodorus* and *Trichodorus*), ring (*Criconemella*) and spiral (*Heicotylenchus* and *Rotylenchus*) nematodes (McKenry et al., 1985).

1.2.2 Endo-plant parasitic nematodes

Endo-parasitic nematodes penetrate plant roots entirely and spend most of their life cycle inside the root. This type of parasites can be further divided into two types.

i) Migratory endo-plant parasitic nematodes

ii) Sedentary endo-plant parasitic nematodes

1.2.3 Migratory endo-plant parasitic nematodes

Migratory endo-parasitic nematodes enter the host and migrate through host tissues causing extensive damage. With the help of their stylet they simply suck the cytoplasm of plant cells eventually killing the cells. They do not have permanent feeding sites. However, because of their migration, they cause wounds and necrosis in plant roots. Additionally, fungi and bacterial secondary infections may also occur more often and damage the entire root system (Zunke, 1991). This group includes species of three families; *Pratylenchidae*, *Aphelenchoididae* and *Anguinidae*. These

species show remarkable variations in their life cycle and lead to fascinating contracts in host-parasite interaction.

1.2.4 Sedentary endo-parasitic nematodes

Sedentary endo-parasitic nematodes are the most damaging nematodes worldwide because of their endo-parasitic strategies. They includes; cyst nematodes (*Heterodera spp*.), root-knot nematodes (*Meloidogyne spp*.), citrus nematodes (*Tylenchulus spp*.) and reniform nematodes (*Rotylenchulus spp*.). However, the cyst and root-knot nematodes are considered the main nematodes in this group. In most of these species, the second stage juvenile (J2) is the infective stage and moves through root tissues to find a suitable site for colonization. Once a feeding site is selected they become immobile and transmit specialized proteins to dissolve the neighboring cells and establish giant feeding sources. Initially, they are completely embedded inside the roots but later the cyst nematodes protrude from the roots. The sedentary endo-parasites exhibit a complex interaction with their host and are accountable for the majority of the damage to agricultural crops. Therefore, most research is focused on cyst (*Heterodera*) and root-knot (*Meloidogyne*) nematodes that stimulate specialized feeding structures in plant roots.

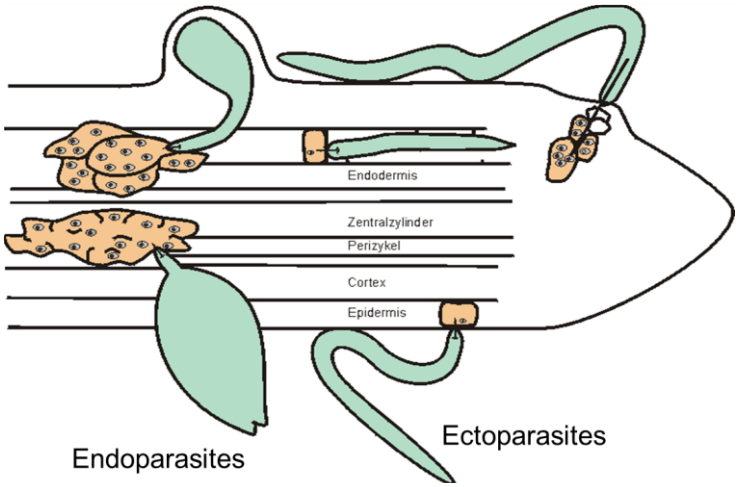


Figure 1.3: Parasitic strategies of ecto and endo-plant parasitic nematodes (Source: Florian M. W.Grundler).

1.3 Root-knot nematodes

Root-knot nematodes are a diverse group of PPNs that feed on plant roots, resulting ingalls or knot formation. They exhibit a wide host range and are the most destructive genera of various annual and perennial field crops. The genus *Meloidogyne* includes more than 90 species; however, *M. incognita*; *M. arenaria*, *M. halpa* and *M. javanica*

are the most disturbing pests worldwide (Eisenback et al., 1991). These organisms possess an embryonic stage, four juvenile stages (J1, J2, J3 and J4) and an adult stage. Eggs are deposited in gelatinous masses and survive in soil or plant root residues. The first molt takes place in the eggs, giving rise to second-stage infective juveniles (J2). Mobile juveniles enter the root close to the tips. Once inside the root, nematode move intercellularly towards the elongation zone where they make a U-turn and start moving upwards. Initially they select five to seven cells in vascular parenchyma and encourage the formation of permanent feeding sites "giant cells" (Hussey and Grundler, 1998).

The nematodes start feeding from the giant cells and become sedentary. Upon establishments of giant cells, they release secretions through stylet into the host cells leading to the hypertrophy and hyperplasia of infected cells. That is followed by swelling of surrounding cells causing the appearance of typical galls or knot-like structure on roots. During this development process, second-stage juveniles increase in size and molt to third, fourth stage juveniles (J4) and finally to adult females. However, under un-favorable conditions such as limited nutrients supply and high population densities, males can also occur in parthenogenetic species. Females becoming sexually mature secrete a gelatinous matrix from its vulva (Maggenti and Allen, 1960) that pores through the surrounding tissue (Orin and Frank, 1990). Females can deposit their eggs in this gelatinous matrix and it also plays an important role in protecting eggs from soil microbes.

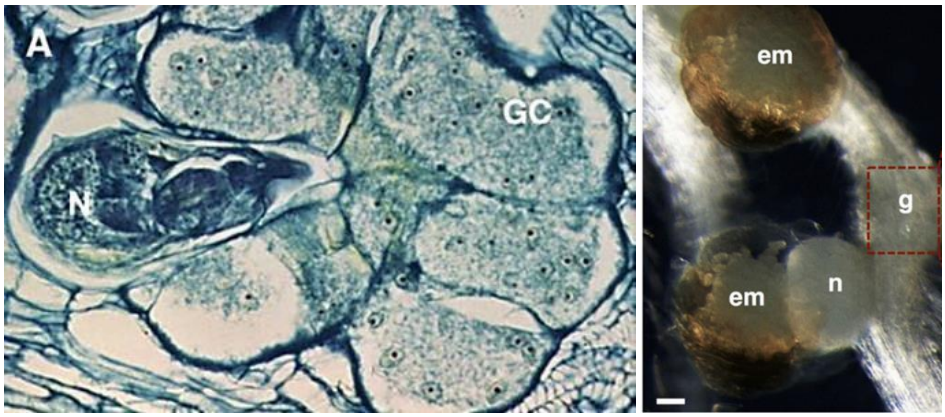


Figure 1.4: Morphological presentation of giant cells and gall. A giant cell cross-section initiated by *Meloidogyne incognita* (A) (Davis and Mitchum, 2005). A gall containing a mature female nematode and associated gelatinous matrix filled with egg masses induced in Arabidopsis roots (B). Giant cell (GC); Egg masses (em); Gall (g); Nematode (n).(Source: Rodiuc et al., 2014).

1.4 Cyst nematodes

Cyst-forming nematodes are another important group of PPNs, comprising 75 or more species. These species are organized into six genera of families *Heteroderidae*: *Globodera*, *Heterodera*, *Punctodera*, *Cactodeva*, *Afenestrata* and *Dolichodera*. Species belonging to the genera "*Heterodera* and *Globodera*" genera have been of great economic importance for several decades, however *Heterodera schachtii* on sugar beet, *Globodera rostochiensis* and *Globodera pallid* on potatoes are believed to be the major species of cyst nematodes worldwide (Zunke and Eisenback, 1998). Their life cycle starts with the hatching of second stage juveniles (infective stage), often stimulated by root metabolites of the host plant. They use their stylet and cell wall degrading enzymes produced in the sub-ventral gland cells to invade the roots of host plant and move intracellularly to vascular cylinder where they induce a feeding site. Starting with an initial syncytial cell (ISC), with the help of effectors proteins produced mainly in the dorsal glands a large number of neighboring root cells are dissolved into a syncytium which is the permanent source of nutrients. Male nematodes become mobile and leave the roots to fertilize the females, whereas most of the female body is exposed but remains attached to roots enclosing hundreds of eggs after mating. Later on the female dies and its body becomes a tanned cavity "cyst" which protects the eggs and can survive in the soil for several years (Siddiqi, 2000) until favorable conditions are available for further hatching.

Nevertheless, the structure and chemical compositions of the cyst wall of *Heterodera spp*. is of great importance because it's a protective envelope inside which hundreds of eggs can survive for extended periods of time. Doliwa (1956) indicates that the larval cuticle and cyst wall of *Heterodera spp*. are made of sclera protein without chitin. However, Ellenby (1963) found evidences of some polyphenols in the white swollen females of *H. rostochinensis;* whereas the amino acids (mainly proline, glycine and alanine) compositions suggested that collagens are predominating in the cell wall and cuticle of *Heterodera spp*. (Clark et al., 1968). Cyst nematodes are considered as pests of temperate regions, but some are also present in tropical and sub-tropical regions (Evans and Rowe, 1998).They cause substantial losses to various important crops including rice, soybeans, potatoes and cereals. However, as many environmental, cultural and biological factors are involved in causing yield losses to

crops, it is impossible to evaluate the total economic losses caused by cyst nematodes alone.

In favorable conditions, several cyst nematode species in tropical areas (Perry and Gaur, 1996) for example, *H. glycines* produce eggs in egg sacs that hatch readily in water, whereas in unfavorable conditions, greater hatch stimulation is required. Temperature and soil conditions play an important role in hatching of cyst nematodes. Rough and coarse-textured soil stimulates hatching, providing a favorable environment for aeration and migration of nematodes, whereas water-logging and drought inhibit hatching. Present study, focuses more on cyst nematode *H. schachtii* as a model organism for the interaction with Arabidopsis to investigate their susceptibility to variants which is described in detail in the next paragraph.

1.4.1 Sugar beet cyst nematode (*H. schachtii*)

Sugar beet (*Beta vulgaris L.*) is the largest crop used for sugar production as about 40% of the world sugar is extracted from this crop. Sugar beet is grown in 42 different countries all over the world (Asadi, 2007). Various biotic and a-biotic stresses challenge the growth and yield of sugar beet; however, plant-parasitic nematodes are one of the main parasitic factors that influence sugar beet plantsand result in severe yield losses. One of these nematodes, the sugar beet cyst nematode, *H. schachtii* (Schmidt et al., 1871) is the most important and disturbing type (Greco et al., 1982). The yield losses caused by this nematode may exceed to 80% worldwide (Curto, 2008). Thes biotrophs that plague plant roots can cause severe yield losses to other agronomic crops as well.

As far as their life cycle is concerned, second-stage infective juveniles of *H. schachtii* enter the roots by piercing epidermal cells and moving intracellularly to the vascular cylinder where they select a suitable initial syncytial cell (ISC) to initiate the feeding site (Golinowski et al., 1996; Sobczak et al., 1997). ISC selection may take several hours; however, once the ISC is established, the nematodes become sedentary. The ISC expands through local dissolution of cell walls leading to a multinucleate, hypertrophied syncytial nurse cell system formation. Syncytium is the only source of nutrients for the sedentary nematodes. A cocktail of nematode-origin secretions that are released into the infected host tissues facilitate the processes of nematode migration, ISC establishment and syncytium development. In the third juvenile stage,

females become sedentary, their bodies swelling to lemon shape, and the swollen portion of the adult female is exposed to the root surface while their head stays embedded in the root.The adult males leave the roots and follow the females for fertilization. After fertilization the females transform into a body cavity "cyst" containing hundreds of eggs. Cysts can protect the eggs for several years until favorable conditions are available for hatching. Thus, syncytia associated with female nematodes stay active, providing life-sustaining nutrients, whereas syncytia associated with males become necrotic after the adult males leave the roots (Wyss 1997; Decraemer and Hunt 2006).

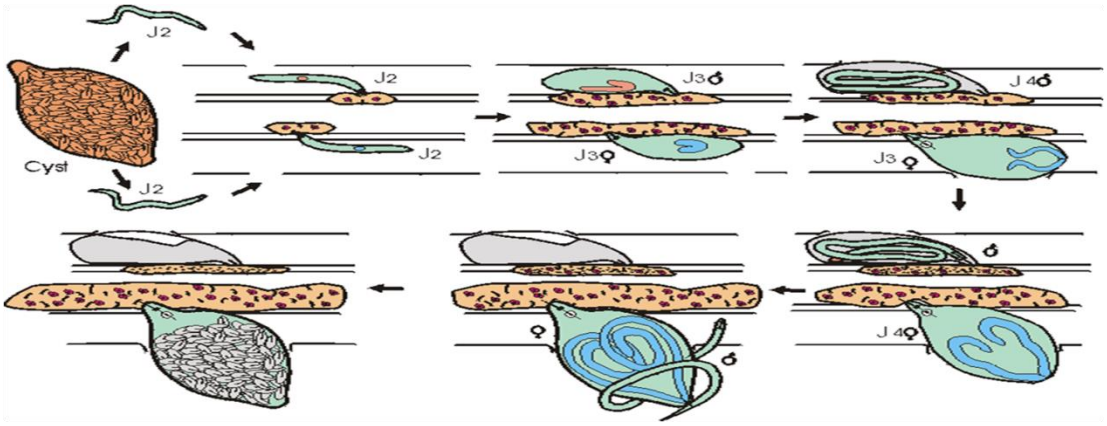


Figure 1.5: Life cycle of *H. schachtii*. After hatching from the eggs, infective juveniles move towards the roots, enter the root cortex and continue to move through cells until they find the permanent feeding site. Once they find the initial feeding cell, they become sedentary and molts into a third- and a fourth-stage juvenile to become an adult male and female. Females transport their dorsal secretions to dissolve the neighboring cells to form a multinucleate permanent feeding site "syncytium". The females become enlarged and come out with head attached to the root. Adult males moving around fertilize the female and after fertilization the female becomes a brownish cavity "cyst" containing 200-300 eggs (Source: Florian M. W. Grundler).

As far as Arabidopsis-nematodes interactions are concerned, *Arabidopsis thaliana* has been shown to be a good host for many plant parasitic nematodes (Sijmons et al., 1992) and several special properties such as translucent roots growing on artificial media have made it possible to study the behavior of cyst and root-knot nematodes inside the roots (Wyss and Grundler, 1992; Sijmons et al., 1994). *H. schachtii* can complete its life cycle within 6 weeks on Arabidopsis roots (Sijmons et al., 1991) and now microarray analysis has accelerated the transcriptomic studies of syncytia induced in Arabidopsis roots. A series of studies have been conducted to understand the physiological changes and regulation of gene expressions within syncytia induced

by cyst nematodes. Puthpffet al., (2003) examined *H. schachtii* infected roots at 3 days post infection (dpi) and identified 128 genes covering 30% of the genome with altered mRNA expressions after nematode infection (Puthoff et al., 2003). In another microarray study 1400 protein transport coding genes were reported for *M. incognita* infected roots of Arabidopsis (Hammes et al., 2005). Transcriptome analysis of syncytia induced by *H. schachtii* in soyabean and Arabidopsis roots at 5 and 15 days post infection (dpi) showed differential expressions of genes compared with uninfected roots (Szakasits et al., 2009) and with gene ontology (GO) analysis of up and down-regulated genes it was concluded that genes related to high metabolic activities had preferentially increased expressions. In addition, the enrichment of sugars and amino acids indicates that amino acids and carbohydrates play an important role in cyst nematodes development (Siddique and Grundler, 2015).

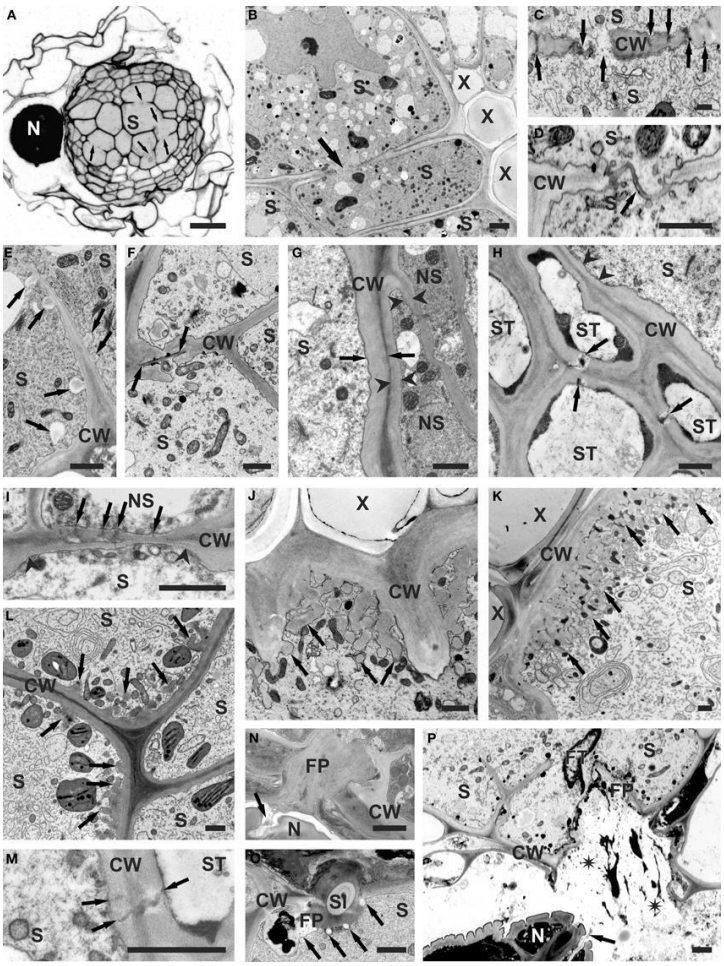


Figure 1.6: Cross-sections in *H.schachtii* induced syncytia in Arabidopsis roots. Syncytium structure with arrows indicates cell wall openings (A). Ultra-structure of

syncytium with arrow indicates cell wall dissolution (B). Widening of plasmodesmata (arrows) leading to openings of cell wall (C). Cell wall opening without involvement of plasmodesmata and arrow indicates middle lamella covered with plasmalemma (D). Paramural bodies formed at the inner side of cell wall (E). Casparian stripe covered with newly deposited cell wallin syncytium induced in the endodermis (F).Thickness comparison between syncytial and non-syncytial cell wall (G). Thin part of outer syncytial cell wall with arrowheads facing sieve tube and Plasmodesmata between sieve tubes is shown with arrows (H). Arrows pointing Plasmodesmata at thin partof outer syncytial cell wall (between arrow heads) facing non-syncytial parenchymatous cell (I). Single cell wall ingrowths (J) and well developed system of cell wall ingrowths (arrows) at syncytial wall facing vessels (K). Immature cell wall ingrowths on wallbetween syncytial elements pointed with arrows (L). Plasmodesmata between sieve tube and syncytial element (M). Feeding plug in syncytial cell wall and secretions from nematode amphids is pointed with arrow (N). Feeding plug with inserted cross-sectioned nematode stylet and callosedepositions (white spots) are indicated with arrows (O). Broken feeding plug in syncytial walland cytoplasm is marked with asterisks (P). Light microscopy (A) and transmission electron microscopy micrograms (B–P) of syncytia at 2dpi (E, O), 5dpi (A, B, D, F, N, P), 10dpi (G, H, I, J, M), and 13dpi (C, K, L) days post inoculation. Cell wall (CW); feeding tube (FT); feeding plug (FP); nematode (N); syncytium (S); non-syncytial cell (NS); stylet (Sl); sieve tube (ST); xylem vessel (X). Bars = 20 μm (A) and 1 μm (B–P) (Source: Bohlmann and Sobczak., 2014).

1.5 Sex determination in nematodes

Cyst nematodes are sexually dimorphic and variations in the sex expression of nematodes are associated with several interacting partners like genetic mechanism of sex determination, the environmental impact on sex differentiation, their developmental rate and the survival rate of both sexes. To understand the mechanism of sex ratio variations in nematodes are of vital importance in a population dynamic as it determines not only the population increase from generation to generation but also damage proportion which is most often directly relative to female nematodes. So far, it appears that males and females in bi-sexual species occur in an equal proportion and the chromosomal mechanism of sex determination is of XX♀-XO♂ type (Triantaphyllou, 1973). However, no sex chromosomes have been known among plant-parasitic nematodes and possibly XX♀-XY♂ prevails with sex chromosomes.In parthenogenetic species only females are produced or very small proportion of male is developed by an unknown mechanism (Triantaphyllou, 1964), whereas among hermaphroditic species variable percentages of male and females are produced.

Although several studies have been conducted to investigate whether sex variations among nematodes are epigenetic or solely environmental, nevetherrheless, it still it is pre-mature and unclear. To date, most studies believe that environment plays a

decisive role in determining the sex expressions in nematodes. First indications that availability of healthy nutrition to the developing nematodes may influence its sex of nematodes were observed on *Mermithids*. Cobb et al. (1927) observed in small infestations, only a few *Mermithids* were recovered from their host insects and most of them were generally females. Nevertheless, in moderate infestations, male and female nematodes were almost equal in numbers, whereas in heavy infestation with relatively large number of nematodes, the sex ratio shifted to favor males. Christie (1929) confirmed these observations and demonstrated that on feeding 4-5 *Mermithid subnigrescens* eggs to insect host, 92% of the nematodes recovered were females but when 20-30 eggs were fed to insects, 85% recovered nematodes were males. Therefore it was suggested that crowded conditions or high infestation limits the space and food which changes the physiology and biochemistry of the insect host thus effecting the sex differentiation of nematodes. In addition, Tyler (1933) illustrated that adverse nutritional conditions during the early developmental stagesmight be important to change the physiological balance of certain individuals. Linford (1941) favored these arguments that unfavorable conditions cause an increase in male proportion but more recent studies have partially clarified the mechanism of sex differentiation among *Heterodera* and *Meloidogyne* genera.

The occurrence of unbalanced sex ratios of root-knot nematodes in natural population and in green houses is quite common and it is generally believed that the sex expression of *Meloidogyne* is influenced by the environment. Experimental evidence that environment impacts its sex ratio is explained by the anatomical study of embryogenesis (Triantaphyllou and Hirschmann, 1960) and it was concluded that under favorable conditions *Meloidogyne* larvae develop normally into pear-shaped females with two testes, whereas the pattern of larval development changes under adverse conditions and become vermiform adult males with one testis. These preliminary observations convinced the scientists that “environment is a sex determining factor”. Similarly, unbalanced nematode sex ratios were also observed in several members of the *Heteroderidae* family. Molz (1920) observed a considerable variation in male numbers of sugar beet cyst nematode (*H.schachtii*), which is directly, associated with unfavorable growth conditions of the host plant. This was the start of a plant pathology era through the investigation of how the environment can influence that sex differentiation of *Heterodera* larvae. Studies to investigate the

mechanism, by which environmental factors influence the sex ratio of nematodes, have been limited to only a few species of *Heterodera*. Ellenby (1954) reported that in *H. rostochiensis* more males were developed on lateral roots (less nutrition available to the developing larvae on lateral roots) than on primary roots of potato host plant and concluded that "sex differentiation is under environmental influence".In addition, he also suggested that differential death rate of both the sexes could explain his data. DenOuden (1960) reviewed these findings and conducted some experiments on potato seedlings. He inoculated potato seedlings with single larvae of *H. rostochiensis* and absolutely obtained more female numbers than males suggesting that nutritional status could influence the sex expression of nematodes rather than their differential death rates.

Similarly, Ross and Trudgill (1969) showed that under low infection densities, all the larvae find good infection sites with sufficient space to develop giant cells and under such balanced conditions most of the larvae become females, whereas with high infection densities, all the larvae got settled close to each other with limited space to develop giant cells and under such unfavorable conditions most of the larvae become males. Uneven sex ratios have been observed often in *H. schachtii* with conflicting results on their incidence. Apfel and Kämpfe (1957) found that variation in sex ratio of *H. schachtii* is influenced by the infection density and host plant and these findings were tailed by Kämpfe and Kerstan (1964) who confirmed the impact of numerous environmental factors such as lack of light, nitrogen deficiency and unfavorable water supply on the sex ratio of nematodes.

Ross (1964) proposed that fluxes in temperature can influence the male to female sex ratio in *H. glycines*. According to him, smaller numbers of male populations were found in *H. glycines* cultures at 24C°, whereas this ratio was increased to 1:1 at 28C°and this relative increase in the male to female ratio was ascribed to the influence of higher temperature. Further increase in male to female ratio at 30C° was credited to the death of many female larvae at these high temperatures. In addition, Steel (1971) suggested that resistance of the plant roots to deep root penetration is associated with the increased percentages of *H. schachtii* males. He described that males can develop from larvae with their heads embedded into plant roots, whereas females can develop from the larvae that deeply penetrates the roots with their heads planted into the vascular parenchyma of the roots. Therefore, it was concluded that

factors resisting the access of larvae to vascular cylinder of roots including the tough epidermis in old roots indirectly resist the development of female larvae and will favor male to female ratio.

Grundler (1991) hypothesized that an exogenous application of sucrose could enhance the nutritional status of nematodes feeding site "syncytium" which influence the development of male and female juveniles. Under sterile conditions he applied different concentrations of sucrose containing solutions to *H. schachtii* juveniles and confirmed that under healthy nutritional conditions more females were developed, whereas under adverse or unhealthy nutritional conditions most of the juveniles remained undeveloped. These finding were a great addition to believe that changing host nutritional status in turn leads to alter the sex expression of some cyst nematodes and thus environment controls the sex differentiation.

More recently, Anjam et al. (2017) (unpublished) proposed that male and female associated syncytia in *Arabidopsis thaliana* host plant could be a source for sex expression regulators that may contribute in sex determination of cyst nematode *H. schachtii*. They performed a transcriptome analysis of male associated syncytia (MAS) and female associated syncytia (FAS) at early feeding stages and selected some potential candidates including *lptg6*, *lng1* and *irx12* which were differentially up-regulated in female associated syncytia (FAS). In addition, T-DNA mutant lines were subjected to infection assays and a remarkable decrease in female numbers associated with syncytia size was observed. Finally, it was concluded that these observations support the role of environment on the sex expression of cyst nematode *H. schachtii*.

Although early work favored the environmental influence on the sex differentiation of nematode larvae, still these conflicting results are unable to form a definitive statement on the underlying cause of unbalanced sex ratio in genus *Heterodera*. Considering the previous findings, in the present study with the help of genome-wide association mapping (GWAS), an effort has been made to figure out the genetic variants activating the variations in the average number of male and female nematodes per plant in Arabidopsis. In addition, female to male sex ratio traits were also brought under investigation and it is assumed that underlying allelic variations in Arabidopsis genome in response to these cyst nematodes might be associated with altering their sex ratios.

1.6 Arabidopsis a model organism for genome analysis

Arabidopsis thaliana commonly known as thale cress or mouse-ear-cress is a small plant belonging to the mustard family (Brassicaceae) with broad natural distribution throughout Europe, North America and Asia. It has become the organism of choice for a wide range of molecular studies in plant biology (Meyerowitz and Somerville, 1994). The significant advances in Arabidopsis research reflects that this simple angiosperm can serve not only for plant biology but also addresses the fundamental questions of biological structure and functions common to all eukaryotes. Important features of Arabidopsis include small size, small genome size, short generation time and prolific seed production through self-pollination. The entire life cycle is completed in 5 to 6 weeks. The 135Mbp genome is organized into five haploid chromosomes containing 25,000 genes.

Prior to the 80s, when dramatic advances in molecular genetics fundamentally changed the landscape of biology, it was not obvious that plant science would play a central role in the approaching revolution. Plant genomes were large and complex, life cycles were long, and most of the favored genetic models at the time were difficult to transform. Even the future of plant genetics as a discipline was uncertain. In due course, manipulation of resources including chemical and insertional mutagenesis, efficient crosses and Agro-bacterium mediated plant transformations, mutations of diverse phenotypes and a mapping of a variety of chromosomes of mutant genes in Arabidopsis (Koncz et al., 1992) enabled plant biologists to remain at the fore front of modern biology.

On the other hnad, Arabidopsis, displays a wide range of genetic and trait variations. Availability of hundreds of accessions and genomic resources has increased the potential of *A. thaliana* for studies of natural genetic variations in recent years. Initially, natural variation controlled by multi-genesand influenced by environmental factors was resistant to molecular dissection. To examine this variation required a significant genomic regional analysis to be conducted as a first step, which has proved very efficient by using recombinant inbred lines (RILs) in Arabidopsis. Then, identification of associated genes and single nucleotide polymorphism (SNP) using particular genomic region was the major challenge and is now feasible by combining unique high-throughput genetics and functional genomic strategies. Today, the

analysis of Arabidopsis genetic variation is providing unique knowledge from functional, ecological and evolutionary perspectives.

A large amount of research and information on *A. thaliana* is compiled in online sources such as; The Arabidopsis Information Resource (TAIR) (http://www.arabidopsis.org), Arabidopsis Biological Resource Centre (ABRC) (http://abrc.osu.edu/) and Nottingham (European) Arabidopsis Stock Centre (http://arabidopsis.info/). The genome sequence and nucleotide polymorphism data has greatly accelerated the rate of gene discovery and profitably extended the reach of genetic analysis into many research areas including cell biology and metabolism. Here, another attempt has been made to investigate the quantitative trait loci (QTL) associated with Arabidopsis susceptibility against cyst nematode. For these studies a natural population of 220 accessions was selected all over the world. The nematode infection assays for all 220 accessions were conducted but due to the un-availability of sequences of all variants, only 148 accessions were further processed for nucleotide polymorphisms analysis using Arabidopsis 1001 genome databases.

1.7 Genetic recombination and linkage analysis

In sexual reproduction, homologous chromosomes inherited from each of one's parents exchange their genetic information, resulting in the formation of recombinant chromosomes. The genetic recombination carries a combination of traits that are different from those of each of the parents. In crossing over, these new recombinations occur at any position along chromosomes; however, the frequency of these genetic recombinations between two positions solely depends on the distance apart. Therefore, the genes physically close to each other on the chromosome are difficult to segregate during crossing over and are thus considered linked. In other words, there are fewer chances of recombination between linked genes and more likely to be inherited together. Moreover, the frequency of genetic recombination is directly proportional to the distance between two loci. The greater the distance between two loci, the higher will be the frequency of recombination, whereas the lower the distance between loci, the lower will be the frequency of recombination. Nevertheless, for geneticists, the segregation of closely related loci after crossing over has been quite valuable to discern their distance on chromosomes. For this, genetic linkage analysis is a useful approach to assess the physical position of causal variants and subsequent associations with a particular trait.

1.7.1 Gene mapping

This is a method to identify the locus of a gene and distance between them. However, with the advent of molecular markers, it has become faster and more efficient to track the movement of linked loci on chromosomes. These markers describe the relationship between an inherited trait and its genetic cause, which might be due to polymorphism in nucleotides like insertion, deletion or substitutions of nucleotides. The residence of genes associated with particular traits has been evaluated using several genome mapping strategies.

Since decades the dissection of genetic architecture conferring phenotypic traits has been an ongoing challenge for geneticists. Two complementary approaches have primarily been used for mapping a genome; i) QTL mapping (using experimental population or bi-parental population) and ii) association mapping or LD-mapping (using natural population of diverse lines or germplasm collection). Both methods identify the quantitative trait loci (QTLs) with marker-trait associations. However, the fundamental difference between them is the mapped populations which directly determine mapping resolution and power. Generally, association mapping is linked to two main categories; i) candidate-gene association mapping, which deals with the polymorphism in a candidate gene already supposed to influence the trait variations and ii) genome-wide association mapping (GWAS), which scans the entire genome to discover the association signals for specific traits (Risch andMerikangas, 1996).

1.7.2 Significance of molecular markers in plants

Genetical progressin crop species is always required to improve their economic importance. Since decades increase in crop yield, resistance to pathogens and better adoptability to different environmental conditions has always been the main focus of geneticists and breeders. Genetic manipulations in plants to obtain the desired characteristics with genetic engineering, genomic and marker-assisted selection have substantially influenced plant breeding. Individuals within a population exhibit heritable genomic variations caused by mutations, deletions, insertions and translocations.

Genetic markers are biological features determined by allelic forms of genes with a known position on a locus and can be transmitted from one generation to another. Generally, they represent as a reference points or signs in close proximity of target genes. These markers do not have direct effect on phenotypic traits because they are

located near or associated with genes influencing the trait (Rabiei, 2007). They can be used as experimental tags to track individuals, genes, nucleus or chromosomes (Botstein et al., 1980).

With advances in genome sequencing technologies and advent of genetic markers with known genomic locations, it has become feasible to map whole genome and figure out the allelic variants with single-base resolution. Genetic markers are classified into two categories; i) Classical markers such as; morphological (phenotypic traits such as; flower color, seed shape, growth habits or pigmentation biochemical (isozymes, including allelic variants of enzymes) and cytological markers, ii) DNA markers (variation in DNA) such as; PCR (polymerase chain reactions) and DNA sequencing (Collard et al., 2005) including Amplified Fragment Length Polymorphism (AFLP), Random Amplified Polymorphic DNA (RAPD), Restriction Fragment Length Polymorphism (RFLP) and Single Nucleotide Polymorphism (SNP).

In the eighties, because of their co-dominance and reproducibility, RFLPs markers were considered the most widely used in plant molecular genetics (Lander and Botstein, 1986) however; they were considered expensive and time-consuming rendering them obsolete. Nevertheless, in the beginning of nineties, the detections of polymorphisms with PCR technology replaced the low-throughput RFLP markers and the applications of RAPD, AFLP, and SSR markers became obvious in various plant systems. Nevertheless, during the last five years, the dominance of these PCR based markers gave way to SNP markers which were first discovered in human genome and are believed to provide a universal form of genetic variations among same species (Ghosh et al., 2002).

They are more abundant and commonly found in plants and animals and their frequencies ranges by one SNP every 100-300bp (Edwards et al., 2007). With such large marker numbers it has become feasible to map the entire genome to find out highly associated markers for quantitatively inherited traits which is called genome-wide association studies (GWAS), or association genetics (McCarthy et al., 2008). They are bi-allelic that means within the population there are two commonly occurring base-pair possibilities for each SNP position on chromosome. Although thousands of SNP markers are being used in human and animal genome analysis but their use in plants is still in its infancy.

We consider Arabidopsis, a model organism for phenotyping and genetic studies, SNP markers have widely been used to identify loci underlying several traits. A high density 250,000 SNPs chip has been employed to map a panel of hundreds of natural inbred lines for numerous traits such as; shade avoidance (Atwell et al., 2010), in addition to regulation of allyl glucosinolate in various plant processes (Francisco et al., 2016), as well as differences in flowering times and pathogen resistance (Aranzana et al., 2005). In the present study, considering the protective detection of polymorphism, genome-wide mapping in Arabidopsis of over 250,000 SNP markers were used to investigate genetically associated variants to three different traits the response to cyst nematode *H. schachtii*.

1.7.3 GWAS in crops

Natural variation in crop plants is transpired mainly due to impulsive mutations in their wild ancestors. Breeding and domestication of crops have had an intense impact on the variation in genetic architecture present in modern crops. Knowing the genetic fundamentals of phenotypic variation and effect of crops domestication can help us employ these genetic resources to improve crop productivity. Therefore, using the huge germplasm and genetic approaches such as genetic populations, genome sequences, and genome-wide association studies (GWAS) as well as several other transformation techniques, scientists are able to determine the natural variations by correlating the phenotypic variations with causal sequence variants.

Genome-wide association studies (GWAS) offer the best use of naturally happening recombination events to identify the genetic loci underlying various complex traits with a high resolution. Initially, it has been generally used in identifying genes involved in various human diseases (Altschuler et al., 2008).With identification ofmillions of SNPs in human genome, this approach has been well established to create a high-density haplotype map of human genome (Hapmap, 2005). Nevertheless, with advances in various sequencing techniques, now it is becoming a prevailing tool to study the genetic architecture of different phenotypic variations in crops (Rafalski et al., 2010). Available diverse genotypes that require genotyping once followed by phenotyping under different environmental conditions, supports mapping strategies for specific QTLs in crops.

Barely, maize, wheat, soybean and rice are considered appropriate crops for GWAS (Huang et al., 2010; Pariyar et al., 2016, Kump et al., 2011 and Jia et al., 2013).

Genetic manners for several traits such as; leaf size, leaf angle, disease resistance and flowering times in maize have been investigated using GWAS (Tian et al., 2011; Kump et al., 2011and Buckler et al., 2009) were identified. Jia et al., (2013) performed GWAS across 916 genetically diversified traditional landraces and modern cultivars of millet and identified many novel genomic loci for various agronomic traits in multiple environmental conditions. In another study, 917 accessions were brought under investigation by GWAS to identify the QTLs for plant height and inflorescence traits in sorghum (Huang et al., 2014). Similarly, 615 barley cultivars were subjected to GWAS to trigger the genetic variants underlying ten agronomic and 32 morphological traits (Huang et al., 2014).

Since many years, GWAS has been quite successful to identify and characterize the putative QTLs conferring various traits of interest in soybean. Wang et al., (2008) identified and confirmed several markers associated with iron deficiency chlorosis (IDC) trait in two advanced breeding line populations of soybean using simple single repeat (SSR) markers. Considering these facts, this study was followed by Jun et al., (2008) and new genomic regions (QTLs) associated with seed protein contents were identified. Sonah et al. and Bastien et al. (2013, 2014) performed GWAS for 130 soybean breeding accessions to identify the associated QTLs and candidate genes underlying resistance to *Sclerotinia* stem rot (*Sclerotinia sclerotiorum*). Recently, over a diverse set of 298 soybean accessions GWAS was conducted for seed protein and oil content trait using 55,000 SNP markers and associated QTLs for high protein contents were detected at chromosome 20 (Hwang et al., 2014).

The response to soybean cyst nematodes (SCN), GWAS has also been effectively conducted to locate the associated genomic regions and identify the potential genetic variants underlying resistance to SCN. Li et al., (2011) identified six genetic markers significantly associated with SCN resistance across 159 soybean accessions genotyped for 55 SSR markers. In another study, over 282 soybeans breeding lines were mapped for thousands of SNPs and several major linked QTLs including rhg1 and FGAM1genes were identified conferring resistance to SCN *Heterodera glycine* (Bao et al., 2013).

Bi-parental populations have commonly been used to characterize the putative QTLs controlling resistance to soybean cyst nematode (SCN), *H. glycine* Ichinohe. Vuong et al., (2015) performed GWAS for 553 soybean plant introductions (Pls) to detect the

QTLs associated with SCN resistance. For this analysis 45,000 SNPs were used resulting in the identification of 14 loci comprising of 60 SNP markers that are significantly associated with SCN resistance. Recently, Zhang et al., (2016) screened 235 wild soybean (Glycine soja) to dissect the genetic basis for resistance against SCN and in addition to confirming previously known genes they identified other disease resistance related proteins with a leucine-rich region (LRR), a mitogen-activated protein kinase (MAPK) and a MYB transcription factor as promising candidate genes. More recently, GWAS was successfully deployed on a diverse population of 161 winter wheat accessions that were genotyped for 90,000 SNP arrays to identify the genetic regions associated with resistance to *Heterodera filipjevi* nematodes. With this study eleven novel QTLs were located on different chromosomes including 1AL, 2AS, 2BL, 3AL, 3BL, 4AL, 5BL and 7BL underlying the potential candidate genes (Pariyar et al., 2016).

Such precise and confined results indicate that GWAS can be employed effectively to define the associated genomic regions in diverse germplasms which facilitates positional cloning of the causal genes. Nevertheless, as with all scientific tools, it is imperative to understand the issues and limitations associated with GWAS studies. Population structure and kinship matrix are the basic challenges leading to false marker-trait associations resulting with false positives. Therefore, the linear mix model (LMM) is the most suitable method accounts for marker-trait associations and addresses these challenges including population structure and kinship matrix across the crop population. In addition, GWAPP, a web-based application for GWAS in Arabidopsis, was developed with an execution of a linear mix (LMM) model, also known as accelerated mix model (AMM) (Seren et al., 2012) that has been successful in investigating the genetic variants underlying various phenotypic traits.

1.7.4 GWAS in Arabidopsis

Discovering the genetic basis of complex quantitative traits in plants such as; growth rate, flowering time and especially yield has been a major focus for scientists to improve crops and general plant adaptations to different stimulating conditions. *Arabidopsis thaliana* has been a unique and appealing model organism to study natural variations and adaptations. Besides, available genomic resources, small genome size, habitat diversity and wide distribution (Shindo et al., 2007) leading to

different challenging conditions makes it an attractive organism to study the genetic basis of phenotypic variations.

In addition, sequenced genome and availability of hundreds of natural isogenic lines have made Arabidopsis an appropriate model for GWAS (Atwell et al., 2010). Once these accessions are genotyped, they can be phenotyped multiple times for various traits of interest in different environmental conditions. Arabidopsis is well-suited for association studies mainly for two reasons; first, a decrease in heterozygosity level due to inbreeding among associations enables haplotype inquiry by genotyping the accessions directly. Second, based on population size, LD decays have been estimated within 50-250 kb (Nordborg et al., 2002). Moreover, supporting GWAS concept in *Arabidopsis*, genetic variants for several traits including salt tolerance (Baxter et al., 2010), shade avoidance (Filiault et al., 2012), 107 different phenotypes (Atwell et al., 2010), flowering traits (Li et al., 2010) and glucosinolate levels (Chan et al., 2011) have successfully been dissected with this mapping approach.The current advances in GWAS mapping to explore the molecular processes of natural variation in Arabidopsis offers an opportunity to increase our knowledge regarding plant-stress interaction (biotic and a-biotic).

As far as Arabidopsis-nematode interaction is concerned, two of the most economically important groups, root-knot and cyst nematodes, readily infect Arabidopsis roots (Sijmon et al., 1991) providing an excellent opening to explore the molecular processes involved in the development of complex feeding structures induced. However, GWAS in Arabidopsis in response to cyst or root-knot nematodes has never been investigated. In this study, a geo graphically diversified natural population of 148 *Arabidopsis* accessions was used to investigate cyst nematodes susceptibility associated with quantitative traits loci (QTLs) across the whole genome and residing causal genes. Here, we have provided a new potential resource of resistance to cyst nematodes.

1.7.5 Quantitative trait loci (QTL)

In nature, many agronomic traits such as quality and productivity, tolerance to biotic and a-biotic stresses as well as several other disease-resistance traits require a cumulative action of many genes and environmental factors that complicate the breeding process. These complex traits are referred to as quantitative traits, polygenic or multi-factorial traits, and the region within the genome, where genes associated

with a certain trait reside is known as a quantitative trait locus (QTL). For a long time, in genetics and breeding, identify genes that confer trait variations has been a main goal.

During the 1980s, statistical approaches relied on means, variances and co-variances of relatives and have been extensively used to examine quantitative traits. With these studies it became feasible to partition the total phenotypic variance into environmental and genetic variances and further analysis of genetic variance in terms of additive, dominance and epistatic effects. In addition, it became also possible to estimate the hereditariness of the trait with minimum number of genes affecting the traits of interest. However, so far, a little information was available about chromosomal location and their ability to control traits. For any trait there were several segregating genes approximately with additive effect (Kearsey and Pooni, 1996) which were named "polygenes" by Mather (1949).

In parallel, the development of DNA molecular markers and statistical packages to analyse marker-trait association and examine the variations in quantitative traits has revolusiotionized this research area. With these improvements in molecular technology and availability of molecular markers, it is quite feasible to map quantitative trait loci (QTLs) successfully (Paterson et al., 1988; Stuber et al., 1992). QTLs as a term were first coined by Gellderman (1975) and its analysis is based on the principle of association between quantitative phenotype and genotype of markers. In this type of mapping, the population is portioned with markers into genotypic groups with/without a particular marker locus to determine whether significant differences exist between groups with respect to traits being considered (Tanksley 1993; Young 1996). Significant phenotypic differences among population groups indicate that the marker locus used to partition the population is linked to the QTL controlling the trait. QTL and marker gynotype association is challenging with natural population and exists only where the marker is completely linked to the QTL. Therefore, for QTL analysis segregating populations such as recombinant inbred lines (RILs), F2-derived populations, near-isogenic lines (NILs) and backcross populations are used. Analysis for QTL-trait associations is performed by the following approaches; single marker approach, simple interval mapping (SIM), composite interval mapping (CIM) and multiple intervals mapping (MIM).

Using experimental populations to identify QTLs, linkage mapping displays various limitations including low resolution power, limited number of recombinations, not suitable for the estimation of dominance effect and time requirements. Therefore, since 2001 natural population based mapping (association mapping or genome-wide association mapping) has emerged as a powerful tool for the detection of natural variations influencing complex traits, with the advantages of high mapping resolution and high allelic recombination without requiring any detailed mapping of populations.

1.7.6 Genome-wide association mapping (GWAS)

Many phenotypic traits are quantitative in nature and complex in etiology with multiple environments; therefore, to understand the genetically elucidated reasons, GWAS provide an unprecedented corridor to track the inherited and genetic architecture of complex traits. Over the past decades of using this approach, enormous progress has been made to mapmany traits in humans and animals. The major goal in medical genetics is to identify the genomic regions that are highly associated with various diseases including heart diseases, mental illness and cancer (Plomin et al., 2009). However, sequenced human genome has facilitated the the use of GWAS to access an individual's genotype at thousands of SNPs across the entire genome (Donnelly, 2008; Hindorff et al., 2009).

Price in, 2006 defined this system as that it is an approach based on linkage disequilibrium (LD) to map linked QTLs across the genome and to identify the genetic variants affecting the particular traits of interest. It is based on the principle that traits that have recently entered a population are more likely to be still associated with the genetic sequence of the original evolutionary ancestors. It focuses on an association between genetically variant alleles (SNPs) and the particular traits in a population across the complete genome. For association mapping it is pre-requisite to have an extensive knowledge regarding SNPs across the whole genome. In addition, the availability of high-density SNPs support the success of this approach by scanning the entire genome, often small haplotype blocks that are significantly correlated with quantitative trait variations. It is difficult to perform this approach for an organism whose genome has not been fully sequenced (Yu et al., 2008).

GWAS offers several advantages over traditional linkage mapping to dissect quantitative traits at higher resolutions because of higher recombination rate and density of SNP markers across the genome among the population. In addition, SNP

markers help to calculate precise population structures for GWAS, hence, the knowledge of genotypes pedigree or crosses are not required (Myleset al., 2009). For decades, various linkage analyses have been conducted in plant species while few QTLs were identified and dissected at the genetic level (Price, 2006). Therefore, association mapping has emerged to prevail over these limitations and complex trait variations at genetic level (Nordborg and Tavare, 2002).

1.7.7 Linkage disequilibrium (LD)

Linkage disequilibrium (LD) is a non-random association of alleles at different loci. However, loci are believed to be in linkage disequilibrium when the association frequency of their diverse alleles is found to be more or less than expected if the loci are segregating independently (Slatkin, 2008). In evolutionary biology, LD is of great importance as it indicates complete information of the history of natural selection as well as, various mutations and gene conversions across different loci. Several natural factors such as; mutation, recombination rate, genetic linkage and population structure have had much impact on the extent of LD in human genome. However, a recent study in Arabidopsis demonstrated that extensive inbreeding could produce high LD as a self-mating system may increase the homozygous recombination rates per meiosis (Nordborg et al., 2002), leading to maintaining LD over longer physical distances. The extent of LD is based on the physical associations of loci on chromosomes. Closely associated loci exhibit stronger LD than loci placed farther apart from each other on a chromosome. Moreover, decay of LD determines the amount required to tag a haplotype.

Association mapping considers the use of natural population of diversified cultivars with the purpose of obtaining more recombination event numbers that contributes to a higher resolution to find novel regions associated with particular traits. However, it suffers some challenges and limitations as well especially when the trait under examination is highly associated with population structure. This population-based survey or approach in plants includes the following steps: Selection of natural population or genetically diversified germplasm collection, measurement of phenotypic characteristics with multiple replications genotyped with several molecular markers (AFLPs, SSRs and SNPs), quantification of LD using molecular marker data, evaluation of population structure and kinship to limit false associations, correlation of phenotypic and genotyping data using appropriate statistical approaches

(GLM, SA, MLM), identification of linked QTLs of interest and finally characterization of candidate genes underlying traits **(Fig.1.7)**.

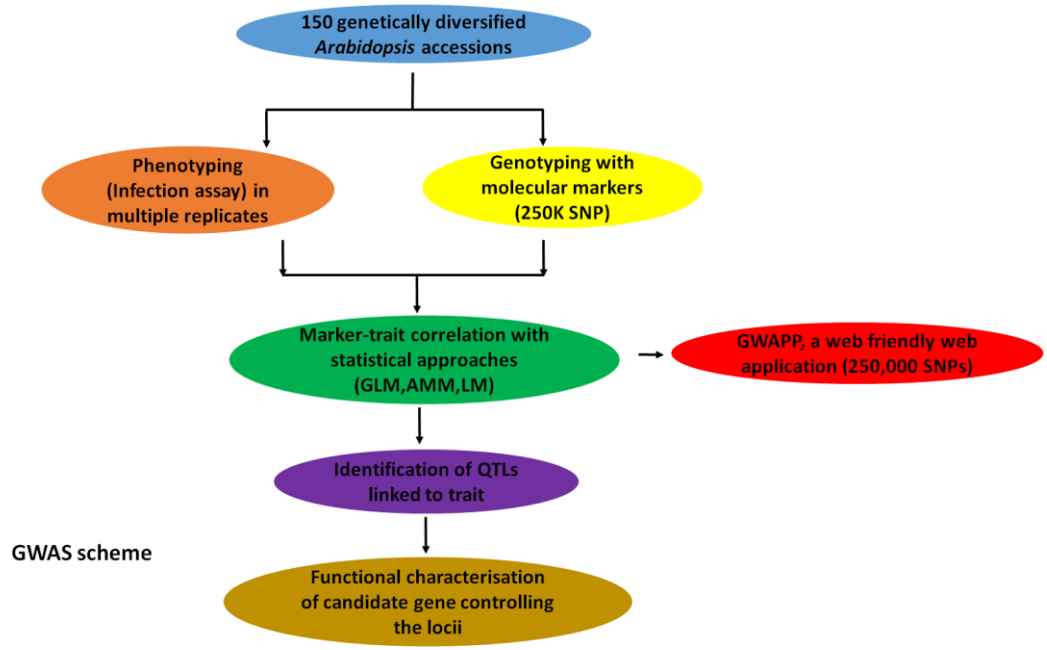


Figure 1.7: Schematic presentation of association mapping for tagging a gene of interest using germplasm (Abdurakhmonov and Abdukarimov, 2008). It describes all the steps involved in phenotype-genotype correlation using different statistical approaches (AMM, LM and GLM) and finally identification of linked QTL underlying causal genes.

1.8 Objectives of the study

Many efforts have been made to discover the susceptibility and resistance resources in Arabidopsis in response to various pathogens. Several QTLs for developmental traits have been identified via traditional bi-parental and genome-wide association mapping. However, in this study, an additional effort has been made to identify susceptibility resources in a diverse natural population of 148 Arabidopsis accessions against *H. schachtii* and in order to identify remarkable variations in the development of male and female nematodes and cyst sizes per accessions. Therefore, we hypothesized that several putative susceptibility associated QTLs will be identified and potential candidate genes residing these QTLs will be further predicted. However, the main objectives of this study are:

i. To screen 148 Arabidopsis accessions against cyst nematode *H. schachtii*.
ii. To genotype 148 accessions for marker-trait association using a web-based GWAPP application incorporated with 250k SNP Dataset v3.06; TAIR9 array.
iii. To identify putative QTLs linked to *H. schachtii* susceptibility.
iv. To identify potential candidate genes residing the significant QTLs.
v. To characterize the role of candidate genes in nematode parasitism.

Chapter 2

2.0 Materials and Methods

2.1 Phenotyping

2.1.1 Plant material and growth conditions

For the genome wide association studies (GWAS) an assembly of geographically distributed natural population of 148 *Arabidopsis* accessions was used. Three independent experiments were performed, each comprising of 50 accessions with 10 technical replicates and 1000 plants. A highly susceptible accession Col-0 from Columbia was used as a reference control. During statistical analysis the value of Col-0 was considered as 100%.

2.1.2 Seeds sterilization

Seeds were sterilized for 4-5 min in 0.7% sodium hypochlorite, and subsequently washed three times with sterile water for 1 min at room temperature. Seedlings were grown in 9 cm Petri dishes containing Knop medium under the optimized growth conditions at 22C° with twelve hour's photoperiod.

2.1.3 Knop media ingredients

Table 2.1: Stock Solutions per liter. The stock solution is prepared by combining the following chemical ingredients.

Stock solution	Chemicals	g/L
Stock solutionI	KNO3 MgSO4 - 7H2O	121.32g/L 19.71g/L
Stock solutionII	Ca(NO3)2 - 4 H2O	120g/L
Stock solutionIII	KH2PO4	27.22 g/L
Stock solutionIV	FeNaEDTA	7.34g/L
Stock solution V	H3BO3 MnCl2 CuSO4 -5 H2O ZnSO4 - 7 H2O CoCl2 - 6 H2O H2MoO4 NaCl	2.86g/L 1.81g/L 0.073g/L 0.36g/L 0.03g/L 0.052g/L 2g/L

To prepare one liter of Knop media, following quantity of stock solutions, sucrose and Diachin agar is used **(Table2.2).**

Knop Medium (Preparing 1 liter of Knop solid medium).

Table 2.2: Composition of Knop medium per 1 liter

Composition	Quantity
Sucrose	20 g/L
Diachin Agar	8 g/L
Stock solution I	2 ml/L
Stock solution II	2 ml/L
Stock solution III	2 ml/L
Stock solution IV	0.4 ml/L
Stock solution V	0.2 ml/L
B5 Vitamins	1 ml/L

MS media

Table 2.3: Preparation of Murashige and Skoog medium for the culturing of *Arabidopsis thaliana*

Ingredients	Quantity (g/L)
Murashige and Skoog mediumincludingVitamins and MES buffer	4.9 g L-1
Sucrose	20 g L-1
Agar Agar	10 g L-1
ddH2O	1L (1000ml)
pH	6.5

MS-Gelrite media

Table 2.4: Preparation of MS-Gelrite media for *Meloidogyne incognita* assay (1000ml)

Ingredients	Quantity
Murashige and Skoog medium including Vitamins and MES buffer	4.7g/L
Sucrose/Sugar	20g/L
ddH20	1 L
PH	5.7
Gelrite agar	5g/L

LB-medium

Table 2.5: Preparation of LB-medium (1L)

Ingredients	Quantity
Tryptone	10g
Yeast extract	5g
NaCl	10g
ddH2O	1L
PH	7.0

2.1.4 Nematode Infection Assay

H. schachtii cysts were harvested from the stock culture of mustard plants growing on Knop medium and collected into a cyst collecting funnel containing sterilized 3mM ZnCl2 to stimulate the hatching. After one week larvae (J2) were harvested and sterilized with 0.05mM HgCl2 subsequently washed three times with sterile water for 3 min at room temperature. Twelve days old seedlings were inoculated with 60-70 J2 per plant. Infected plants were placed back horizontally in growth room. To analyze the infestation rate, at twelve days post infection (12dpi) ratio of the number of males and females per plant was studied under the stereo microscope Leica KL 200 LED. At this stage nematode are developed well enough to differentiate the male and female. At 14dpi briefly the syncytia size with associated female nematodes was measured under the microscope Leica 2006. For each line 25-30 females were randomly selected and photographed.

2.1.5 Cysts collecting funnel preparation

An autoclaved funnel containing 3Mm Zncl2 to stimulate the hatching process was used to collect the cysts (300-400).

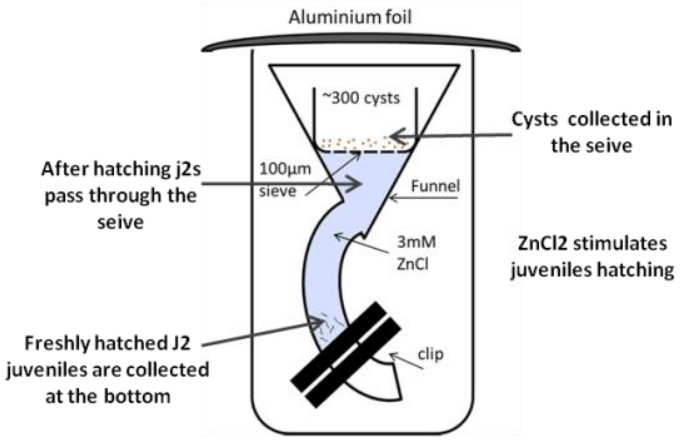


Figure 2.1: Cyst collecting funnel

2.2 *M. incognita* infection assay

2.2.1 *M. incognita* egg isolation from tomato roots

Infected tomato plants were collected from soil and the roots containing galls and carefully rinsed with water to remove the soil particles. Roots were cut into 1-3cm pieces and put into the Blender containing 500 ml 1.5% sodium hypochlorite (NaoCL). Root samples were mixed in blender for 15 seconds low and 15 seconds at high speed. Suspensions were transferred to a 5L bowl and filled with water up to 5L volume. To collect the eggs the sieves were arranged like 500, 250, 150, 100, 50, and 25um. The suspension was carefully poured through a 250um sieve nested over 25um sieve to collect the eggs. The sieves and roots material was rinsed with water and eggs were collected in a small 50um plastic beaker.

2.2.2 Eggs Purification

The eggs suspension was transferred in 50ml plastic tubes and spinned down at 3000rpm for 1 mint (accel 9, brake 0). The brown pellet was suspended into 35% autoclaved sucrose solution up to 35ml and with an additional layer of 1-1,5cm distilled water, spinned at 2000rpm for 3 mints. The floated eggs at the interface between sucrose and water were collected and transferred to a small 25um metal sieve. Immediately, eggs were rinsed several times with distilled water to remove sucrose. The eggs were collected in a 50ml Falcon tube.

2.2.3 Egg surface sterilization

The eggs were incubated for maximum 3 mints in 10% bleach and later washed 3 times with distilled water to remove the bleach completely. Cleaned eggs were transferred to a 50ml sterile plastic tube and 2ml of 22,5mg/ml gentamycine sulfate and 150ul Nystatin antibiotics were added. The eggs solution was poured in an autoclaved hatching chamber and placed at 27 ˚C for hatching.

2.2.4 Sterilization of *M. incognita* juveniles

After 4, 5 days of hatching, the solution around the sieve inside the hatching chamber was collected in a 50ml falcon tube and centrifuged at 4000 rpm for one min (accel 9 and brake 0). The pellet was suspended 1ml of 0.5% (W/V) streptomycin-penicillin solution for 20 mints. The nematodes were spun down for 1 mint (accel 9 and brake 0) and supernatants were removed carefully without disturbing the pellet. The juveniles were then incubated in 1ml of 0.1% (W/V) ampicillin-gentamycin solution

for 20 mints and again spun down at 4000 rpm for 1 min. After removing supernatants the nematodes were washed in 1ml sterile water for 5 mints. Tubes were centrifuged at 4000 rpm (accel 9 and brake 0) for 1 mint and supernatants were discarded. The juveniles were re-suspended and incubated for no longer than 3 mints in 0.1% (W/V) chlorhexidine solution and centrifuged at 4000 rpm (accel 9 and brake 0) for 1 min. The supernatants were discarded and juveniles were washed for 3 times with sterile water.

2.2.5 Infection to Arabidopsis with *M. incognita* J2´s

Twelve days old plants were subjected to surface sterilized *M. incognita* juveniles. Plants were infected with around 100 nematodes and stored in darkness to increase the infection. After 20 dpi the average numbers of galls were counted using a binocular.

2.3 GWA mapping

The average trait value of the biological replicates for each accession was taken. For the evidence of SNPs association with sex ratio of nematodes phenotype a web-based application GWAPP (https://gwas.gmi.oeaw.ac.at) (Seren et al., 2012) using accelerated mixed model (AMM) and linear regressions (LM; Seren et al., 2012). During analysis all the SNPs with minor allele frequency less than 0.5% was eliminated from the data. To avoid the intensity of false marker-trait association a nominal 5% threshold with Bonferroni correction was used. P-value bias due to population stratification was evaluated with Q-Q plots. The LD was visualized in the flanking region of the *AtS40-3* gene. In addition, the single nucleotide polymorphism (SNPs) among lowly and highly susceptible accessions with candidate genes from the novel genomic region was examined using Arabidopsis 1001 genome browser (http://signal.salk.edu). The DNA and protein sequences were aligned with CLUSTAL-omega (https://www.ebi.ac.uk).

2.4 Sequence analysis

DNA and proteins sequences were analyzed and compared with the sequences available in the database at National Centre of Biotechnology and Resource Centre (NCBI). To examine the differences in some of our extreme accessions with all five candidate genes from the novel QTL, we analyzed DNA and proteins polymorphism in the un-translated region (UTR), promoter, coding and non-coding regions of

candidate gene with Salk 1001 genome data base (http://signal.salk.edu). Interestingly, we found that there were nucleotide deletions in the promoter regions shared by *AtS40-3* and *PPR-proteins* in lowly susceptible accessions, whereas there were no such deletions in highly susceptible accessions. In addition, we found some remarkable changes in amino acids among lowly susceptible and highly susceptible accessions with *AtS40-3*genes. Later on these single nucleotide changes were also confirmed in some of extreme accessions with *AtS40-3* with real time PCR.

2.5 Gene expression analysis by real time PCR

To study which genes trigger the variation in *Arabidopsis* susceptibility to *H. schachtii* from highly associated genomic region, we assumed that the expression of genes involved in this pathway should be enriched in the plant tissues close-fitting the traits. We analyzed the expression of all candidate genes in 20 kb region surrounding the GWA peak by qRT-PCR in some of our lowly susceptible (Xan-1 and Van-0) and highly susceptible accessions (Zdr-1 and Kro-0) without nematodes infection. RNA was extracted from lowly susceptible and highly susceptible accessions and transcribed into cDNA using random primers [oligo (dN)6] and SuperScript III reverse transcriptase (Invitrogen), following the manufacturer's procedure and instructions.

Gene-specific primers were selected and were checked for gene specificity within the *Arabidopsis* genome. Actin gene was used as internal references. The samples for lowly and highly susceptible accessions were analyzed containing 10ul Fast SYBR Green Master Mix (Applied BioSystem), 10mM reverse and forward primers 1ul of each, cDNA 1ul and ddH2o 7ul in total volume of 20ul. The Arabidopsis gene Actin was used as an internal control. The analysis executed in a MicroAmp® fast optical 96 well reaction plate (Applied Biosynthesis, Darmstadt, Germany), with an ABI Step One Plus TM Real Time PCR System (Applied Biosynthesis, Darmstadt, Germany). The reaction was conducted in 40 cycles; 10 min at 95C°, each cycle at 95C° for 15s and for 60s at 60C°. Variations in the transcripts were calculated using ΔΔCt method (Pfaffl, 2001; Schmittgen and Livak, 2008). Three biological and three technical replicates were used.

2.6 Loss-of-function mutant analysis

To determine which gene underlie the variation in this strong genomic region, the T-DNA lines containing the axon insertions of all these candidate genes were obtained

from The European Arabidopsis Stock Centre (http://arabidopsis.info). The list of the total mutants used is obtained in the **Table 2.6.** To analyze the phenotype Infection assays were carried out for each mutant with three biological replications.

Table 2.6: Knock out mutant used in study

Nr.	Locus	ID	Source	Annotation
1	At4g18970	N678581	Salk_062226C	GDSL-like Lipase
2	At4g18980	N403134	GK-033F02	AtS40-3 senescence regulator protein
3	At4g18990	N661208	SALK_149853C	XTH29
4	At4g19000	N500734	Salk_000734	Transcription factor IWS2

Table 2.7: Primers for genotyping

Nr.	Genes	Forward Primers	Reverse Primers
1	At4g18970	CTGAGCTTCTTGGTTTCGATG	ACTACCGATGACAACATTCGC
2	At4g18990	TGATGGTGATGATGGTATCATGT	CAATAACACACTCCGGTGGA
3	At4g19000	GAGACCGCTCTTGATCAACTG	GAAGCGGTTCTGTAGATTTCG

Table 2.8: Primers for expression analysis

Nr.	Genes	Forward Primers	Reverse Primers
1	At4g18970	CTGAGCTTCTTGGTTTCGATG	ACTACCGATGACAACATTCGC
2	At4g18980	TAATGAGATCGAATCCCTTGACTT	ATTACTTCTCAAGCAAACCCCATA
3	At4g18990	TGATGGTGATGATGGTATCATGT	CAATAACACACTCCGGTGGA
4	At4g19000	ATCGAAATTCACAGGGAGGA	ATTTGCCACTTGGAACAAGG

Table 2.9: Primers for qRT-PCR

Nr.	Genes	Forward Primers	Reverse Primers
1	At4g18970	TGGGTTTAGGGTGACAAACG	ACAACATTCGCAGCTTCTCC
2	At4g18975	TGGGTCTGGACAGAAAGCAC	ATTTGCAACGCCTTAGCTGC
3	At4g18980	GAGGGAGAAATGACACCGCC	TCCAAAAAGCCGGTCATCCT
4	At4g18990	AGCCCTCGACGGTTGTAATG	TTCTCGGAACCGTCGCATAG
5	GFP-6	ACGTGTCTTGTAGTTCCCGT	ACTACCTGTTCCATGGCCAA
6	Hygromycin	CCGGTCGGCATCTACTCTAT	TTTCGATGATGCAGCTTGGG
7	AtS40-3 promoter (Xan-1)	CTCGCTAATAATTGTCTTTCCATCA ATC	ACGGTAGAAAAAAGAAGATCCACA GC

2.7 Genotyping and expression analysis

To ensure the true homozygous mutant, plants were grown in soil at green house and genotyped using primers obtained by the signal web resource http://signal.salk.edu (**Table 2.7**). Expression for the loss of function of genes was confirmed with real time PCR using gene specific primers **(Table 2.8)** covering the insertion.

2.7.1 DNA isolation and PCR

For the DNA isolation leaves from both wild type and mutants were obtained, grinded manually with plastic grinding sticks and isolated using modified CTAB method (Murray and Thompson, 1980; Sambrook and Russell, 2001). The quality and quantity of isolated DNA was measured in Nano Drop 2000C spectrophotometer (Peqlab, Erlangen, Germany). Primers for the candidate genes were designed covering insertion by Primer3 Plus software**.** Gene amplification for the knock-out mutants was performed in 25ul reaction mixture following standard PCR conditions: 5X Buffer: 5ul, 10mM dNTPs mix: 0.5ul, left primer: 1ul, right primer 1ul: left boarder primer: 1ul, Tag DNA polymerase: 0.15ul, ddH2o: 15.35ul and DNA template: 1ul. PCR was performed, 4 min at 95C°, followed by 30 cycles for 1 min at 94C°, 1 min at 60C°, I min at 72C° and final extension temperature at 72C° for 5 min. PCR products were stained on 1% agarose gels containing 100 ml of 1xTBE buffer with 5ul of peg-Green at 90C° for 50 min. Gel Documentation System in BIO-RAD, Gel DocTM XR was used for visualization and photographed under ultraviolet light.

2.7.2 RNA isolation and real time PCR

To analyze the expression of homozygous mutants with real-time PCR, total RNA was isolated from roots of mutant plants and Col-0 wild type plants according to manufacturer’s instruction. The quality and quality of isolated RNA was examined using a Nano-drop 2000c Spectrophotometer. To assure the purity of RNA, total isolated RNA was cleaned using RQ1-RNase-Free DNase protocol (Promega). A total mixture of 10ul containing RNA 10ul, RQ1 RNase-Free DNase 10X Reaction Buffer 1ul, and RQ1 RNase- Free DNase 1ul was prepared and incubated for 30 min. Then 1ul of RQ1 DNase stop solution was added to terminate the reaction.

Finally, to inactivate the DNase incubated at 65C° for 10 min. 2ul of clean and stable RNA was used to synthesis to generate the first strand cDNA (Schmittgen and Livak, 2008) using cDNA Reverse Transcription Kit" (Applied Biosystems, Darmstadt, Germany) following the manufacturer's protocol. The cDNA was synthesizedwith10 x RT buffer 2ul, 25mM dNTPs mix 2ul, 10mM Random primers 2ul, RNase inhibitor (1U/ul) 1 ul, reverse transcriptase (5U/ul) 1ul, RNA 2ul, and ddH2o 10ul in a total volume of 20ul using T gradient PCR cycler 8 (Biometra GmbH, GÖttingen, Germany).

2.7.3 Genotyping for Gabi-Kat mutants

To assure the real homozygous knock-out mutantsfrom Gabi-Kat, plants were grown on media containing anti-biotics sulfadiazine. In case of homozygous, all the plants were germinated and survived, for heterozygous few plants survived and rest did not survive after a few days of germination, whereas for Col-0 wild type, all plants died after a few days of germination **(Fig. 2.2)**.

Homozygous plants with 100% germination on Sulfadiazine antibiotics

Heterozygous plants with 50% germination on Sulfadiazine antibiotics

Heterozygous plants with no germination on Sulfadiazine antibiotics

Figure 2.2: Homozygosity test for Gabi-Kat mutants. Arabidopsis plants were grown on Sulfadiazine antibiotics and their germination rate was analyzed accordingly.

2.8 Gateway Cloning

The cloning procedure was followed by Gateway Cloning Kit from Thermo Fisher Scientific (Inc., Massachusetts (US)). This technique uses a set of recombination sequences, the “Gateway att” sites, and two enzyme mixes, called “LR-Clonase”, and “BP Clonase”. After cloning a promoter into an “entry vector”, this promoter can be sub-cloned in any other final “destination vector”. AtS40-3 promoter with nucleotide deletions from lowly susceptible accession Xan-1 and AtS40-3 promoter without nucleotide deletions from highly susceptible accession Col-0 was cloned into a Donor vector PMDC-207 using BP-reaction. Gateway attB1, and attB2” sequences are added to the forward and reverse primers. The primers were designed with flanking attB sites.

Table 2.10: AtS40-3 promoter specific primers with attB gateway sites

Nr	*attB* Gateway sites	Primers
1	GGGGACAAGTTTGTACAAAAAAGCAGGCTG	(F_p)CCTCGCTAATAATTGTCTTTCCATCAATC
2	GGGGACCACTTTGTACAAGAAAGCTGGGTC	(R_p)CTAGAAAGTAAGGAAAAATATAGAGTACG

DNA from lowly susceptible (Xan-1) accession was amplified with specific primers containing gateway sites under following conditions.

2.8.1 BP reaction

The PCR product is then mixed with Gateway "Donor vector PDonor 207" and BP clonase enzyme mix. The enzyme mix catalyses the recombination and insertion of the "*attB*" sequence containing PCR product into "*attP*" sites in the Donor vector (PDonor-207). The B sites match perfectly to the complementary P flanking sites of the Donor vector. Once the cassette is part of the target plasmid; it is called an entry clone with recombination sites "*attL*".

Reaction ingredients

*att*B-PCR product: 1-5 µl, pDONRTM vector (supercoiled, 150 ng/ul): 1 µl, TE buffer, pH 8.0 to 8 µl, BP ClonaseTM0.5 µl

Reaction procedure

The reaction mixture was incubated at 25c° for 60 min. Proteinase K was added and mixture was again incubated for 10 min at 37c°. The mixture was transformed to a competent *DH5α E. coli* cellsvial and the heat shock was performed for 55 s at 42 °C. The cells were streaked out on a LB plates containing gentamycin and overnight cultured. The colonies were selected and proven by colony PCR to contain the desired PCR product. Positive colonies were sent for Sanger sequencing (GATC Biotech AG, Konstanz (D)). After getting the complete sequences, the sequenced colony was cultured overnight containing the antibiotics of destination vector (PMDC-107) and plasmid was extracted with a Plasmid Extraction Kit (Macherey-Nagel, Düren (D)). In addition, a stock culture was prepared by dissolving the transformed cells in LB media in 25 % glycerol and stored at -80°C.

2.8.2 LR reaction

BP-reaction was followed by LR-reaction. The flanking sites of the entry clone will now be called L-sites and are complementary to R sites of the destination vector. The transformation was done accordingly as previously described. However, the selection of the positive clones was achieved with LB-medium containing Kanamycin antibiotics. An additional glycerol stock was prepared and stored at -80°C.

2.8.3. *Nicotiana* epidermis infiltration

Nicotiana benthamiana seeds were grown on soil containing pots under standard greenhouse conditions. Two weeks after germination, seedlings were transferred to new pots (only one plant/pot). After 5-10 weeks the plants were mature enough, used for leaf infiltration.

The a*grobacterium tumefaciens* cells were overnight cultured containing antibiotics (10 µg/mL of gentamicin, 50 µg/mL of rifampicin, and 50 µg/mL of kanamycin) at 28°C and re-suspended in infiltration buffer (50 mM MES, 2mM Na3PO4, 0.5 % 2D-Glucose, 100µM Acetosyringone). The infiltration was performed with a syringe in the lower sponge mesophyll of the tobacco leaf. The leaves were covered for white paper bag and placed in growth chamber. After 4-5 days RNA was extracted from leaves and GFP expression was analyzed with qRT-PCR. The optical density (OD) was confirmed around 0.8 and bacteria were harvested by centrifuging at 4000 rpm for 8 mints at room temperature. The pellet was re-suspended in infiltration buffer and co-infiltrated with RNA silencing inhibitor P19. After 2-4 hours of incubation, infiltration was performed with a 1ml syringe without needle in the lower sponge mesophyll of the tobacco leaf. The infiltrated plants were incubated in growth chamber for 5-10 days. Later on, the expression of GFP was analysed with qRT-PCR with gene specific designed primers. Hygromycine was treated as an internal control.

Chapter 3

3.0 Results

3.1 Phenotyping

To assess the variation in host responses to cyst nematode *Heterodera schachtii*, an assembly of geographically distributed natural population of 148 Arabidopsis accessions from 29 different countries was analyzed via nematodes infection assays. These accessions were ordered from Nottingham Stock center (see list of all lines in table S1). These accessions are a part of a set of natural accessions, which were sequenced using Illumina HiSeq 2000 platform by J. Ecker laboratory at the Salk Institute-USA as part of the 1001 Genome Project (NASC ID: N76636). Originally, all 195 accessions were ordered; however, due to germination and contamination issues, data from only 148 accessions were collected.

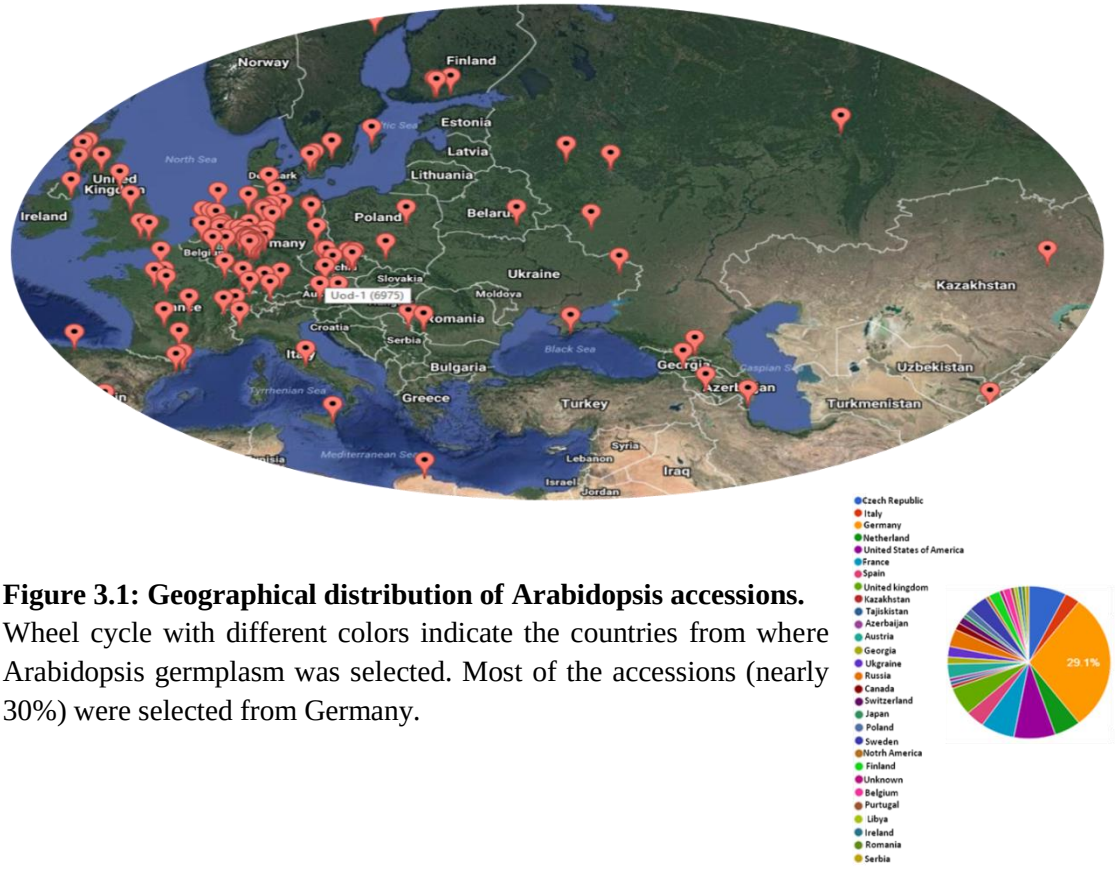


Figure 3.1: Geographical distribution of Arabidopsis accessions. Wheel cycle with different colors indicate the countries from where Arabidopsis germplasm was selected. Most of the accessions (nearly 30%) were selected from Germany.

3.2 Putative traits under study

Cyst nematodes are sexually dimorphic; however, sexual differentiation is greatly influenced by different environmental conditions. Under favorable conditions (healthy plant with sufficient nutrition) more females are produced, whereas under adverse conditions (poor plant health and lack of nutrition) more males are produced. Here, an assembly of the population of Arabidopsis was presented to cyst nematode *H. schachtii.* According to a standardized infection assay *H. schachtii* nematodes developed 10-12 days post infections (dpi) where they could be differentiated easily into males and females. It was assumed that across the entire genetically diversified natural inbred lines, allelic variations regarding male and female nematode numbers may occur. The aim of the study was to investigate the genetic architecture for these quantitative variations with genome-wide studies (GWAS). GWAS has been widely used to study genetic basis for various trait in Arabidopsis, but the response to cyst nematode *H. schachtii* has never been investigated.

In the following assays, three traits: i) the average number of female nematodes, ii) the average number of male nematodes and iii) the average female-to-male sex ratio of nematodes (iii), were analyzed to dissect the genetic basis for the quantitative variations in these traits. All the genotypes were categorized into three different ranks based on their intensity of susceptibility compared with Col-0 control plants. Accordingly, in case of the average number of female nematodes, accessions were categorized into three groups; i) lowly susceptible accessions, ii) moderately susceptible accessions, iii) highly susceptible accessions.

In case of the average number of male nematodes, accessions were categorized into three groups; i) accessions with low males, ii) accessions with similar or moderately increased males, iii) accessions with high males.

In case of female to male sex ratio, accessions were categorized into three groups; i) lowly susceptible accessions, ii) moderately susceptible accessions, iii) highly susceptible accessions.

3.3 Trait 1: Average numbers of female nematode/accession

148 Arabidopsis accessions were exposed to cyst nematode *H. schachtii* via infection assays. After 12dpi the average female nematodes population was assessed per accession. Statistical analysis revealed a great natural variation in the formation of the number of female per accession. Data was normalized to percentage by setting Col-0 as 100% and categorized into three groups: i) accessions with low female numbers (lowly susceptible), ii) accessions with similar or moderately increased female numbers (moderately susceptible), and iii) accessions with high female numbers (highly susceptible) **(Fig. 3.2)**. Thus accessions possessing females numbers ≤ 100% were considered lowly susceptible and accessions with females ≥150% were considered highly susceptible to cyst nematodes. Accordingly, 39.1% were lowly susceptible, 28% were moderately susceptible and 31.4% were highly susceptible.

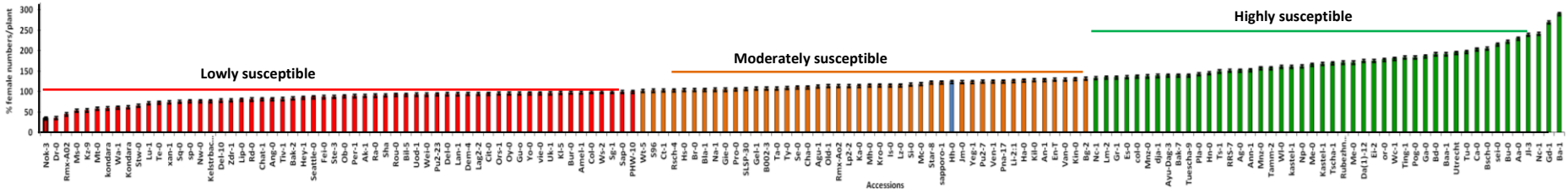
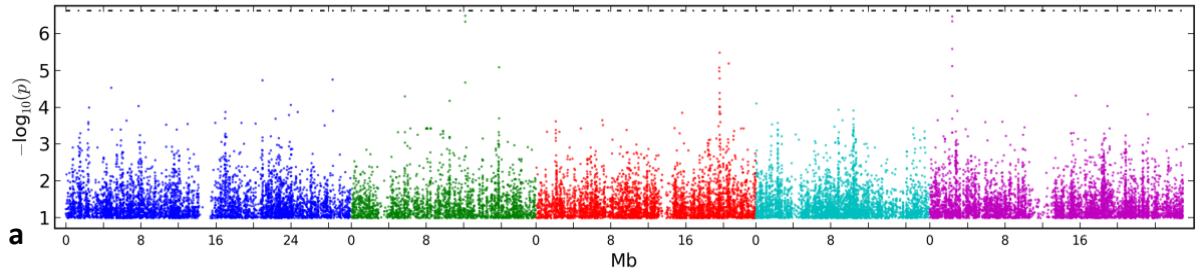


Figure 3.2: Phenotypic variation in the average number of female nematodes in 148 Arabidopsis accessions. The average numbers of female were counted at 12 days post infection (dpi) and accessions were categorized into three groups: i) accessions possessing low female numbers (lowly susceptible), ii) similar or moderately increased female numbers (moderately susceptible), and iii) high female numbers (highly susceptible) compared with Col-0 control plants. These values were converted into percentage (%) by setting the wild type values as 100%. Data was analyzed for the significance difference using T-test ($p<0.05$).

3.3.1 GWAS Results

To investigate the genetic mechanism underlying variation in female formation, genome-wide association mapping was performed. Over 250,000 SNPs were used for mapping the entire genome. GWAS identified several novel QTLs at different chromosomes associated with SNP locus markers using AMM and LM. However, the most significant QTL was identified at chromosome 5 with some significant markers such as; SNP positioned at 2392347bp with-log10 P (value) 6.02.

GWAS results using AMM



GWAS results using LM

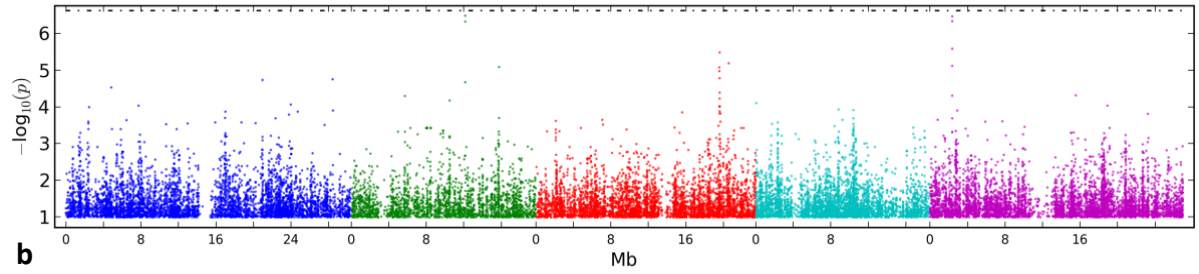


Figure 3.3: GWAS results for the regulation of female nematodes in 148 Arabidopsis accessions. Manhattan plot presents the association result from AMM (a) and LM (b) analysis using 250,000 SNP markers. Different colors indicate each chromosome. The horizontal dash-dot line corresponds to a nominal 5% significance threshold after Bonferroni correction. The x-axis represents the position on chromosomes, whereas Y-axis indicates -log10 (P) values for association of all the possible SNPs. SNPs with minor allele frequency <0.5 were removed.

3.3.2 Isolation of candidate genes

To identify candidate genes involved in female formation, a region of 20 kb neighboring to the most significant SNP locus from the novel QTL was further magnified. Detailed information for this SNP was obtained including pair-wise linkage disequilibrium (r^2) with other neighboring markers. The distribution of segregating alleles across the entire population between lowly and highly susceptible accessions was also determined. Using physical map at this locus**,** we selected six candidate genes; At5g07550 (glycine-rich protein-19), At5g07540

(glycine-rich protein-16), At5g07530 (glycine-rich protein-17), At5g07560 (glycine-rich protein-20), At5g07570 (gylcine-rich protein) and At5g07571 (oleosin family proteins) and considered influencing female nematodes formation **(Fig. 3.3c)**. Several other candidates were found to be interesting as well but based on their functional annotation and their role in association to cyst nematodes these six candidates were selected.

Magnified genomic region

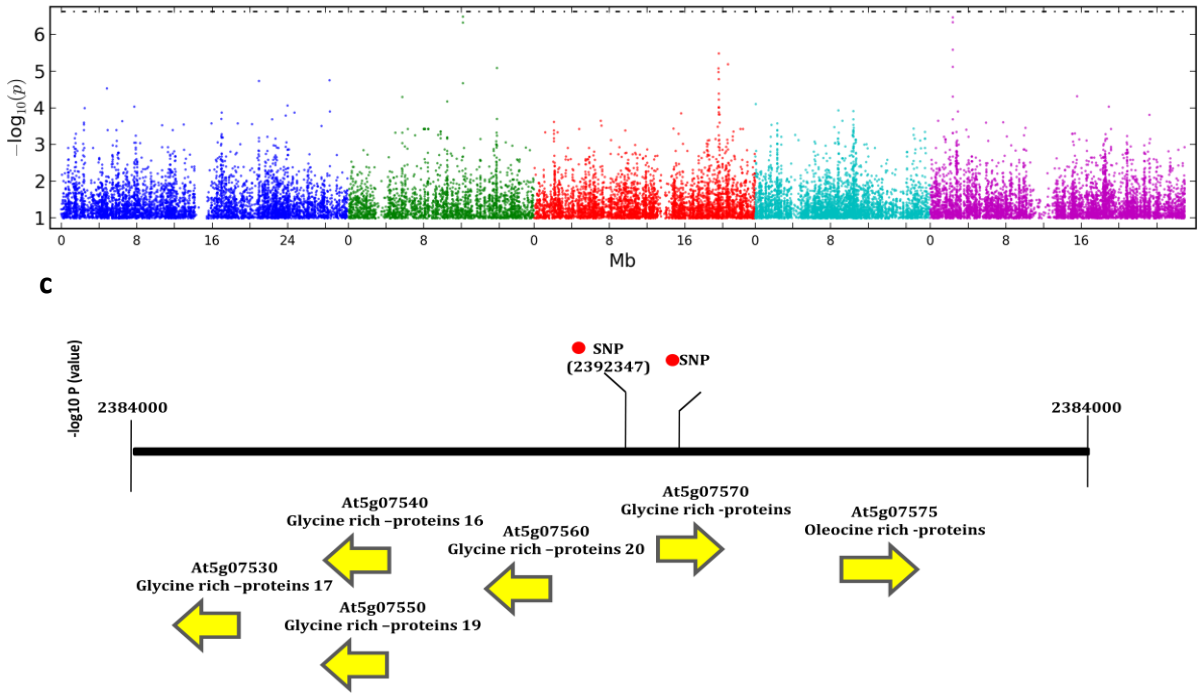
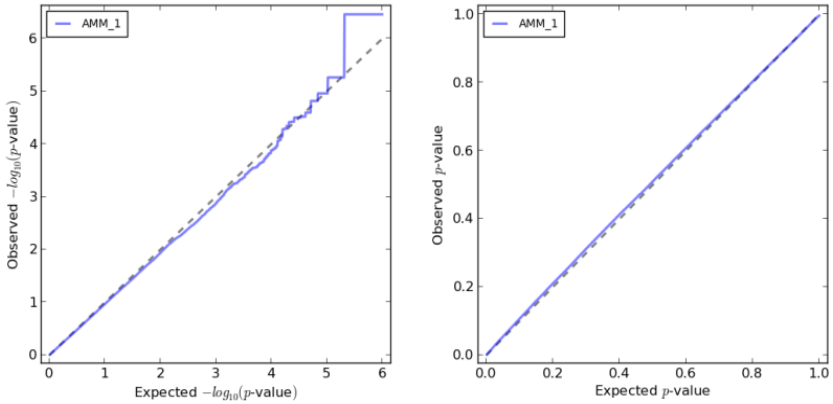


Figure 3.3c: GWAS results for the regulation of female nematodes in 148 Arabidopsis accessions. Manhattan plot presents the association result from LM analysis using 250,000 SNP markers for the average female nematodes per accession trait. In addition, 20 kb region from the putative QTL residing 6 candidate genes; At5g07530, At5g07540, At5g050, At5g07560, At5g07570 and At5g07571 neighboring the significant SNP is also indicated. QQ plots of observed versus expected negative logarithms of the P values of SNPs from AMM for the average female numbers/accession traits were generated. It indicates that false-positive SNP-trait associations due to population structure and kinship were well controlled in the mixed model.

Quantile-Quantile plot (Q-Q plot) using AMM



Quantile-Quantile plot (Q-Q plot) using LM

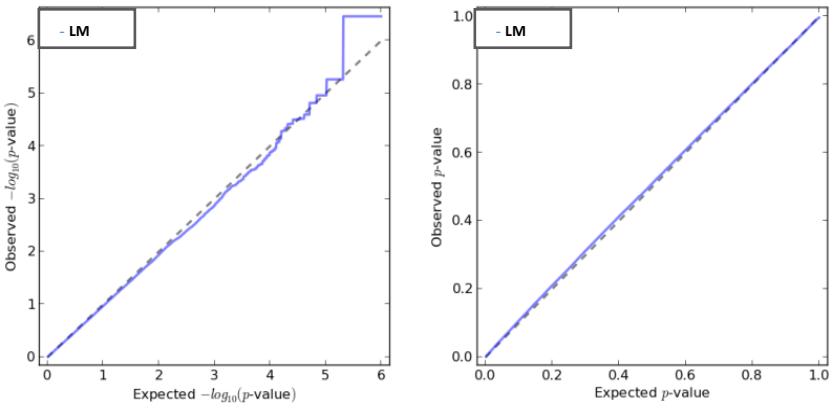


Figure 3.4: Quantile-Quantile plots. Q-Q plots of observed versus expected negative logarithms of the P values of SNPs from LM (a) and AMM (b) for average female numbers/accession trait were generated. It indicates that false-positive SNP-trait associations due to population structure and kinship were well controlled in the linear regression (LM) and accelerated mixed model (AMM).

Table 3.1: Summary of candidate significant SNP and nearest genes from the novel genomic region at chromosome five. Six putative genes with locus At5g07530, At5g07540, At5g07550, At5g07560 and At5g07571 neighboring the significant linked SNPs were selected for further analysis. In addition, the strength of SNPs association with the particular trait is also mentioned in terms of -log10 P (value). The distance of genes from SNPs with their functional annotation is also mentioned.

Nr.	SNP	Gene distance	Gene locus	Annotations
1	2392347bp	3098bp	At5g07550	Glycine-rich protein-19
2	2392347bp	5098bp	At5g07540	Glycine-rich protein-16
3	2392347bp	7823bp	At5g07530	Glycine-rich protein-17
4	2392347bp	1272bp	At5g07560	Glycine-rich protein-20
5	2392347bp	-3982bp	At5g07570	Glycine-rich protein
6	2392347bp	-5076bp	At5g07571	Oleosin family proteins

3.3.3 Selection of lowly and highly susceptible accessions

Finally, to analyze variations in female nematode numbers more intently, we selected twenty lowly susceptible (Nok-3, Dr-o, Rmx-AO2, MS-0, Kz-9, Mt-0, Kondara, Wa-1, Stw-0 and

Xan-1) and highly susceptible (Tu-0, Ca-0, Bsch-0, Sei-0, Bu-0, Aa-0, Jl-3 and Nc-1) accessions. Lowly susceptible accessions such as; Nok-3, Dr-0, Rmx-Ao2 and Ms-0 showed a remarkable variation with notably decreased female numbers. In contrast, highly susceptible accessions such as; Ba-1, Gd-1, Nc-1 and Jl-3 contained two times the females in comparison with Col-0 control plants. These variations in female numbers underline the importance of allelic natural variation and isolation of their genetic variants in Arabidopsis in different environments.

3.3.4 Infection assays of lowly and highly susceptible accessions

Lowly and highly susceptible accessions were infected with *H. schachtii* nematodes several times and the average numbers of developed female were counted at 12dpi.

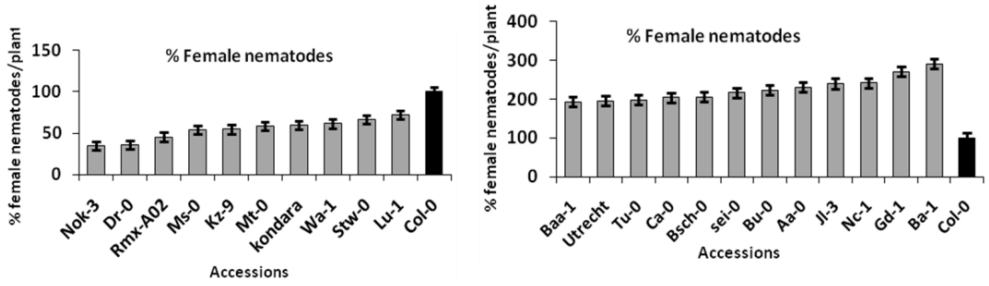


Figure 3.5: Phenotypic variations among lowly and highly susceptible accessions. Comparison of average numbers of female nematodes/accessions between lowly and highly susceptible accessions. Statistical data indicates female numbers among lowly susceptible accessions are highly variable. Numbers of female nematode/accession were counted after 12dpi.

Based on these variations, it was assumed that segregating marker-locus in these genotypes might be tightly associated with causal alleles and is confined to a particular genomic region influencing female formation across the entire population.

3.3.5 Associated alleles with highly and lowly susceptible accessions

Candidate SNP marker locus revealed an additional information of accessions carrying alleles more specifically that allele A was segregated across lowly susceptible accessions except Rmx-A02 with allele G, whereas allele A was normally distributed across highly susceptible accessions. These consistent allelic differences indicate that this candidate SNP marker locus might be associated with quantitative variations in female nematode development.

Table 3.2: List of lowely susceptible accessions with origin and associated alleles

Nr.	Accession	Alleles	Country	Latitude	Longitude
1	Nok-3	A	Germany	52.24	4.45
2	Dr-0	A	Germany	51.05	13.73
3	Rmx-A02	G	USA	42.04	-86.51
4	Ms-0	A	Russia	55.75	37.63
5	Kz-9	A	Kajikistan	49.50	73.10
6	Mt-0	A	Libya	32.34	22.46
7	Kondara	A	Tajikistan	38.48	68.49
8	Wa-1	A	Poland	52.30	21.00
9	Stw-0	A	Russia	52.00	36.00
10	Xan-1	A	Azerbaijan	38.65	48.80

Table 3.3: List of highly susceptible accessions with origin and associated alleles

Nr.	Accession	Alleles	Country	Latitude	Longitude
1	Ba-1	G	UK	56.55	-4.80
2	Utrecht	G	Netherland	52.09	5.11
3	Tu-0	G	Italy	45.00	7.50
4	Ca-0	G	Germany	50.30	8.27
5	Bsch-0	G	Germany	50.02	8.67
6	Sei-0	G	Italy	46.54	11.56
7	Bu-0	G	Germany	50.50	9.50
8	Aa-0	G	Germany	50.92	9.57
9	Jl-3	G	Czeck republic	49.20	16062
10	Nc-1	G	France	48.62	6.25

3.3.6 SNPs and promoter analysis

To examine single nucleotide polymorphism (SNPs) and amino acid changes among twenty lowly and highly susceptible accessions, we compared DNA and protein sequences of all candidate genes using Arabidopsis 1001 Genome Browser (http://signal.salk.edu). It turned out that SNPs were present with random distributions in the coding sequences of several candidates including At5g07570 (Glycine-rich repeat) and At5g07571 (Oleosin-rich repeat). However, there were no consistent differentiating SNPs and amino acid variations between lowly and highly susceptible accessions with any of the candidate genes that would narrow down the search for true causal genes effecting the formation of female nematodes. In addition, we analyzed the SNPs (nucleotide deletions, additions or substitutions) in the promoter regions of all candidate genes and did not observe any promising changes that would lead to expression alterations of candidates between lowly and highly susceptible accessions.

It is pre-mature to conclude that these genes play an important role in influencing the variation in female formation without complete molecular characterization including

knockout mutant analysis and molecular transformations of all potential candidates in relation to a particular phenotype. Nevertheless, these are potential candidates that might play an important role with additive effects in controlling the quantitative genomic region (QTL) at chromosome five. Therefore and in order to attain precise and more accurate evidence regarding differentiated SNPs and amino acid changes among extreme bulks with causal alleles, we extended the GWAS study to include another quantitative trait “the average numbers of male nematode per accessions”.

3.4 Trait 2: Average numbers of male nematode/accession

To understand the genetic architecture for male nematodes variations, systematic infection assays were made for genetically diversified population.Three experiments were performed each comprising of 50 accessions. The numbers of male nematode were counted after 12 dpi. To normalize the quantitative data, the absolute values were transformed into percentage (%) by setting Col-0 control accessions as 100%. Statistical data was categorized into three groups: i) accessions with decreased male numbers, ii) accessions with similar or moderately increased male numbers, and iii) accessions with highly increased male numbers **(Fig. 3.6)**. Accessions possessing male numbers ≤ 100% favored development of low males; accessions possessing males 100-150% favored a considerable increase, whereas accessions with males ≥150% favored high male numbers development.

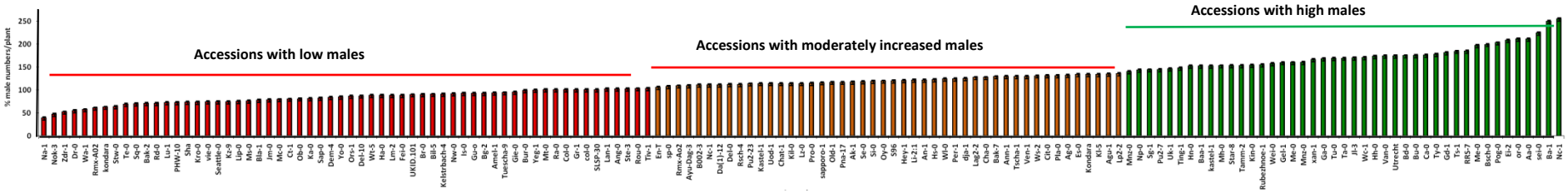
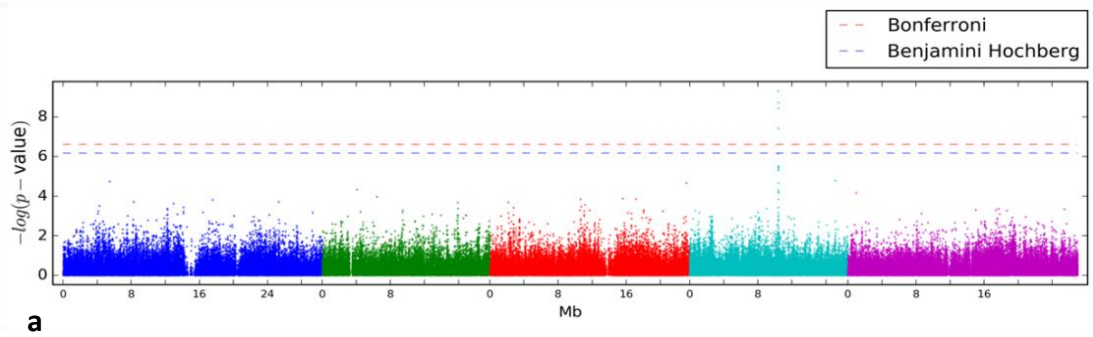


Figure 3.6: Phenotypic variation in average number of male nematodes in 148 Arabidopsis accessions. The average number of males was counted at 12 days post infection (dpi) and these accessions were categorized into three groups: i) accessions with low male numbers, ii) accessions with similar or moderately increased male numbers, and iii) accessions with highly increased numbers of male nematode compared with Col-0 wild types. These values were converted into percentage (%) setting Col-0 wild type values at 100% as a reference control. Data was analyzed for the significant difference using T-test ($p<0.05$).

3.4.1 GWAS results

Based on quantitative variation in male proportion in Arabidopsis accessions, GWAS was performed to identify QTLs associated with male formation. Over 250,000 SNPs were used for genome wide association mapping. GWAS identified several linked QTLs on different chromosomes with many SNP markers. The most significant QTL was detected at chromosome four in association to SNP with -log10 P (value) 9.32. However, three more neighboring SNPs with significance above 5% threshold with Bonferroni correction were considered highly associated with male nematode formations.

GWAS results using AMM



GWAS results using LM

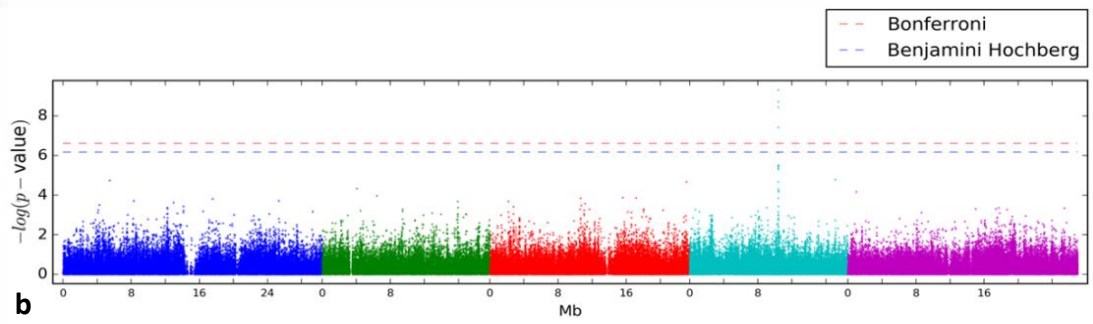


Figure 3.7: GWAS results for the regulation of male nematodes in 148 Arabidopsis accessions. Manhattan plot of GWAS results for the average male nematodes per plant using AMM (a) and LM (b). Different colors are representing each chromosome. The horizontal dash-dot line corresponds to a nominal 5% threshold with Bonferroni correction. The x-axis represents the position on chromosomes, whereas Y-axis indicates -log10 (P) values for association of all possible SNPs.

3.4.2 Isolation of candidate genes

To identify causal genetic variants influencing male formation, a region of 20kb from the novel QTL was explored **(Fig. 3.7c)**. Several candidate genes were found to be interesting but based on their functional annotation and the response to pathogen interactions, five of the novel genes with locus At4g18970 (*GDSL-lipase*), At4g18975 (*PPR-proteins*), At4g18980

(*AtS40-3*), At4g18990 (*XTH-29*) and At4g19000 (*IWS2*) were selected in association with significant SNP marker.

Magnified genomic region

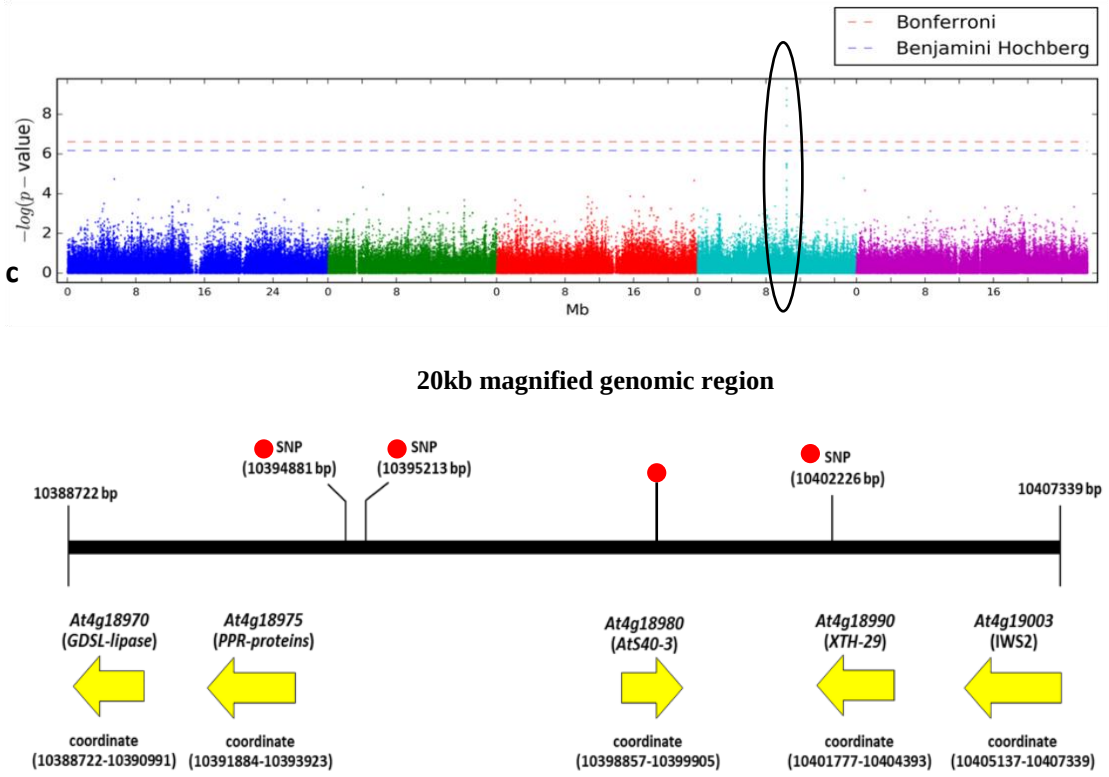
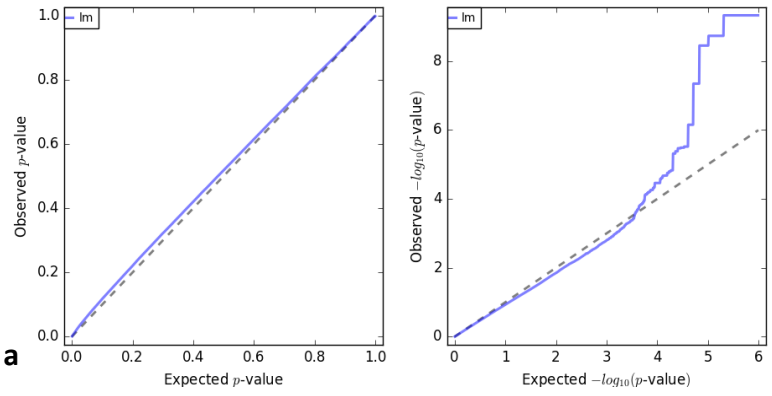


Figure 3.7c: GWAS results for the regulation of male nematodes in 148 Arabidopsis accessions. Manhattan plot of GWAS result for the average male nematodes per plant using AMM. In addition, a region from 10388722bp to 10407339 (18617bp) was magnified to identify the most significant SNP markers in association to causative genes. Four SNP markers positioned at 10394881bp, 10395213bp, 10394611bp, 10402226bp and 10403363bp above the threshold were selected to be truly associated with variation in male nematode numbers. Five candidate genes such as; at4g18970, at4g18975, at4g18980, at4g18990 and at4g19000 were selected in close proximity to the most significant SNPs.

Quantile-Quantile plot (Q-Q plot) using LM



Quantile-Quantile plot (Q-Q plot) using AMM

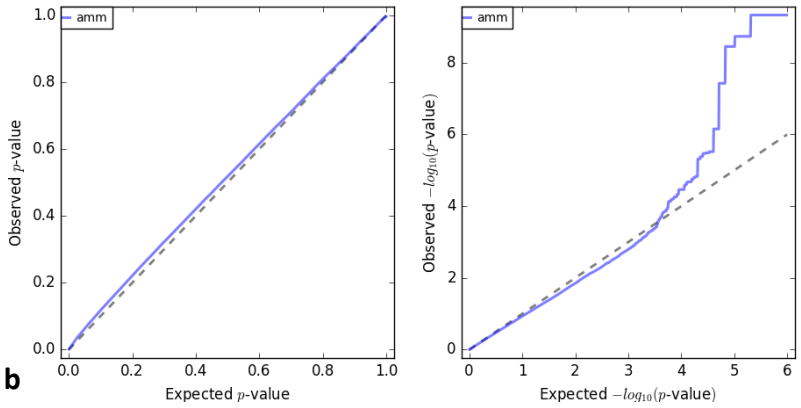


Figure 3.8: Quantile-Quantile plot. Q-Q plots of observed versus expected negative logarithms of the P values of SNPs from LM (a) and AMM (b) for the average female numbers/accession trait were generated. It indicates that false-positive SNP-trait associations due to population structure and kinship that is well controlled in the linear regression (LM) and accelerated mixed model (AMM).

Table 3.4: Summary of possible significant SNPs and the nearest genes from the novel genomic region at chromosome four. Five putative genes with locus At4g18970 (*GDSL-lipase*), At4g18975 (*PPR-proteins*), At4g18980 (*AtS40-3*), At4g18990 (*Endotransglucosylase-29*) and At4g19000 (*IWS2*) neighboring the significantly linked SNPs were selected for further analysis. In addition, the strength of SNPs association with particular trait is also mentioned in terms of -log10 P (value).

Nr.	SNP	Chr	-log10 P	Nearest genes	Gene annotation
SNP 1	10394881	4	9.32	At4g18975	Pentatricopeptide repeat PPR-proteins
SNP 1	10394881	4	9.32	At4g18970	GDSL-lipase/estrases super family
SNP 2	10395213	4	8.73	At4g18980	AtS40-3 a nuclear targeted protein
SNP 3	10394611	4	6.15	At4g18980	AtS40-3 a nuclear targeted protein
SNP 4	10402226	4	8.45	At4g18990	Endotransglucolysase (XTH-29)
SNP 5	10403363	4	7.43	At4g19000	Transcription factor (IWS2)

3.4.3 Selection of extreme accessions with low and high males

Finally, two extreme bulks showing lowest and highest males were selected based on the allelic polymorphism at associated SNP locus for analysis of candidate genes as well as for functional characterizations. Na-1, Nok-3, Zdr-1, Dr-0, Wa-1, Rmx-AO2, Kondara and Stw-0 possessed low male numbers, whereas Bsch-0, Pog-0, Ei-2, Or-0, Aa-0, Sei-0, Ba-1 and Nc-1 showed increased male numbers compared with Col-0 control plants.

3.4.4 Infection asssays of accessions with high and low males

Accessions with low and high males were selected and infected with *H. schachtii* nematodes several times to analyse the density of male nematodes formation. Numbers of developed male were counted after 12 dpi in each accession and a high variation was observed.

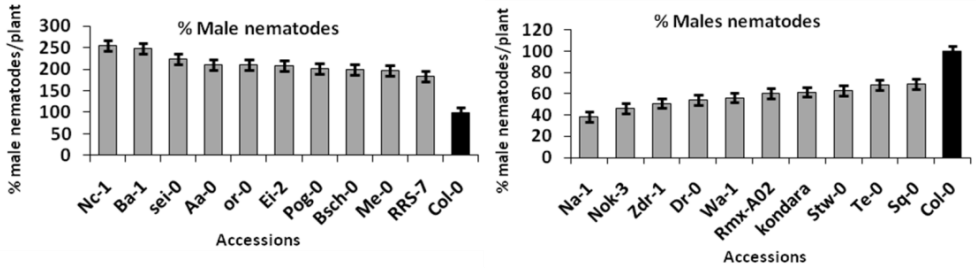


Figure 3.9: Phenotypic variation among extremely variable accessions. The average numbers of male nematodes per accession between extremely variable accessions. Statistical data indicates variation with male population among accessions possessing low males is remarkable. The average numbers of male nematode per accession were counted after 12 days post infection (dpi).

3.4.5 Associated alleles with extreme accessions with low and high males

Candidate SNPs marker locus positioned at 10394881bp revealed that allele A was segregated across lowly susceptible accessions, whereas allele G was normally distributed across highly susceptible accessions, except Nc-1, Ba-1 and Nc-0 carrying allele A. These allelic differences indicates that this candidate SNP marker locus might be one of the SNPs associated with variation in male nematodes.

Table 3.5: List of accessions with low males and associated alleles

Nr.	Accession	Alleles	Country	Latitude	Longitude
1	Na-1	A	France	47.50	1.50
2	Nok-3	A	Netherland	52.24	4.45
3	Zdr-1	A	Czeck republic	49.39	16.25
4	Dr-0	A	Germany	51.05	13.73
5	Wa-1	A	Poland	52.30	21.00
6	Rmx-Ao2	A	USA	42.04	-86.51
7	Kondara	A	Tajikistan	38.48	68.49
8	Stw-0	A	Russia	52.00	36.00
9	Te-0	A	Finland	60.06	23.30
10	Bak-2	A	Georgia	41.79	43.48

Table 3.6: List of accessions with high males and associated allele

Nr.	Accession	Alleles	Country	Latitude	Longitude
1	Nc-1	A	France	48.62	6.25
2	Ba-1	A	UK	65.55	-4.80
3	Sei-1	G	Italy	46.54	11.56
4	Aa-0	G	Germany	50.92	9.57
5	Or-0	G	Germany	8.01	50.38
6	Ei-2	G	Germany	6.30	50.30
7	Pog-0	G	Canada	-123.21	49.27
8	Bsch-0	G	Germany	8.67	50.02
9	Mc-0	A	UK	-2.30	54.62
10	RRS-7	G	USA	-86.43	41.56

3.4.6 SNPs and amino acid changes

To evaluate SNPs and amino acid changes with all five candidate genes (*GDSL-lipase*, *PPR-proteins*, *AtS40-3*, *XTH-29* and a transcription factor *IWS2*) among extremely variable accessions (with low and high male numbers) we used Arabidopsis 1001 genome browser (http://signal.salk.edu). In comparison between accessions with low and more males, we found random distribution of SNPs and amino acid changes in all candidate genes. But this polymorphism was not consistent between these accessions. In addition, with promoter analysis of these candidates we found that *PPR-proteins* and *AtS40-3* carry bi-directional promoter sharing transcriptional and expression regulations. Similarly, in case of *XTH-29* and *IWS2*, the transcriptional activities were also controlled by their common bi-directional promoter.

Taken together, SNPs were detected across all candidates; nevertheless they did not follow any consistent patterns. Therefore, we can speculate that these genetic variants might have additive effects in controlling QTLs associated with male formation. Bulks of extreme accessions display a high variable phenotypic effect regarding male and female nematode which needs to be investigated more meticulously.

3.4.7 Sex ratio in female associated extreme accessions

We had remarkable variation in female numbers among extremely variable accessions with this quantitative trait. Nok-3, Dr-0, Rmx-A02, Kz-9, Mt-0, MS-0, Kondara, Wa-1, Stw-0 and Xan-1 experienced lowest female numbers, whereas Ba-1, Utrecht, Tu-0, Ca-0, Bsch-0, Sei-0, Bu-0, Aa-0, Jl-3 and Nc-1displayed notable increase in female numbers. The average numbers of female nematodes were divided by the average numbers of males in each accession to obtain their sex ratio. We found a high variation in female to male sex ratio among these extreme accessions.

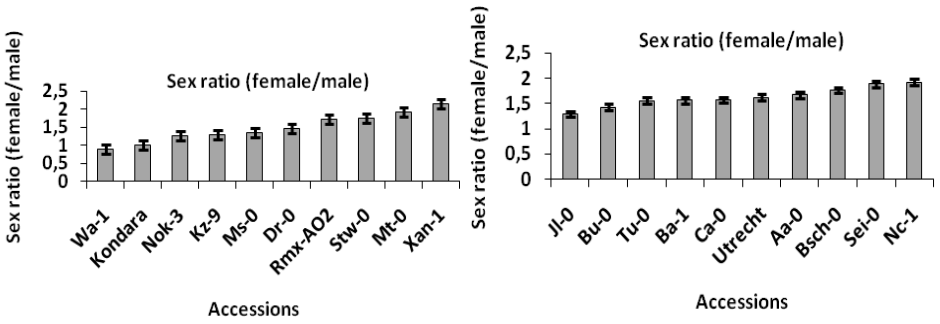


Figure 3.10: Quantitative sex ratio variation in lowly and susceptible accessions. Lowly susceptible accessions show more sex ratio variation (ranged from 0.7 in Wa-1to 2 in Xan-1) as compared with highly susceptible accessions.

3.4.8 Sex ratio in male associated extreme accessions

We had high variations in male numbers between extremely variable accessions. Na-1, Nok-3, Zdr-1, Dr-0, Wa-1, Rmx-A02, Kondara, Stw-0, Te-0 and Bak-2 developed low males, whereas Nc-1, Ba-1, Sei-0, Aa-0, Or-0, Ei-0, Pog-0, Bsch-0, Mc-0 and RRS-7 developed more males compared with Col-0 control plants. To analyze female to male sex ratio, the average number of female nematodes were divided by the average number of male nematodes in each accession. These accessions showed a high variation in female to male sex ratio.

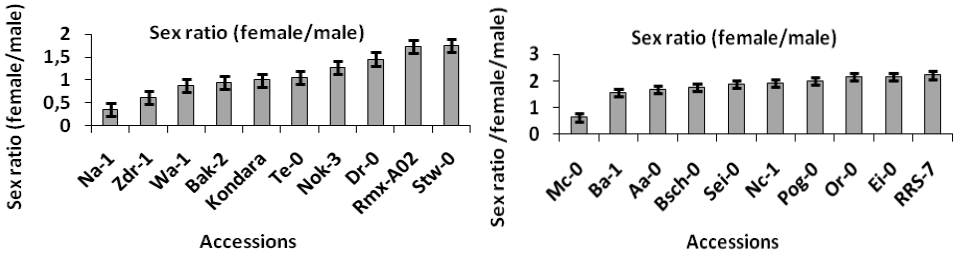


Figure 3.11: Quantitative sex ratio variation in accessions with low and high males. Accessions with low males show more variation (ranged from 0.4 in Mc-0 to 2.3 in Stw-0) in contrary to accessions with high males.

Female to male sex ratio variation among these extremely variable accessions with both traits; the average number of male nematodes/accession and the average number of female nematode/accession, represents the perspectives of genetic studies underlying these variations. Sex ratio variation in accessions with low females were remarkable. It ranged from 0.8, 0.9 (in case of Wa-1 and Kondara) to 2, 2.1 (in case of Xan-1 and Mt-o). Similarly, a high variation in sex ratio among acessions with low males was observed. It ranged from 0.3 (in case of Nc-1) to 1.7 (in case of Stw-0 and Rmx-AO2). Therefore, GWAS studies were

switched to identify QTLs associated with sex ratio variation and functional analysis of residing genes.

3.5 Trait 3: Sex ratio of nematodes/accession

We collected the male and female data after 12 dpi and converted it to female to male sex ratio for the all accessions. Although in general the accessions showed random variations but for better understandings and to normalize the quantitative data, we transformed it into percentage (%) setting Col-0 control values as 100%. Based on the female/male sex ratio, we categorized all genotypes into three susceptibility groups (sex ratio 0-0.5, lowly susceptible; sex ratio 0.5-1, moderately susceptible; sex ratio >1, highly susceptible). Accordingly, 46.6% were lowly susceptible, 26% were moderately susceptible, and 25.4% were highly susceptible.

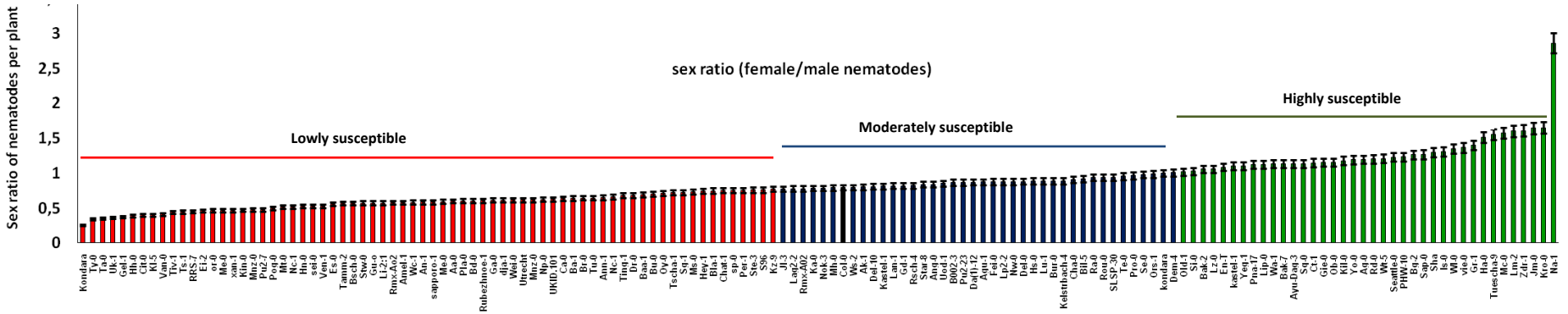
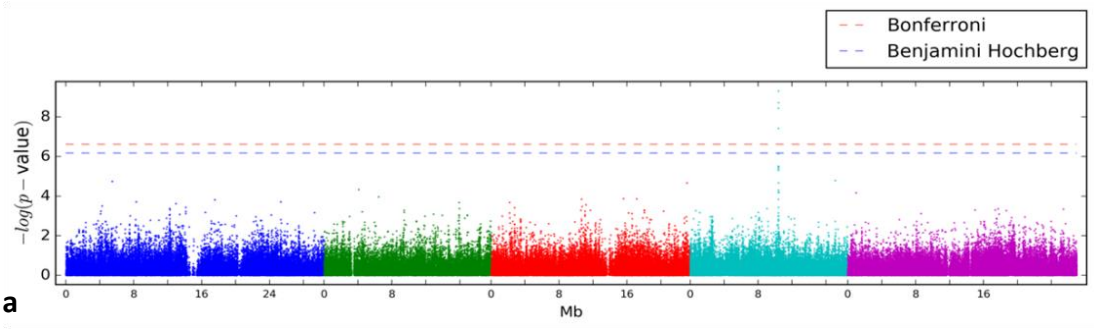


Figure 3.12: Phenotypic variation in the average female to male nematodes sex ratio in 148 Arabidopsis accessions. Statistical data for this trait showing that accessions were categorized into three groups: i) accessions with decreased female to male sex ratio (red shaded), ii) accessions with similar or moderately increased female to male sex ratio (blue), and iii) accessions with highly increased female to male sex ratio (green) of nematodes compared with Col-0 wild types. The average sex ratio was determined at 12 days post infection (dpi) and these values were converted into percentage (%) by setting Col-0 wild type values as 100% used as a reference control. Data was analyzed for the significance difference using T-test ($p<0.05$).

3.5.1 GWAS results

GWAS results revealed many QTLs at different chromosome with several significant SNPs. The most prominent QTL was detected at chromosome four in association with four significant SNP markers. The identified QTLs for female to male sex ratio and the average numbers of males/plant were co-localized on same chromosome. Apparently this co-localization of similar QTLs between the two traits might be due to more variations driven in male numbers/accession among the whole population. Five significant SNPs above 5% threshold with Bonferroni correction considered highly associated with variation in the average sex ratio of nematodes. Entire genes including their un-translated, coding or the non-coding regions, within 20 Kb strech were supposed to be associated with one of these significant SNPs. Additionally, we calculated pair-wise linkage disequilibrium (LD) in this novel region indicating the strong association of haplotypes with higher LD value denoting strong LD.

GWAS results using AMM



GWAS results using LM

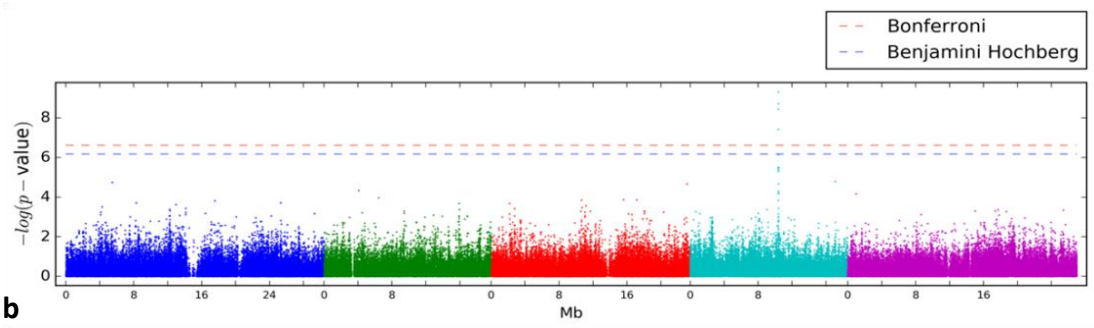


Figure 3.13: GWAS results for the regulation of female to male sex ratio of cyst nematodes in 148 Arabidopsis accessions. Manhattan plots of GWAS result for the average female to male sex ratio of cyst nematodes per plant using AMM (a) and LM (b). Different colors are representing each chromosome. The horizontal dash-dot line corresponds to a nominal 5% threshold with Bonferroni correction. The x-axis represents the position on chromosomes, whereas Y-axis indicates -log10 (P) values for association of all possible SNPs.

3.5.2 Isolation of candidate genes

To access candidate genes associated with significant SNP markers locus, 20Kb neighboring region from the putative genomic region was further investigated (**Fig. 3.13c**). To access which of the genes in proximity to highly associated genomic region underlie variation in sex ratio, a total of the five candidates with locus At4g18970 (*GDSL-lipases/estrases* family), At4g18975 (*PPR-proteins* super family), At4g18980 (*AtS40-3* a nuclear targeted proteins that modulates senescence in plant leaves), At4g18990 (*Endotransglucosylases/hydrolases XTH-29*) and At4g19000 (*IWS2* a transcription factor) neighboring to significant SNPs were selected for further investigations.

Magnified genomic region

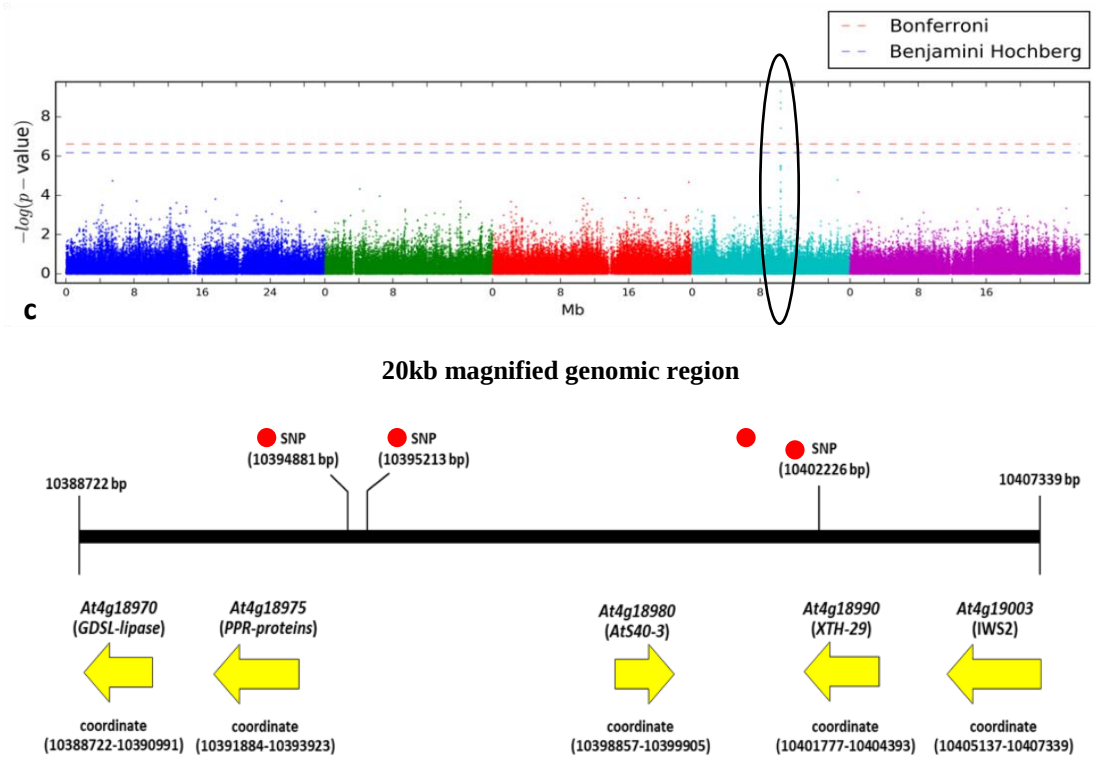
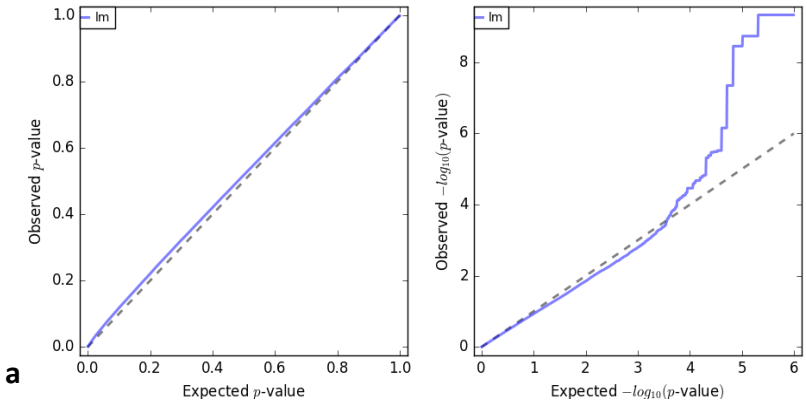


Figure 3.13c: GWAS results for the regulation of female to male sex ratio of cyst nematodes in 148 Arabidopsis accessions. Manhattan plots of GWAS result for the average female to male sex ratio of cyst nematodes per plant using AMM. A region from 10388722bp to 10407339 (18617bp) was magnified to identify the most significant SNP markers in association with causative genes. Four SNP markers positioned at 10394881bp, 10395213bp, 10394611bp, 10402226bp and 10403363bp above the threshold were selected to be truly associated with variation in male nematode numbers. Five candidate genes such as; at4g18970, at4g18975, at4g18980, at4g18990 and at4g19000 were selected in close proximity of the most significant SNPs (c).

At4g18970 and At4g18975 were selected neighboring to SNP (-log10 P-value 9.32, MAF: 0.389, MAC: 51) with un-known annotation positioned at 10394881. Another significant SNP (-log10 P-value 8.73, MAF: 0.38, MAC: 50) positioned at 10395213 flanking 3643bp upstream to the At4g18980 was taken into account for further characterization. Two SNPs flanking the synonymous region of *XTH-29* positioned at 10402226 (-log10 P-value 8.45, MAF: 0.469, MAC: 61) and 10403363 (-log10 P-value 7.43, MAF: 0.469, MAC: 53) were considered potential candidate markers influencing the novel genomic region at chromosome four. However, it can be hypothesized that many of these potential regions contain at least one of the candidate gene influencing the variation in sex ratio of cyst nemattode *H. schachtii*.

Quantile-Quantile plot (Q-Q plot) using LM



Quantile-Quantile plot (Q-Q plot) using AMM

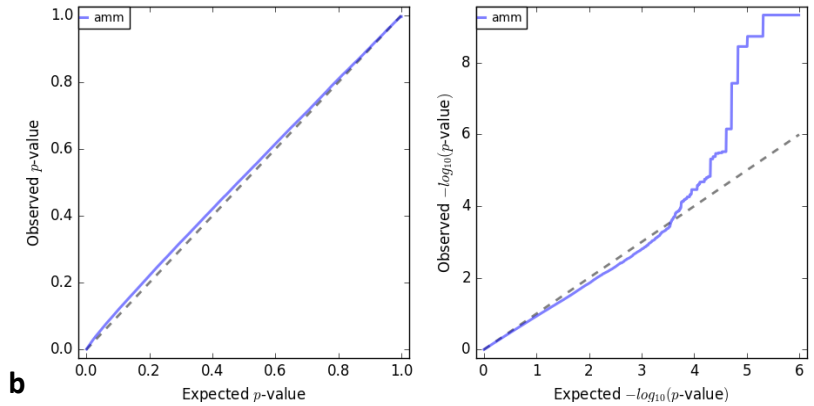


Figure 3.14: Quantile-Quantile plots. Q-Q plots of observed versus expected negative logarithms of the P values of SNPs from LM (a) and AMM (b) for the average male numbers/accession trait. This indicates that false-positive SNP-trait associations due to population structure and kinship were well controlled in the linear regression (LM) and accelerated mixed model (AMM).

Table 3.7: Summary of candidate significant SNPs and the nearest genes from the novel genomic region at chromosome four for the average sex ratio of nematodes per plant trait. Five putative genes with locus At4g18970 (*GDSL-lipase*), At4g18975 (*PPR-proteins*), At4g18980 (*AtS40-3*), At4g18990 (*Endotransglucosylase XTH-29*) and At4g19000 (*IWS2*) neighboring the significant linked SNPs were selected for further analysis. In addition, the strength of SNPs association with the particular trait is also mentioned in terms of -log10 P (value).

Nr.	SNP	Chr	-log10 P	Nearest genes	Gene annotation
SNP 1	10394881	4	9.32	At4g18975	Pentatricopeptide repeat PPR-proteins
SNP 1	10394881	4	9.32	At4g18970	GDSL-lipase/estrases super family
SNP 2	10395213	4	8.73	At4g18980	AtS40-3 a nuclear targeted protein
SNP 3	10394611	4	6.15	At4g18980	AtS40-3 a nuclear targeted protein
SNP 4	10402226	4	8.45	At4g18990	Endotransglucolysase (XTH-29)
SNP 5	10403363	4	7.43	At4g19000	Transcription factor (IWS2)

3.5.3 Selection of extremely variable accessions associated with sex ratio

With remarkable quantitative variations in female to male sex ratio of cyst nematodes, a bulk of twenty extremely variable accessions named; Kondara, Uk-1, Xan-1, Van-0, Ta-0, Ty-0, RRS-7, Ei-2, Gel-1 and Ts-1 (lowly susceptible to cyst nematodes) and Na-1, Zdr-1, Kro-0, Jm-0, Mc-0, Ha-0, Is-0, Sap-0 and Bg-2 (highly susceptible to cyst nematodes) was selected for further detailed investigations. These accessions were exposed to nematodes several times to know the robustness of our infection mechanism and similar results were obtained. These irregularities in female to male nematode numbers indicate that this bulk plays an important role in examining the genetic basis for the QTLs associated with variation in sex ratio. The origin of these accessions with latitude and longitude is given below **(Table 3.8, 9)**.

3.5.4 Alleles associated with lowly and highly susceptible accessions

Alleles associated across lowly and highly susceptible bulks at the most significant SNP (positioned at 10394881bp) locus revealed that allele G was segregated across lowly susceptible accessions, whereas allele A was normally distributed across highly susceptible accessions. Furthermore, although QTLs for male numbers and sex ratio were colocalized but in case of sex ratio there was no abnormalities of extreme accession shifts.

Table 3.8: List of lowly susceptible accessions

Nr.	Accession	Alleles	Country	Latitude	Longitude
1	Kondara	G	Tajikistan	38.48	68.49
2	Uk-1	G	Germany	48.03	7.767
3	Van-0	G	Canada	49.3	-123
4	Xan-1	G	Azerbaijan	38.65	48.79
5	Ty-0	G	UK	56.42	-5.23
6	Ta-0	G	Czech republic	49.5	14.5
7	RRS-7	G	USA	41.56	-86.43
8	Ei-2	G	Germany	50.3	6.3
9	Gel-1	G	Netherland	51.01	5.867
10	Ts-1	G	Spain	41.71	2.93

Table 3.9: List of highly susceptible accessions

Nr.	Accession	Allele	Country	Latitude	Longitude
11	Col-0	A	USA	38.3	-92.3
12	Na-1	A	France	47.5	1.5
13	Kro-0	A	Germany	50.07	8.96
14	Zdr-1	A	Czech republic	49.38	16.25
15	Jm-0	A	Czech republic	49	15
16	Mc-0	A	UK	54.61	-2.3
17	Ha-0	A	Germany	52.37	9.73
18	Is-0	A	Germany	50.5	7.5
19	Sap-0	A	Czech republic	49.49	14.24
20	Bg-2	A	USA	47.64	-122.3

3.5.5 Infection assays of lowly susceptible accessions

Lolwy susceptible accessions were exposed to *H. schachtii* infection to investigate the average numbers of male and female in each accessions and their subsequent female to male sex ratio. Variation in male (a), female (b) and female to male sex ratio (c) is described below.

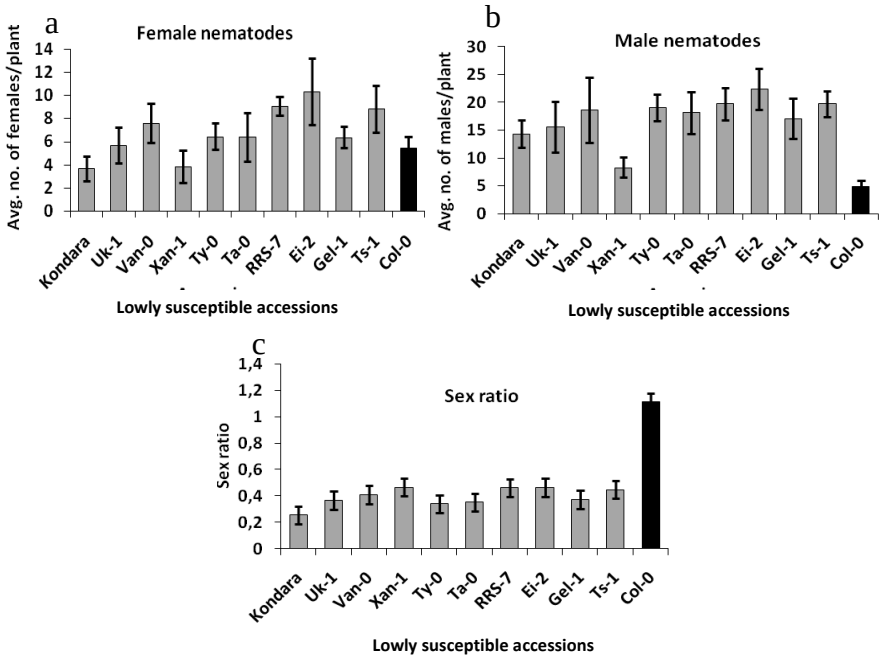


Figure 3.15: Phenotypic variations among lowly susceptible accessions. Variation in cyst nematodes infection among lowly susceptible accessions such as; Kondara, Uk-1, Van-0, Xan-1, Ty-0, Ta-0, RRS-7, Ei-2, Gel-1 and Ts-1. The development of males, females and their altered sex ratios is presented separately; the average number of female nematodes per accession (a), the average number of male nematodes per accession (b) and the average female to male sex ratio per accession (c). Col-0 wild type plants were used as an internal control. Data was analyzed for the significance difference using T-test ($p<0.05$).The astericks indicates the significance difference to wild types.

3.5.6 Infection assays of highly susceptible accessions

Highly susceptible accessions were exposed to *H. schachtii* infection to investigate the average numbers of male and female in each accessions and their subsequent female to male sex ratio. Variation in formation of males (a), females (b) and female to male sex ratio (c) is described below.

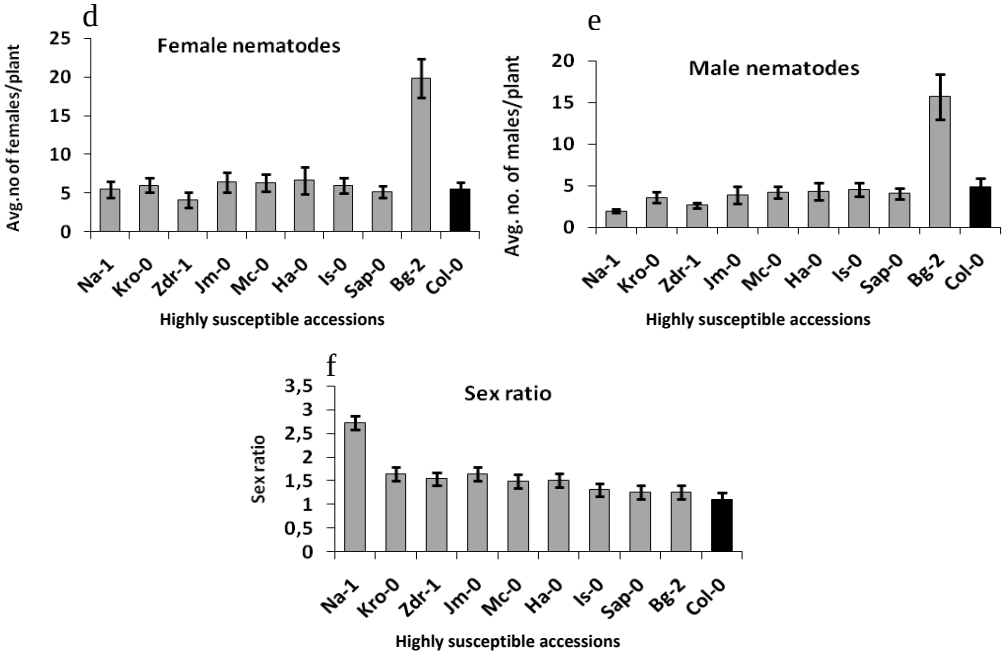


Figure 3.16: Phenotypic variation among highly susceptible accessions. Variation in cyst nematodes infection among highly susceptible accessions such as; Na-1, Kro-0, Zdr-1, Jm-0, Mc-0, Ha-0, Is-0, Sap-0 and Bg-2. The development of the average number of female nematodes per accession (d), the average number of male nematodes per accession (e) and the average female to male sex ratio (f) compared with Col-0 wild type plants. Data was analyzed for the significance difference using T-test ($p<0.05$).

3.5.7 Analysis of genome sequences

To evaluate SNPs and amino acid changes, we used The Arabidopsis 1001 Genome Browser (http://signal.salk.ed). We compared the DNA sequences of twenty lowly and highly susceptible accessions for all the five candidate genes (*GDSL-lipase*, *PPR-proteins*, *AtS40-3*, *XTH-29* and *IWS2*). It turned out that SNPs and amino acid variations were present with the *AtS40-3* gene in lowly susceptible accessions and to some extent with *GDSL-lipases/estrases* in highly susceptible accessions, whereas *PPR-proteins*, *XTH-29* and *IWS2* showed a random distribution instead of consistent polymorphism of SNPs across lowly and highly susceptible accessions. These consistent SNPs in non-synonymous region with *AtS40-3* was found only in lowly susceptible accessions (Kondara, Ta-0, Ty-0, Xan-1, Van-0, RRS-7, Uk-1, Is-0, and Ei-2) **(Fig. 3.17),** whereas no such consistent polymorphism was observed in highly susceptible accessions (Na-1, Kro-0, Zdr-1, Jm-0, Mc-0, Gr-1, Ha-0, Is-0, Sap-0 and Bg-2).

In addition, DNA and protein sequences in lowly and highly susceptible accessions with AtS40-3 were compared and observed with Multiple Sequence Alignment (MSA) Clustal-Omega (https://www.ebi.ac.uk). Polymorphism in the coding region of AtS40-3 includes the substitution of negatively charged Aspartate (D) to Glutamate (G) positioned at 39, positively charged Arganine (R) to Lysine (K) positioned at 46, and Threonine (T) to Lysine (K) positioned at 63. These single nucleotide changes were also examined biologically with simple RT-PCR in some of extremely lowly susceptible accessions as well which supports the evidences of these amino acid substitutions. These SNPs were also found with *PPR-proteins* but there were no differentiating amino acid changes among extreme accessions.

Similarly, with *GDSL-lipase* these substitutions were the other way around and only a few SNPs were found in some highly susceptible accessions such as; Zdr-1, Jm-0, Gr-1, Is-0 and Sap-0), whereas no such changes were observed in lowly susceptible accessions (Kondara, Ta-0, Ty-0, Xan-1, Van-0, RRS-7, Uk-1, Is-0, and Ei-2). Therefore, considering these differentiating SNPs with *AtS40-3* we assume that this might be the major susceptibility regulator influencing the variations in sex ratio of nematodes across Arabidopsis accessions.

3.5.8 SNPs in AtS40-3 protein coding region

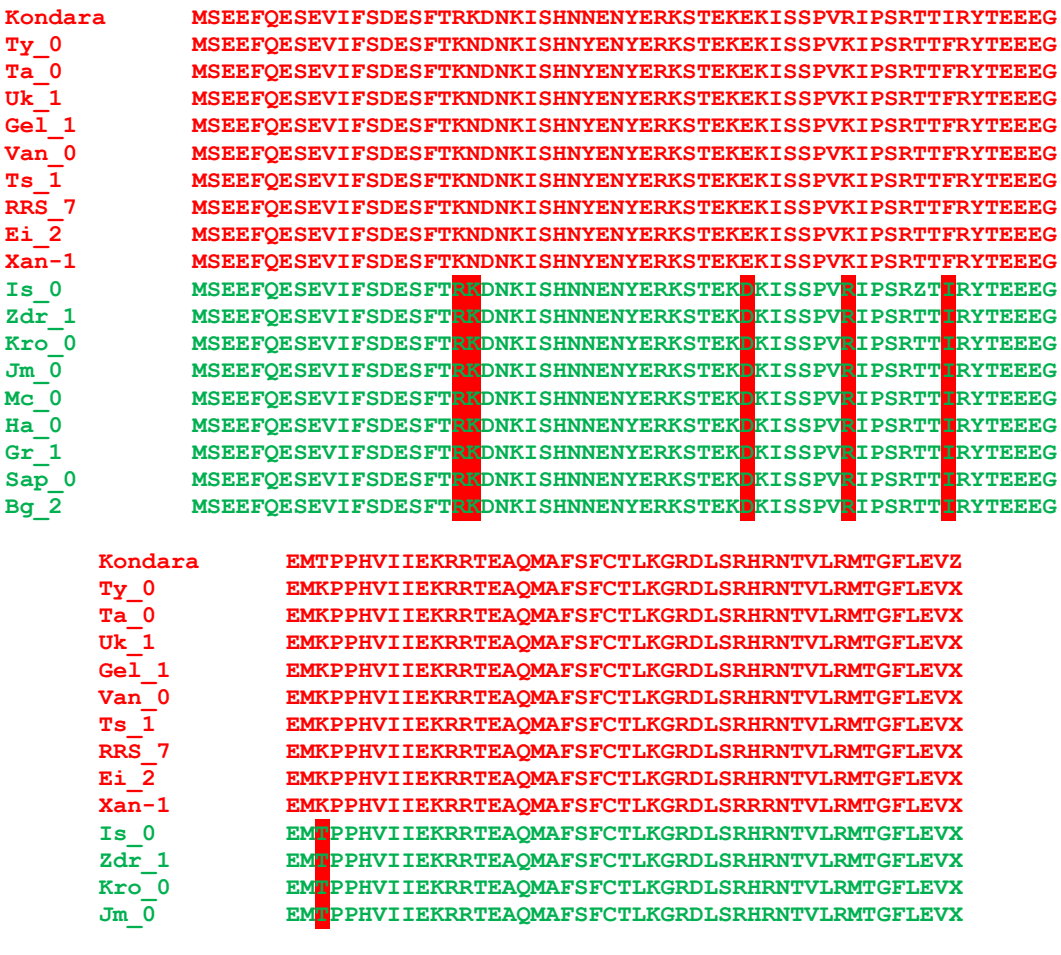


Figure 3.17: *AtS40-3* protein sequence alignment among lowly susceptible (red) and highly susceptible (green) accessions. Red shades indicate amino acid substitutions. Polymorphism was detected among lowly susceptible accessions, whereas no such changes were observed among highly susceptible accessions with *AtS40-3* gene. Amino acid substitutions includes Aspartic acid (D) to Glutamic acid (E) (positioned at 39), Arganine (R) to Lysine (K) (positioned at 46) and Threonine (T) to Lysine (K) (positioned at 63). Protein sequences were aligned with multiple sequences alignment tool "clustal omega" (https://www.ebi.ac.uk).

3.6 Expression analysis of candidate genes in extreme accessions

To analyze the expression of candidate genes in root tissues without infection to nematodes, some accessions Xan-1 and Van-0 (lowly susceptible to cyst nematodes) and Zdr-1 and Kro-0 (highly susceptible to cyst nematodes) were selected and are described below with their geographical origins, altitudes and latitudes (**Table 3.10**).

Table 3.10: List of two lowly and highly susceptible accessions with origins.

Nr.	Lowly susceptible accession	Origin	Latitude	Longitude
1	Xan-1	Azerbaijan	38.65	48.79
2	Van-0	Canada	49.30	-123

Nr.	Highly susceptible accession	Origin	Latitude	Longitude
1	Zdr-1	Czech republic	49.38	16.25
2	Kro-1	Germany	50.07	8.96

We set out to define which genes underly the variation in Arabidopsis susceptibility to *H. schachtii* from this highly associated genomic region. We assumed that the expression of selected genes involved in pathway should be enriched in plant tissues revealing the traits. We then analyzed the expression of four candidate genes (*GDSL-lipase*, *PPR-proteins*, *AtS40-3* and *XTH-29*) in the 20 kb region surrounding the GWA peak by qRT-PCR in some of extreme accessions Xan-1 and Van-0 (lowly susceptible to cyst nematodes) and Zdr-1 and Kro-0 (highly susceptible to cyst nematodes) without nematodes infection with three biological replications.

Interestingly, *PPR proteins* (At4g18975) and *AtS40-3* (At4g18980) showed strongly reduced expression in lowly susceptible accessions, whereas *GDSL-lipase* (At4g18970) and *XTH29* (At4g18990) didn't show differentiating expression compared with the reference accession Col-0 as an intermediate case. Similarly, no remarkable expressions changes were observed among highly susceptible accessions (Zdr-1 and Kro-0). This suggests that AtS40-3 and PPR

genes might be the causal genes underlying the sex ratio of cyst nematodes and that their expression level determines the variation in Arabidopsis susceptibility to *H. schachtii*.

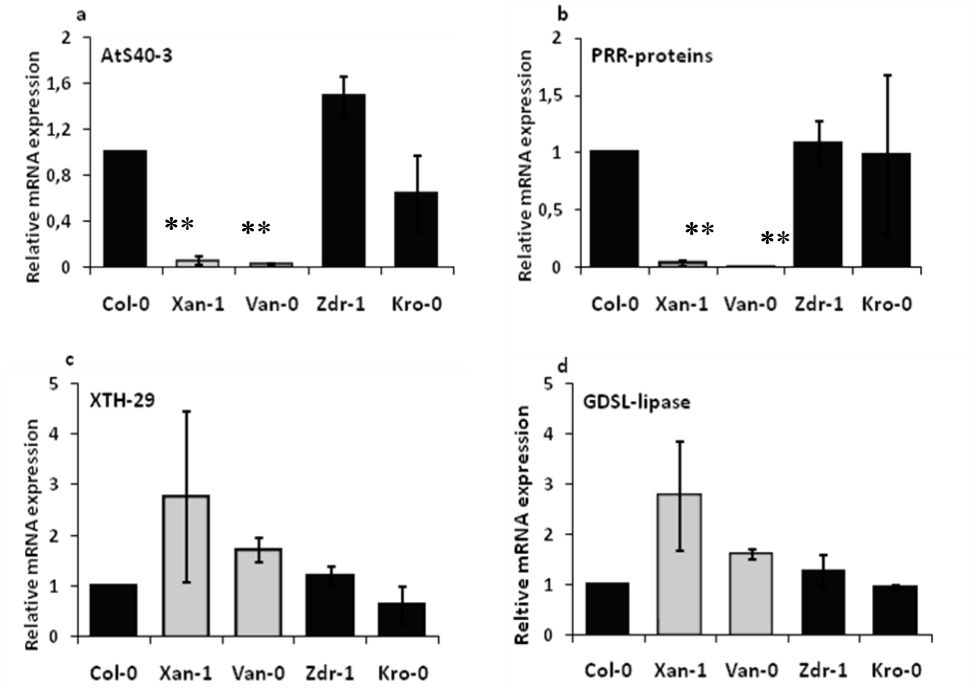


Figure 3.18: The relative mRNA expression analysis for candidate genes. The relative mRNA expression of four selected genes such as; *GDSL lipase*, *PPR-proteins*, *AtS40-3* and *XTH29* in root tissues of lowly susceptible (Xan-1 and Van-0) and highly susceptible (Zdr-1, Kro-0) accessions without nematodes infection. *AtS40-3* and *PPR-proteins* were strongly down-regulated in lowly susceptible accessions Xan-1 and Van-0 compared with Actin as internal reference. Data were analyzed for the significance using T-test ($P \leq 0.05$). The asterisks* indicates the significant differences compared with the wild type.

3.7 AtS40-3 and PPR bi-directional promoter analysis

Considering the variations in candidate genes expressions (*GDSL-lipase*, *PPR-proteins*, *AtS40-3* and *XTH-29*), especially the strong down-regulation of *AtS40-3* (a nuclear targeted protein that modulates senescence) and *PPR proteins* in root tissues of extreme accessions (Xan-1 and Van-0 lowly susceptible to nematodes). We analyzed SNPs in their promoter regions using sequence analysis with Clustal Omega (https://www.ebi.ac.uk) for protein-coding sequences and their corresponding promoter regions. We found that *AtS40-3* and *PPR-proteins* share their bi-directional promoter transcribing these genes in opposite orientation **(Fig. 3.19a)**. In addition, in comparison between lowly and highly susceptible accessions, 253bp downstream to *AtS40-3*, there was a clear pattern of 28-30bp nucleotides deletions in its promoter region among lowly susceptible accessions, while these gaps and deletions were not present in highly susceptible accessions (**Fig. 3.19b)**. Several nucleotide gaps were found

in the regions next to previously described deletions but it was not in consistent order among lowly susceptible accessions.

Taken together, compareed with other candidate genes from the highly associated genomic region (QTL) we could speculate that disturbance in the bi-directional promoter of *AtS40-3* and *PPR-proteins* might be the causes of their strong down-regulations in roots tissues. Therefore, these nucleotide gaps may also be affecting the phenotypic variations in Arabidopsis accessions lowly susceptible to cyst nematodes. In addition, it was interesting to explore DNA binding motifs and their co-factors which were hindering the rna-polymerase bindings to initiate the transcription. However, with the Mulan analysis for the disturbed promoter region (http://mulan.dcode.org/) we found that several TATA box sequences and some other cis-acting elements were also missing among lowly susceptible accessions which were clear convincing findings due to which AtS40-3 and PPR were strongly down-expressed among lowly susceptible accessions.

3.7.1 Schematic presentation of AtS40-3 bi-directional promoter

The schematic presentation of AtS40-3 bi-directional promoter among Uk-1, Kondara, Xan-1, Van-0 and RRS-7 (lowly susceptible bulk) and Kro-0, Zdr-1, Mc-0, Jm-0 and Ha-0 (highly susceptible bulk) is presented below in **Fig. 3.19.**

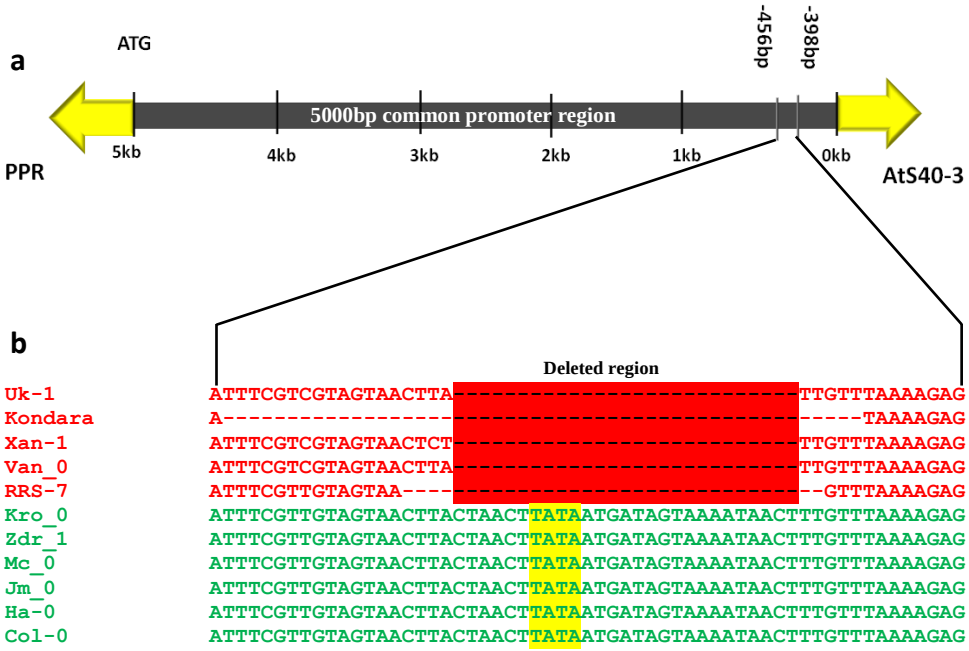


Figure 3.19: Bi-directional promoter analysis of AtS40-3 and PPR genes. A region of 58bp nucleotides indicating major deletion (-398bp- -456bp) in the putative bi-directional promoter of *AtS40-3* and *PPR proteins* (a) in lowly susceptible accessions (Uk-1, Kondara, Xan-1, Van-0 and RRS-7), whereas no such gaps were found among highly susceptible bulks (Kro-0, Zdr-1, Jm-0, Mc-

0, Ha-0 and Col-0) (green) (b). In addition, we can see the missing TATA box among lowly susceptible accessions is also mentioned (yellow shaded).

3.7.2 Predicted *cis*-acting elements in disturbed promoter

Based on the nucleotide gaps among lowly susceptible accessions and reduced expression of *AtS40-3* and *PPR-proteins* among, we hypothesized that these deletions in the promoter region of *AtS40-3* and *PPR-genes* might be one of the reasons for their strong down-regulations in roots tissues. To further investigate this hypothesis, we performed an *in silico* analysis of 28-bp promoter region from a susceptible line Col-0 for the occurrence of important *cis*-acting elements (Chow et al., 2015). Our analysis revealed the presence of some prominent transcription factor binding sequences including a typical TATA box; however, these motifs were deleted in lowly susceptible lines **(Table 3.11)**.

Table 3.11: Promoter sequence (28bp) from Col-0 with predicted putative *cis*-acting elements

Family	Position	Strand	Similarity Score	Hit sequence
AT-Hook	17	+	1	gtaaAATAA
GATA; tify	12	+	1	TGATA
ZF-HD	8	-	1	ATAAT
GATA; tify	13	+	1	GATAG
Trihelix	1	-	0.8	CTAAC
Trihelix	17	-	0.8	GTAAA
Trihelix	22	-	0.8	ATAAC
Motif sequence only	17	-	0.8	GTAAA

3.7.3 AtS40-3 promoter analysis among lowly and highly susceptible accessions

AtS40-3 promoter region from lowly and highly susceptible accessions was selected and analysed single nucleotide polymorphism (SNPs). Lowly susceptible accessions contained several consistent nucleotide deletions, whereas no such nucleotide gaps were observed between highly susceptible accessions. Nucleotide gaps positioned at 10398529-10398563 between lowly susceptible accessions are shown below.

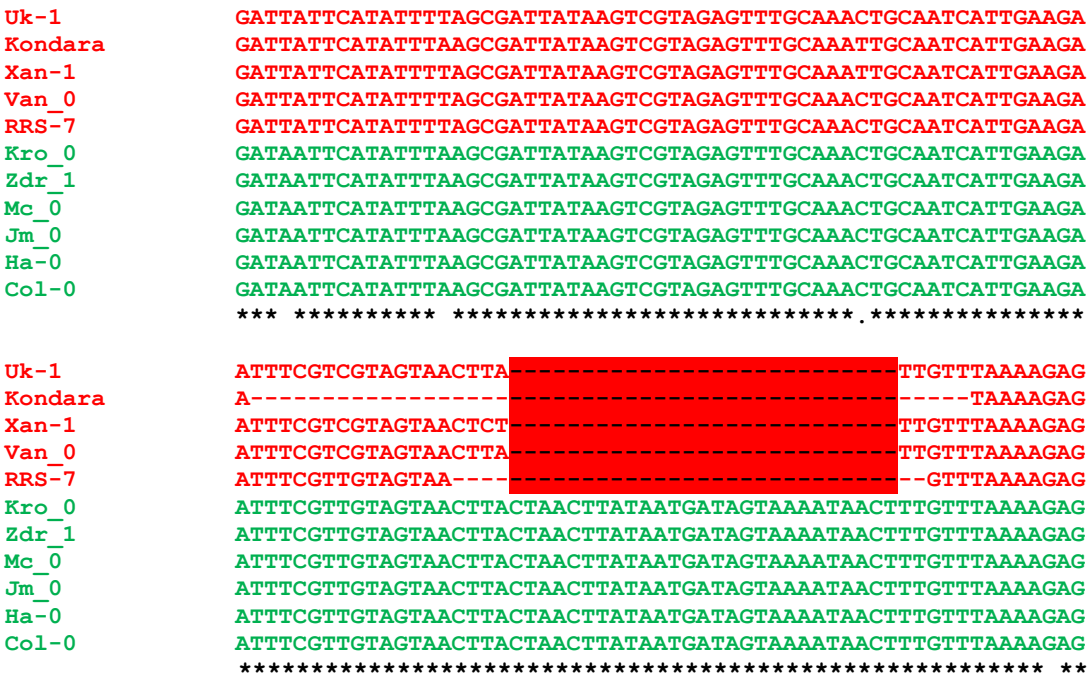


Figure 3.20: Sequence alignment of shared promoter of *AtS40-3* and *PPR-proteins* in four lowly susceptible accessions (Kondara, Uk-1, Xan-1 and Van-0) with nucleotide deletions positioned at 10398529-10398563 (grey shaded) and four highly susceptible (Kro-0, Zdr-1, Jm-0 and Mc-0) accessions without nucleotide deletions. A promoter region of 1500bp upstream to AtS40-3 among lowly and highly susceptible accessions was aligned with CLUSTAL Omega.

3.8 Infiltration of AtS40-3 disturbed promoter in *Nicotiana benthamiana*

To investigate the exact role of the bi-directional promoter of *AtS40-3* and *PPR-proteins*, a region of the 1500bp downstream to *AtS40-3* was cloned from Xan-1 accession (lowly susceptible to cyst nematodes) carrying 28-30 nucleotide deletions parallel to Col-0 wild types (highly susceptible to cyst nematodes) without nucleotide deletions. Later it was fused with green florescent proteins (GFP) and localized into *Nicotiana benthamiana* leaves. The transformation of the construct was verified with simple RT-PCR with GFP primers (**Fig. 3. 22**).

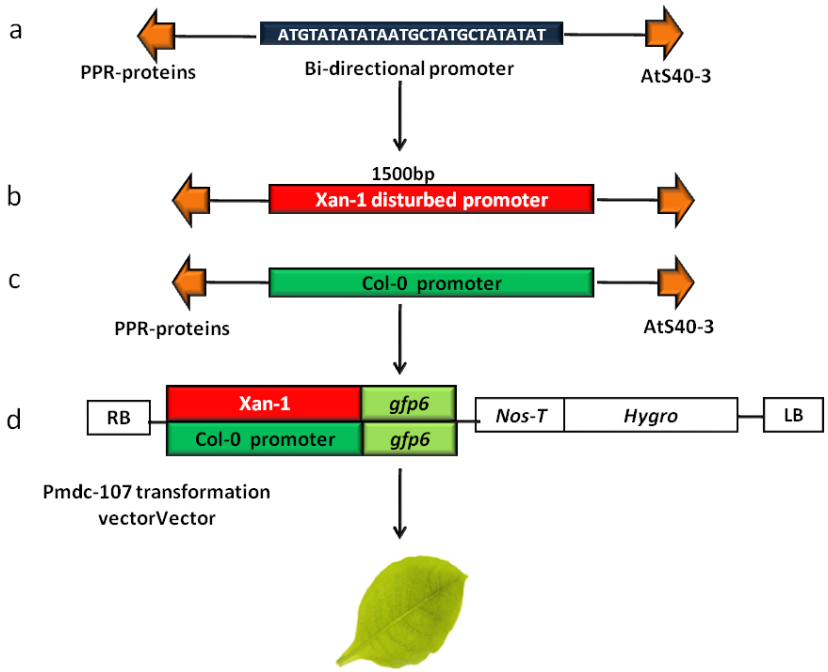


Figure 3.21: Schematic presentation of *AtS40-3* promoter infiltration in *Nicotiana benthamiana*. The transcription regulation of *AtS40-3* and *PPR-proteins* is controlled by their shared bi-directional promoter (a). Disturbed promoter (with nucleotide gaps) of *AtS40-3* (red-colored 1500bp) cloned from Xan-1 accession (b) parallel non-disturbed promoter (without nucleotide gaps) (green-colored 1500bp) cloned from Col-0 accession (c) were fused with GFP (d) and infiltrated into epidermis of *Nicotiana benthamiana* leaves to analyze the GFP expression with different infiltrated constructs. The Hygromycine gene was used as a reference control.

Interestingly it was revealed that there is a significant down-regulation of the GFP gene in lowly susceptible accession Xan-1 compared with Col-0 wild types. The Hygromycine was used as an internal control.

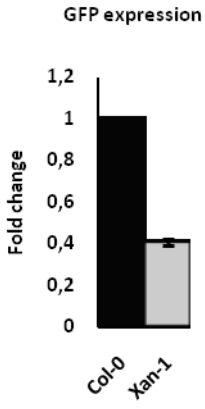


Figure 3.22: Relative mRNA expression of transformed GFP fusion in *Nicotiana benthamiana* leaves. GFP fusion with disturbed promoter from lowly susceptible accession Xan-1 showing the significant reduced expression compared with susceptible accession Col-0. Data were analyzed for the significance using T-test ($P \leq 0.05$). The asterisks * indicates the significant differences to wild type.

Nevertheless, present results of sequence analysis and a detailed molecular characterization of the disturbed promoter of *AtS40-3* gene reveal that variation in nematodes infection in extreme lowly accessions is solely based on the defective promoter of *AtS40-3* with several deletions of nucleotides discovered by GWAS analysis.

3.9 Characterization of loss of function mutants

To assess which of the five candidate genes, *GDSL-lipase* (At4g18970), *PPR-proteins* (At4g18975), *AtS40-3* (At4g18980), *XTH-29* (At4g18990) and *IWS2* (At4g19000) underly variation in selected genomic region, we obtained the T-DNA-inserted loss of function mutant lines for four candidate genes **(Table 2.6).** Due to the unavailability of *PPR* (At4g18975) mutant lines, its functional activity remained unclear. For genotyping and expression analysis, homozygous lines were generated and confirmed for lack of expression via RT-PCR using specific primers flanking insertion (**Table 2.7, 8**). Confirmed knockout lines were further evaluated for their respons to nematode via infection assays.

3.9.1 *GDSL-lipase* Infection assays

To test our hypothesis that *GDSL-lipase* is involved in the regulation of cyst nematode sex ratio, loss of function of *GDSL-lipase* mutant plants was tested by exposing them to cyst nematode *H. schachtii*. As a result, we found significant increase in the number of male and decrease in the number of female nematodes per plant and subsequent a prominent change in female to male sex ratio. These convincing results indicate that *GDSL-lipase* is one of the important positive regulators of Arabidopsis susceptibility to *H. schachtii*. In addition to male and female numbers we also measured female nematodes and their syncytia sizes but no significant differences were observed compared with Col-0 wild types.

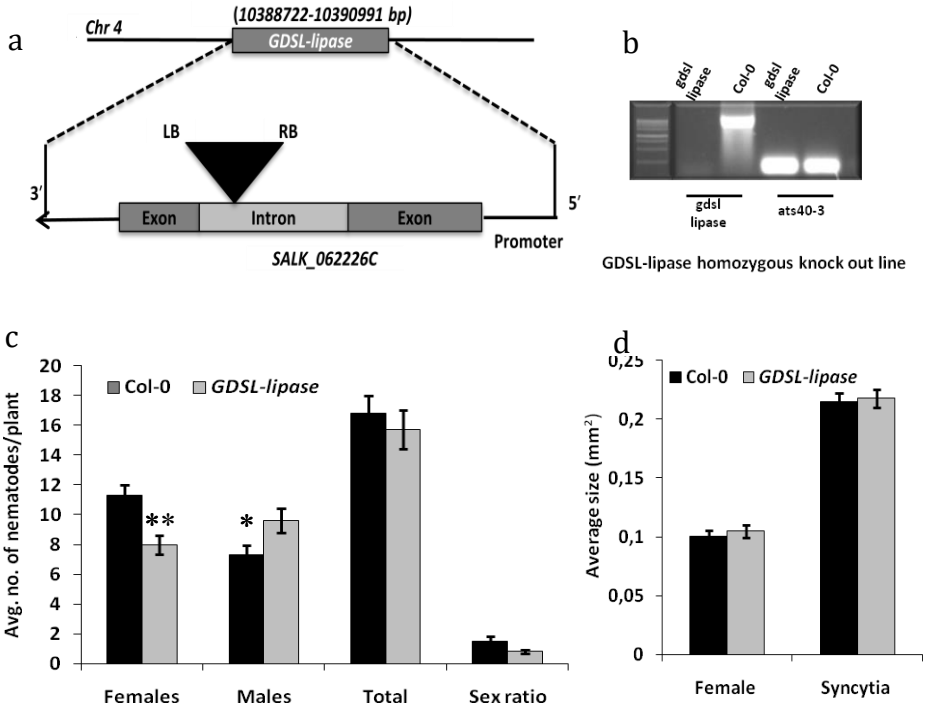


Figure 3.23: Cyst nematode infection assays in *GDSL-lipase* mutant plants. (a) A salk_062226C knock out *GDSL-lipase* mutant line with an insertion in exon is indicated with red triangle. *GDSL-lipase* gene region (positioned at 10388722-10390991) at chromosome four is also mentioned (grey). (b) Confirmation of true homozygous lines with no *GDSL-lipase* expression compared with Col-0 wild type. (c) Quantitative phenotyping of the knock out lines with the average number of female and male nematodes per plant compared with Col-0 wild type. (d) Comparison of female sizes and associated syncytia among *GDSL-lipase* and Col-0 wild type. Data was analyzed for the significance difference using T-test ($p<0.05$).

3.9.2 *AtS40-3* Infection assays

To test our hypothesis that *AtS40-3* is involved in the regulation of nematodes sex ratio the loss of function of *AtS40-3* mutants were tested exposing them to *H. schachtii* infections. As a result, we found significant increase in the number of male and decrease in the number of female nematodes per plant and consequently a prominent change in female to male sex ratio. These convincing results indicate that *AtS40-3* protein is one of the important positive regulators of Arabidopsis susceptibility to *H. schachtii*. In addition to male and female numbers we also measured the syncytia and associated female sizes but no promising differences were observed compared with Col-0 wild types.

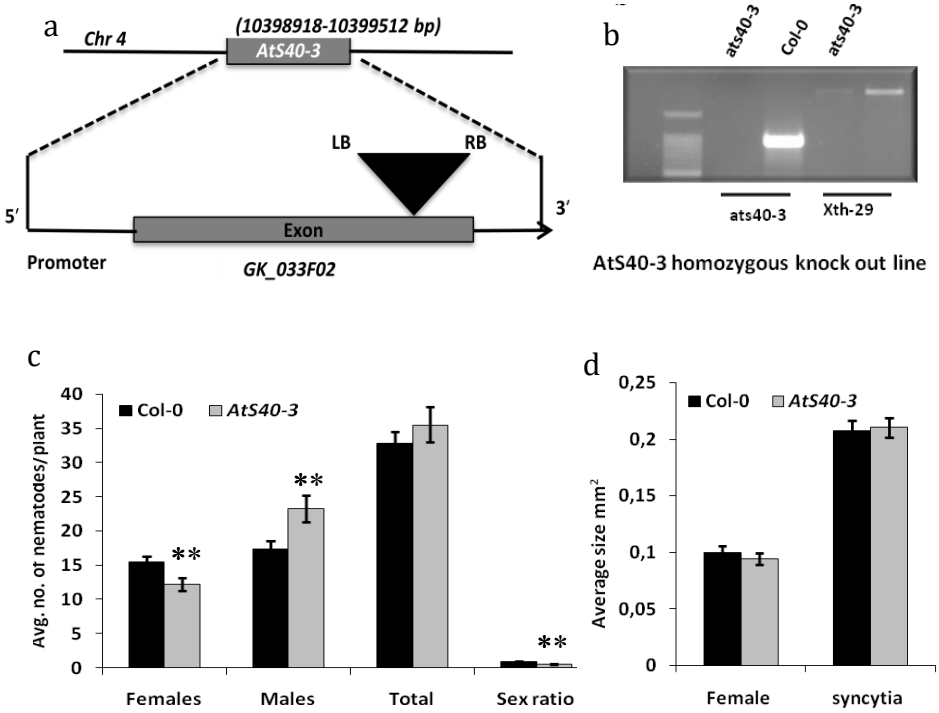


Figure 3.24: Cyst nematode infection assays in *AtS40-3* mutant plants. (a) A GK_033F02knock out *AtS40-3* mutant line with an insertion in exon is indicated with red triangle. Chromosomal position (10398918-10399512) of *AtS40-3* is also mentioned (grey). (b) Confirmation of true homozygous lines with no *AtS40-3* expression compared with Col-0 wild type. (c) quantitative phenotyping of the knock out lines with average number of female and male nematodes per plant compared with Col-0 wild type. (d) comparison of females and associated syncytia among *AtS40-3* and Col-0 wild type. Data was analyzed for the significance difference using T-test ($p<0.05$).

3.9.3 *XTH-29* infection assays

Mutation in the third gene from the novel QTL *Xyloglucan endotransglucosylase /hydrolase-29* in association with SNP positioned 10402226 with -log10 P value 7.45 with regards to infection with cyst nematode *H. schachtii*, showed significant increased female numbers and total number of nematodes per plant compared with Col-0 wild type plants. Comparison of syncytia sizes and associated syncytia with Col-0 wild types did not show any promising differences. However, these results showing increased susceptibility indicate that *XTH-29* enzymes might play a positive role in resisting the cyst nematode invasions and parasitism.

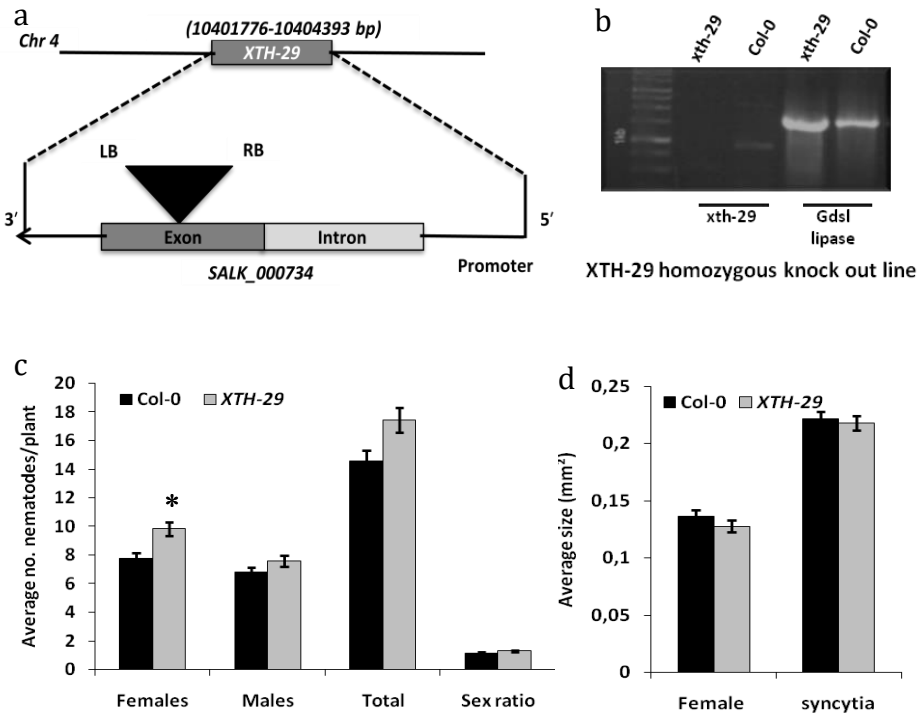


Figure 3.25: Cyst nematode infection assays in *XTH-29* mutant plants. (a) A SALK_000734 knock out *XTH29* mutant line with an insertion in exon is indicated with red triangle. The chromosomal position (10401776-10404393) of *XTH29* is also mentioned (grey). (b) Confirmation of true homozygous lines with no *XTH29* expression compared with Col-0 wild type. (c) Quantitative phenotyping of the knock out lines with the average number of female and male nematodes per plant compared with Col-0 wild type. (d) Comparison of females and associated syncytia among *XTH29* and Col-0 wild type. Data was analyzed for the significance difference using T-test ($p<0.05$).

3.9.4 *IWS2* Infection assays

To test our hypothesis that *IWS2* is the involved in regulation of cyst nematode sex ratio, the loss of function of *IWS2* mutant plants were tested by exposing them to cyst nematode *H. schachtii*. In this case, the new mutant didn't show any remarkable increase or decrease in the average number of nematodes. In addition, we also measured the average size of female and associated syncytia after 14 days post infection (dpi) which did not show any notable changes in comparison with wild types.

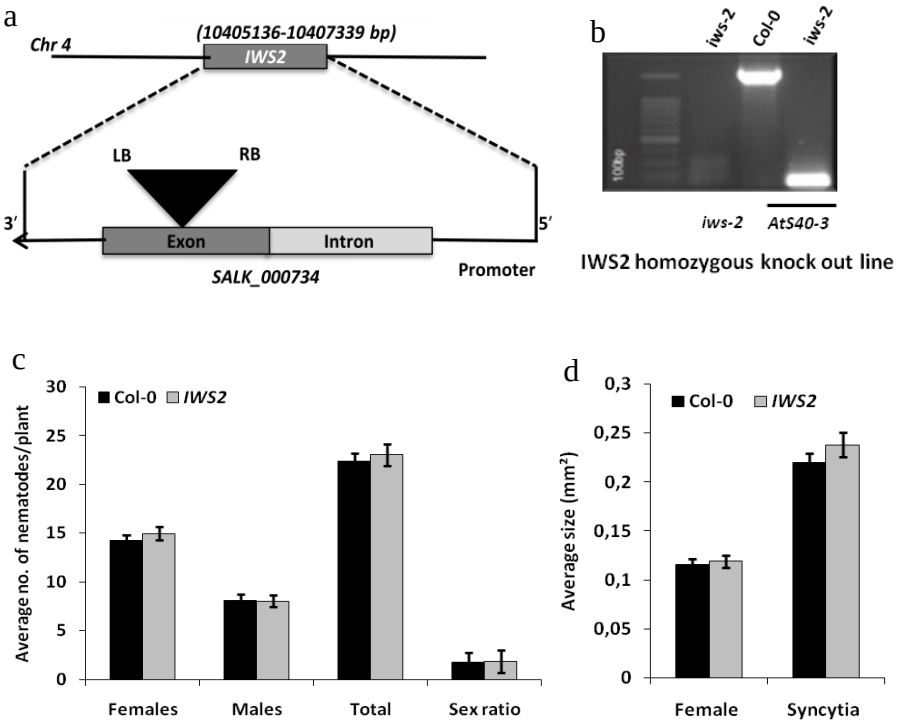


Figure 3.26: Cyst nematode infection assays in *IWS2* mutant plants. (a) A SALK_000734 knock out *IWS2* mutant line with an insertion in exon (yellow block) and insertion with left and right borders is indicated with red triangle. In addition, chromosomal position (10405136-10407339) is also mentioned (in grey). (b) Confirmation of true homozygous lines with no *IWS2* expression compared with Col-0 wild type. (c) Quantitative phenotyping of the knock out lines with the average number of female and male nematodes per plant compared with Col-0 wild type. (d) Comparison of females and associated syncytia among *IWS2* and Col-0 wild type. Data was analyzed for the significance difference using T-test ($p<0.05$).

Taken together these results indicate that GDSL-lipase, AtS40-3 and XTH-29 genes play an additive role in influencing the novel quantitative trait loci (QTL) and controlling the variations in sex ratio of cyst nematode.

3.10 *Meloidogyne* infection assays

The knock out mutant of *AtS40-3* was subjected to root-knot nematodes *Meloidogyne incognita* to analyze their parasitism activity. After three weeks of post infection, the average

gall numbers were investigated and the results revealed that *AtS40-3* was not involved in influencing the average number of galls compared with Col-0 wild type plants. Because there were not differences in the average gall numbers and their corresponding sizes which indicates that *AtS40-3* might only be specific to *H. schachtii* parasitism.

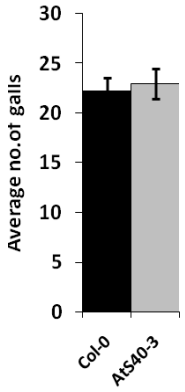


Figure 3.27: *Meloidogyne incognita* infection assay in *AtS40-3* mutant plants. There was no difference in the average number of galls per plant compared with Col-0 wild type plants. Data was analyzed for significance using T-test ($P \leq 0.05$) and n.s. indicates the non-significance.

Chapter 4

4.0 Discussion

In general, changes in habitation causes individuals in a population to struggle for existence, thus regulating their capabilities to cope with new challenges (Pullin, 2002). On the other hand, genetic diversity, which is defined as the frequency of different evolutionary heritable alleles within a population, plays an imperative role in survival and adaptability of plants to new environmental conditions (Frankham, 2005). Natural genetic variation within individuals has been the subject matter of evolutionary genetics and it is considered a source for evolutionary change and for adaptive potential of a species to varying environmental conditions. Exploiting natural variation leads to introduce novel alleles against various biotic and abiotic stresses that improve productivity, adaptation and nutritional value of crop plants. This phenotypic variation in crop plants for various traits is due to underlying genetic complexity with allelic effects that are sensitive to environmental conditions.

Some of the phenotypic differences are due to single gene allelic variants; however, mostly natural variations are quantitative and multigenic, which are determined by molecular polymorphism at multiple loci. These are referred as quantitative trait loci (QTLs) and quantitative trait genes (QTGs). Various methods of genome mapping underlying quantitative variation andtheir genetic architecture have developed few decades ago in crops. Currently, genomic resources have been introduced for several crop plants such as; barley (*Hordeum vulgare*), rice (*Oryza sativa*), and tomato (*Solanum lycopersicum*), enabling the identification of genes and single nucleotide polymorphisms (SNPs) underlying QTL involved in various agronomic traits (Doebley et al., 2006; Sang, 2009). In this study, a natural population of Arabidopsis has been exposed to cyst nematodes to investigate genetic variants involved in formation of male and female nematodes.

4.1. Identification of genetic variants in Arabidopsis using association mapping

The identification of genes that responsible for phenotypic variation has become feasible with several genome mapping approaches such as; traditional linkage or QTL-mapping and association mapping or LD-mapping (Clark, 2007). Linkage-mapping depends upon genetic recombination during the development of mapping populations and has the disadvantages of relatively low mapping resolution, low speed and low allelic richness. In contrary to linkage mapping, association mapping takes the advantage of more genetic recombination, thus providing higher mapping resolutions and greater allele numbers. Natural population based association mapping involves following steps; i) selection of natural population with wide

coverage of genetic diversity, ii) measurement of phenotypic characteristics, preferably, in different environmental conditions with multiple replicates, iii) genotyping a mapping population with molecular markers, iv) quantification of extent of LD, v) estimation of population structure and kinship, vi) with the application of appropriate statistical approaches (MLM, LM, AMM), phenotype-genotype correlation reveals "marker tags" positioned within close proximity of targeted trait of interest and vii) positional cloning and annotation of tagged loci for potential biological function.

Genome-wide association mapping has successfully been implemented in diverse crops for complex traits and functional investigations of evolutionary changes. Barley, rice, wheat, Soybean, millet and cowpea are considered well suited crops for association mapping strategies to identify QTLs for numerous agronomic traits. As far as applications of this approach in plant-nematode interaction are concerned, resistance and susceptibility traits in Soybean (Zhang et al., 2016), rice (Dimpka et al., 2016), wheat (Pariyar et al., 2016) and cowpea (Santos et al., 2018) crops have been studies recently.

4.2. *Arabidopsis thaliana* model organism for association mapping

The model plant *Arabidopsis thaliana* has widely been used to investigate the genetic for various quantitative trait variations. Advanced genomic and biological resources have accelerated the potential of Arabidopsis for studies of natural variation in recent years. The availability of hundreds of natural inbred lines (NIL) has increased the genetic resources to isolate the genetic variants with higher resolution in different environmental conditions. Notably, the molecular analysis of genetic variation has not only led to the correlation of allelic variation of known genes but also to the uncovering of novel genes. In Arabidopsis, genetic variants for several traits including salt tolerance (Baxter et al., 2010), shade avoidance (Filiault et al., 2012), 107 different phenotypes (Atwell et al., 2010), flowering traits (Li et al., 2010) and glucosinolate levels (Chan et al., 2011) have successfully been dissected with association mapping approach. Here, using GWAS, we made an effort to explore the genetic architecture and identify the susceptibility regulators in a diversified natural population of 148 accessions against cyst nematode *H. schachtii*.

4.3. Traits used to investigate genetic allelic variants

Root-knot and cyst nematodes are both biotrophic pathogens of numerous plant species with sedentary modes of parasitism. They cause dramatic changes in the morphology and physiology of their hosts. Cyst nematodes invade the plant roots and move intracellularly to reach the vascular cylinder and thus forming permanent syncytial feeding sites (Grundler et

al., 1998). In contrast to cyst nematodes, root-knot nematodes move intercellularly after penetrating the root, migrating down the plant cortex towards the root tip.The juveniles enter the base of the vascular cylinder, migrate up the root and establish their feeding sites termed as “multinucleate giant cells” (Williamson and Hussey, 1996).

The beet cyst nematode *H. schachtii* is a pathogen of sugar beet (*Beta vulgaris*) but can also complete its life cycle on *Arabidopsis* roots growing on agar plates under sterile conditions. Cyst nematodes are sexually dimorphic and sex differentiation is greatly influenced by different environmental conditions. Under favorable conditions with sufficient nutrition supply more females are produced, whereas under adverse conditions more male nematodes are produced. In this study, we challenged the selected Arabidopsis accessions with *H. schachtii* nematodes through standardized infection assays. After 10-12 dpi they could easily be differentiated into male and females. We analyzed the average male and female numbers in each accession and also their subsequent female to male sex ratio in each accession. As expected, a high natural variation was observed in male and female numbers and their sex ratio in all accessions.

The aim of the study was to investigate the genetic architecture for these quantitative variations in male and female number of nematodes using genome-wide studies (GWAS). GWAS has been widely used to study genetic basis for various trait in Arabidopsis, but the response to cyst nematode *H. schachtii* has never been investigated. The detailed description of these phenotypic variations for each trait is given in the next paragraphs.

4.4. Trait 1: Natural variation in female cyst nematodes in Arabidopsis

Cyst nematodes and root-knot nematodes form complex interactions with their host plants and development of their feeding sites solely based on these interactions. Root-knot nematodes form feeding sites are composed of giant cells (Bird, 1961). The syncytial feeding sites formed by cyst nematode are structurally distinct from giant cell complex. Cyst nematode feeding sites are initiated by single pericycle or procambium cell, whereas feeding sites formed by root-knot nematodes are composed of giant cells derived from 5-7 parenchymal cells. Cyst nematodes rely on plant-derived syncytia as their nutritional source; therefore, unfavorable environmental conditions or weak host plants with least nutrition to develop a fully functional syncytial feeding sites can influence/prevent the development of adult female nematodes (Sobzak et al., 1997).

Female nematodes obtain their food nutrition from syncytia with the help of their stylet and associated feeding tubes. Female cyst nematodes have high nutritional demand and according to estimation the amount of solutes taken up is equal to four times the syncytium volume per day (Sijmons et al., 1991). Females of *H. schachtii* requires 29 times more nutrition than males (Muller et al., 1981) and females induced syncytia are 3.7 times larger than male associated syncytia. Based on previous studies that female nematodes development is influenced by nutritional and genotype-host, we exposed a natural population of Arabidopsis to cyst nematodes with the aim to find out the genetic alleles effecting their development. Infection assays revealed a high variation in female nematode developments in each natural accession.

Based on variation in female nematodes, all genotypes were categorized into three groups; lowly susceptibile, moderately susceptibile and highly susceptible. Accordingly, 39.1% were lowly susceptible, 28% were moderately susceptible and 31.4% were highly susceptible. However, to investigate variation in the average female numbers across the entire population more intensely, we selected twenty extremely variable bulks showing low susceptibility (Nok-3, Dr-0, Rmx-AO2, Ms-0, Kz-9, Mt-0, Kondara, Wa-1, Stw-0, and Xan-1) and high susceptibility (Tu-0, Ca-0, Bsch-0, Sei-0, Bu-0, Aa-0, Jl-3, and Nc-1). We see some lowly susceptible accessions such as; Dr-0, Kz-9 and Wa-1contained only 1.9, 2.5 and 2.6 females respectively, whereas highly susceptible accessions such as; Ba-1, Gd-1 and Nc-1 possessed 17, 14 and 13 females respectively. Variation in female numbers ranged from 1.9 - 13 was remarkable. Approximately, similar variation in female numbers was seen across all 148 accessions. These notable natural variations in female nematodes showed the potential perspectives of their genetic studies in Arabidopsis. Therefore, genome-wide association mapping was conducted to investigate the genetic basis for these phenotypic variations.

4.4.1. QTLs associated with cyst nematode variation

The identification of resistance resources using natural or experimental populations have been demonstrated to be the most practical and efficient manner to control cyst nematodes. In literature, bi-parental crop populations have been commonly used to identify and characterize QTLs conferring resistance to cyst nematodes. In a study, 553 soybean plant introductions were presented to cyst nematode *Heterodera glycine* to identify the QTLs linked to resistance (Vuong et al., 2015). Over 45,000 SNP (SoySNP50K iSelect BeadChip) markers were used to for analysis and GWAS identified 14 QTLs comprising 60 SNPs across the different chromosomes. Recently, GWAS was performed over 161 winter wheat accessions using 90K

iSelect SNPs chip. GWAS identified 11 novel QTLs conferring resistance in wheat against *Heterodera filipjevi* on different chromosomes including eight QTLs residing genes known to be involved in plant-pathogen interaction (Pariyar et al., 2016). These studies show that GWAS can be employed as an effective strategy for identifying allelic variants of complex traits in crops and narrows down the associated genomic regions, which facilitates the positional cloning of the causal genes. More recently, Santos et al., (2018) identified resistant QTLs, designated as QRk-vu9.1, against root-knot nematodes *Meloidogyne javanica* using recombinant inbred lines in cowpea. Mostly identified genes belong to TIR-NBS-LRR family of resistance genes. However, detailed functional characterization on RK-mediated resistance remained unclear.

In present study, GWAS was performed over 148 Arabidopsis accessions to narrow down the susceptibility resources in response to cyst nematode *H. schachtii*. Over 250,000 SNP markers were used for analysis. GWAS analysis revealed a significant QTL conferring variation in female nematodes formation with significance SNP -log10 P (value) 6.01 at chromosome five. The most significant SNP positioned at 2392347bp with its alleles (A/G) and high LD with neighboring SNPs supported its association with phenotypic trait. Neighboring to significant SNP (at 2392347bp), 20kb region was further explored and we found the residing candidates genes; Glycine-rich proteins family comprising six putative genes; at5g07540 (*AtGRP-16*), at5g07530 (*AtGRP-17*), at5g07550 (*AtGRP-19*), at5g07560 (*AtGRP-20*) and at5g07570 (glycine-oleosin protein) **(Table 3.1).** Based on association of the most significant SNP marker locus (2392347bp), it was assumed that one of these candidate genes might be involved in regulating the female nematodes development in Arabidopsis natural population.

4.4.2. Glycine-rich proteins (GRPs) in *Arabidopsis thaliana*

Glycine-rich proteins (GRPs) have been found in the cell wall of many plants with high structural diversity and in addition to extensions and proline-rich proteins (PRPs) are considered a third group of structural protein components. This diverse family is characterized by repetitive glycine-rich domains, frequently with GGGX, GGXXXGG or GXGX repeats (Sachetto-Martins et al., 2000). However, amino terminal signal peptides, cystein-rich domains (Rohde et al., 1990), cold shock patterns (Kingsley and Palis 1994) and RNA-binding motifs (Dreyfuss et al., 1993) have also been identified.

Previous studies show that genes encoding glycine-rich proteins (GRPs) are induced by abiotic and biotic stresses such as wounding (Showalter et al., 1991), osmotic stress (Gómez

et al., 1988) abscisic acid (de Oliveiraet al. 1990), salicylic acid (Hammond-Kosack and Jones 1996), and viral (Linthorst et al. 1990) and fungal infections (Molina et al. 1997). Moreover, cell wall or membrane-associated GRPs are involved in multiple functions in plant defense responses; interacting with kinases, blocking virus-movement, and modulating transcription of defense genes (Park et al., 2001; Ueki andCitovsky, 2002; Tao et al., 2006; Kim et al., 2007; Park et al., 2008; Nicaise et al., 2013; Kim and Hwang, 2015). Pathogens associated molecular patterns (PAMPs) such as flagellin (flg-22), harpin (HrpZ), necrosis-inducing proteins (NPP) and lipopolysaccharide (LPS) can also induce GRPs expression. The bacterial PAMP, HrpZ protein is glycine-rich and elicits hypersensitive response at the infection site (Choi et al., 2013).

Recently a reverse genetic approach revealed that Arabidopsis GRPs could be implicated in secondary cell wall formations and this analysis showed that GRPs are involved in protoxylem synthesis (Yokoyama and Nishitani, 2006). Likewise, *AtGRP-9* is considered involved in lignin biosynthesis in response to salt stress as a result of its interaction with *AtCAD5* (cinnamyl alcohol dehydrogenase) (Chen et al., 2007). In a study, *NtCIGI* gene encoding a cadmium induced glycine-rich protein (cdGRP) identified to be expressed in cell walls of plant vascular tissues. In addition, constitutive expression of cdiGRP inhibited the systemic transport of turnip vein-clearing virus (TVCV), whereas suppression of cdiGRP production allowed TVCV movement in the presence of cadmium. It was found that this inhibition of TVCV was actually due to callose depositions by *NtCIGI*. However, this finding proposed a structural role of GRPs that may function to control plant viral systemic movement (Ueki and Citovsky, 2002).

Another important role of GRPs in plant defense was determined from the interactions of Arabidopsis and *Pseudomonase syringae*. This bacterium uses a type III effector HopU1 to inject the secretions in plants. This type III effector is a mono-ADP-ribosyltransferase (HopU1-His) that possess RNA-binding proteins with RNA-recognition motifs (RRMs) as its substrate. Loss of function mutants of *AtGRP7* were more susceptible than wild-type control plants to *P. syringae* and thus suggesting that *AtGRP7* plays a role in innate immunity. Detailed studies showed that ADP-ribosylation of GRP7 by HopU1 required two arginines within the RRM, indicating that this modification may interfere with GRP7's ability to bind RNA. Besides, GRPs role in plant defense response, its implications as extracellular ligands of kinase proteins has also been observed. A few genes encoding GRPs such as *AtGRP-3* was induced to interact with cell wall associated receptor kinase-1 (*WAK1*) signaling pathway

where its binding is necessary for the activation of KAPP (kinase-associated phosphatase) (Park et al., 2001).

4.4.3. GRPs induced in nematodes associated syncytial feeding sites

Root-knot and cyst nematodes are obligatory plant parasites and their nutritional balanced activities are obtained only from the cytoplasm of living root cells. Once invading the roots, they induce specialized feeding structures within the vascular cylinder. Root-knot nematodes feed from multinucleate giant cells developed by the expansion of cambial cells within the differentiating vascular cylinder, whereas cyst nematodes get their food from syncytium formed by the partial dissolution of cell walls of the neighboring cells (Jung et al., 1998). This formation of nematode feeding sites (NFS) as a result of complex plant-nematode interactions leads to alter the gene expressions of host pants (Sijmons et al., 1994; Williamson and Hussey, 1996).

Studies indicate that glycine-rich protein genes (*Atgrp-6, Atgrp-7* and *Atgrp-8*) are predominantly expressed in the anthers and more specifically in the tapetum layer. Tapetal cells are responsible for nutrition of developing pollen grains and show some functional similarities to nematode feeding sites (NFS) induced in plant roots by sedentary parasitic nematodes (Mascarenhas, 1975). Taken together, previous assumptions revealed that in both systems, cells are multinucleate (D'Amato, 1984), have cell wall ingrowths, and act as transfer cells (Pate and Gunning, 1972). However, in a study, to analyze the expression of GRP genes, transformed Arabidopsis plants containing *Atgrp-7* promoter and glucuronidase (gus) fusion were inoculated with cyst nematodes *H. schachtii* and root-knot nematodes *M. incognita* after different time points. After 3dpi this *Atgrp-gus* fusion transcripts revealed a significant increased expression in the feeding sites of both cyst and root-knot nematodes **(Fig. 4.1)**.

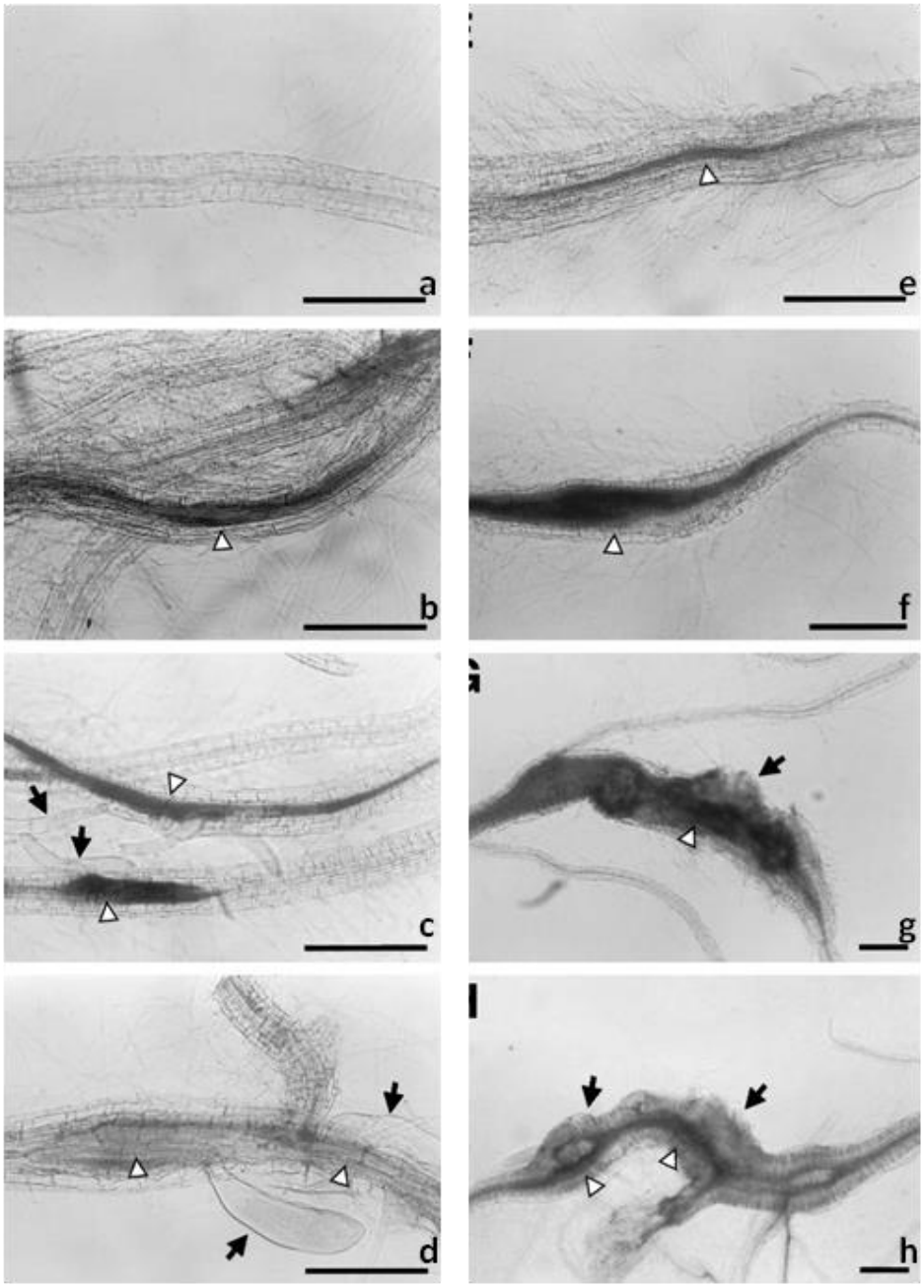


Figure 4.1: *Atgrp*7::GUS expressions upon nematode infections in Arabidopsis plants. Control root without nematode infection (a). Gus expression after 3dpi (b), 7dpi (c) and 14dpi (d) by *H. schachtii*. Gus staining after 3dpi (e), 7dpi (f) and 14dpi (g) and 28dpi (h) by *M. incognita*. Arrows and arrowheads indicate nematodes and feeding sites (Karimi et al., 2002).

In contrary to tapetum specific genes/pollenins, the precise role of *Atgrp* genes in nematodes feeding structures is not fully understood (Frandsen et al., 2001; Sachetto-Martins et al., 2000). Oleosin promoters have been shown to be triggered in young root tips and this assessment was correlated with the presence of lipid bodies (Murphy et al., 2001). Besides, transformed Arabidopsis plants with *Atgrp* promoter-gus transcripts were also induced in

some root tips. There are evidences that numerous lipid bodies are present in syncytia, therefore, taptum genes/pollenin might play a role in lipid stabilization of these nematode feeding sites (Sachetto-Martins et al., 1995). In the present study, Arabidopsis accessions were challenged with cyst nematodes *H. schachtii* and marker-trait correlations identified the associated QTLs residing similar genes (*Atgrp-17*, *Atgrp-18*, *Atgrp-19* and *Atgrp-20*) coding for GRPs as previously described (Karimi et al., 2002). Their cell wall interacting partners/recepters and signaling mechanism against cyst nematode *H. schachtii* is still unknown. However, we assume that these extracellularly localized proteins are essential for female syncytia formations and may play an important role in nutritional stability for growth and development of female nematodes.

4.4.4. Allelic polymorphism at significant SNP locus between extreme accessions

Finally, two extreme bulks showing lowest (Nok-3, Dr-0, Rmx-AO2, Ms-0, Kz-9, Mt-0, Kondara, Wa-1, Stw-0, and Xan-1) and highest (Tu-0, Ca-0, Bsch-0, Sei-0, Bu-0, Aa-0, Jl-3, and Nc-1) susceptibility parameters were selected based on the allelic polymorphism at the associated SNP locus for analysis of candidate genes as well as their functional characterizations. Candidate SNP marker locus revealed that allele A was segregating across lowly susceptible accessions, whereas allele G was normally distributed across highly susceptible accessions. This consistent allelic segregation indicates that this marker locus might be associated with quantitative variation in female nematodes.

We analyzed the allelic polymorphism for each candidate gene in their protein coding, non-coding and promoter regions. SNPs were detected in the protein-coding sequences betweenlow and high susceptibility accessions but this polymorphism was not consistent between accessions.Therefore, due to non-consistent amino acid changes at associated SNP locus among extreme lowly and highly susceptible accessions their functional characterization remains unclear.

Since the presence of female cyst nematodes determines reproduction and infestation level of soil, they are very important in agriculture crops. Female nematodes reproduce hundreds of eggs and thus causing an increase in population dynamics of cyst nematodes to many generations. Cultural practices such as; resistant crop cultivars which suppress the female development and reproduction are mostly used to control cyst nematodes. However, genetic mechanism behind is indefinite. Here, natural resources provide an opportunity for identification of QTLs involved in formation of female nematodes and thus enabling breeders to develop genetic mutations to suppress their formation processes. Our study also supports

the role of different host-plants on development of female cyst nematodes. Identified host factors may provide new opportunities for breeding resistant crop plants.

Considering high variations in female cyst nematodes across the entire natural population of Arabidopsis, investigations on the genetic architecture for this trait was switched to variations in male cyst nematodes in Arabidopsis.

4.5. Trait 2: Natural variations in formation of male nematodes in Arabidopsis

Cyst nematodes are biotrophic parasites that invade the plant roots and cause extensive physiological and structural changes to form syncytial feeding sites. Male associated syncytial sites are considerably smaller than female associated syncytial sites. Variation in male nematodes development is also influenced by environmental conditions. Under adverse nutritional conditions (weak plant) more male nematodes are produced than females. Male cyst nematodes are important for the fertilization of female cyst nematodes to maintain and increase their reproduction from generation to generation. Therefore, it is rather important to investigate the genetic makeup involved in their mechanistic developmental variation. Here, an assembly of natural population of genetically diversified Arabidopsis accessions was infected with cyst nematode *H. schachtii* to understand the genetic basis and genes involved in theirformation variations.

The numbers of male nematodes were counted after 12 days post infection (dpi) in each accession. Systematic infection assays showed a high natural variation in male nematodes numbers in each accession. Based on these variation, all accessions were categorized into three groups; i) accessions with low male, ii) accessions with moderately increased male and iii) accessions with high male. Accordingly, 41.4% accessions possessed low male numbers, 29.8% accessions had moderately increased males, whereas 30% accessions carried high male numbers.

We analyzed male nematode variation in extreme accessions with lowest (Na-1, Nok-3, Zdr-1, Dr-0, Wa-1, Rmx-AO2, Kondara and Stw-0) and highest (Bsch-0, Pog-0, Ei-2, Or-0, Aa-0, Sei-0, Ba-1 and Nc-1) males. Some accessions showed notable variation such as; Na-1, Nok-3 and Zdr-1 exhibited low male numbers 1.9, 2.2 and 2.8 respectively, whereas these numbers increased to 24 (Sei-0), 26 (Ba-1) and 27.5 (Nc-1) respectively. However, these ranges from 1.9 in Na-1 to 27.5 in Nc-1 were remarkable and lead to understand the genetic basis for this quantitative variation involved. Therefore, genome-wide association mapping was performed to identify the QTLs associated with formation of male nematodes.

4.5.1. QTLs associated with male formation

Several studies were conducted on identification of QTLs responsible for cyst nematode variations in crops such as soybean and wheat and genetic variants were identified and characterized successfully, whereas QTLs particular the response to male cyst nematodes *H. schachtii* in Arabidopsis has never been investigated. Here, as above described natural variation in male nematodes provides an opportunity to investigate the genetic reasons for these variations. Therefore, over 250,000 SNPs were used for GWAS analysis and GWAS identified several QTLs associated with male nematode variation on different chromosome. The most significant QTL was identified at chromosome four with several associated SNPs. SNPs with -log10 P value (9.32) was considered the most significant. We explored 20kb genomic region neighboring to this SNP and based on functional annotations, five candidate genes; at4g18970 (GDSL-lipase), at4g18975 (PPR-proteins), at4g18980 (AtS40-3), at4g18990 (XTH-29) and at4g19000 (IWS2) were selected for further characterization.

4.5.2. Allelic polymorphism at significant SNP marker locus in extreme accessions

Finally, twenty extreme accessions; Na-1, Nok-3, Zdr-1, Dr-0, Wa-1, Rmx-AO2, Kondara and Stw-0 (lowest males) and Bsch-0, Pog-0, Ei-2, Or-0, Aa-0, Sei-0, Ba-1 and Nc-1(highest males) were selected to analyze the allelic polymorphism at associated SNP locus for analysis of candidate genes and their functional characterization. Candidate SNP locus revealed that allele G segregated across accessions with low males, whereas allele A segregated in accessions with high males.

Parallel to this association analysis, SNPs in all candidate genes including at4g18970, at4g18975, at4g18980, at4g18990 and at4g19000 were examined. SNPs were detected in translated and un-translated regions of genes such as; at4g18970, at4g18980 and at4g18990 but these SNPs were not consistent between extreme accessions (lowest and highest males). Identified QTLs conferring variation in male and female cyst nematodes indicates that Arabidopsis natural inbred lines (NIL) provide a natural resource to isolate the genetic allelic variants involved in their sex determination and thus offering an alternative to artificial mutant lines.

4.6. Trait 3: Natural variation in cyst nematode sex ratio in Arabidopsis

Cyst nematodes are sexually dimorphic but sex differentiation is greatly influenced by host environment. A series of studies have been conducted to determine sex ratio variation in cyst nematodes *Heterodera* and *Globodera* species but the phenomenon sex variation in male and female are solely epigenetic or environment dependentis still unclear. Previous studies show that many factors are involved in determining their sex ratio such as; temperature (Melton et al., 1986; Ross, 1964), infection density (Ellenby, 1954; Koliopanos and Triantaphyllou, 1972), host resistance (Bridgeman and Kerry, 1980; Halbrendt et al., 1992) and host nutrition (Grundler et al., 1991).

Here, when a natural population of genetically diversified Arabidopsis accessions was exposed to cyst nematodes, notable female to male sex ratio variation was observed, which supports previous assumptions and findings (Anjam et al., 2017) that sex ratio is influenced by environment and host genotypes. The sex ratio was calculated by dividing the average number of females per plant by the average number of males per plant. Based on the female/male sex ratio, we categorized all genotypes into three susceptibility groups (sex ratio 0-0.5, low susceptibility; sex ratio 0.5-1, moderate susceptibility; sex ratio >1, high susceptibility). Accordingly, 46.6% were lowly susceptible, 26% were moderately susceptible, and 25.4% were highly susceptible.

4.6.1. QTLs associated with cyst nematode sex ratio variation

Identification of QTLs with GWAS provides an opportunity to investigate genes involved in their sex determinations. The interrogation of genes regulating both male and female formations in the natural populations is important and their manipulations with genetic engineering might help to develop new mutations and suppress their infestation dynamics in crops. Here, GWAS identified several QTLs at chromosome four influencing sex ratio of cyst nematodes. SNP marker-trait correlation revealed similar significant loci at chromosome four as detected for trait; “the average numbers of male nematode per accession”. We used accelerated mixed model (AMM) and linear regression (LM; Seren et al., 2012) for GWAS because it accounts for population structure and retrieve kinship-matrix to avoid the false marker associations. This co-localization among two variable traits led to answer that it might bebecause of more variations in male numbers than female numbersacross the entire population and this QTL was driven in association to variable male nematode numbers. To confirm this co-localization we repeated marker-trait analysis using additional approaches such as non-parametric tests (KW) and linear model (LM), however, similar results were

obtained. Five SNPs above the threshold after Bonferroni corrections were considered causal SNPs and their significance with associated alleles across all genotypes was also confirmed with Q-Q plot and candlestick chart. These significant markers were positioned at 10394881bp, 10395213bp, 10394611bp, 10402226bp and 10403363bp respectively. SNPs positioned at 10394611bp, 10394881bp and 10395213bp were flanking in the bi-directional promoter shared by AtS40-3 and PPR-proteins. SNPs positioned at 10402226bp and 10403363bp were flanking in the non-synonymous regions of *XTH-29* gene with silenced functions. To explore this putative genomic region, five candidate genes with locus at4918970 (*GDSL-lipase*), at4g18975 (*PPR-proteins*), at4g18980 (*AtS40-3*), at4g18990 (*XTH-29*), and at4g189000 (*IWS2*) in close proximity of 20kb region to the most significant SNP (positioned at 10394881) were selected for further characterizations.

4.6.2. Allelic polymorphism at significant SNP marker locus in extreme accessions

Twenty extreme accessions were used to repeat infection assays, thus confirming the robustness of our method. We observed great variations among these twenty accessions regarding number of male and female nematode and subsequent effects on their sex ratio. The extremely variable accessions with lowest sex ratio (lowly susceptible to cyst nematodes) named as: Kondara, Uk-1, Van-0, Ta-0, Ty-0, Xan-1, RRS-7, Ei-2 and Gel-1 did not show any phenotypic differences regarding plant growth parameters such as; root length, number of lateral roots and number of leaves etc. but exhibit notable decreased number of females compared with Col-0 control plants. Among these accessions “Kondara” displayed thin and very less lateral roots which might be one of the causal factors in nematodes development with reduced nutritional availability. However, all this still pre-mature to claim as long as the genetic architecture is not well explored. As far as formation of male nematodes is concerned with these lowly susceptible accessions, they give rise to significant increased male numbers except accession Ei-2 which possessed similar males compared with Col-0.

Notably, all these lowly susceptible accessions transmitted significant reduction in female to male sex ratios. In contrary to lowly susceptible accessions: Na-1, Kro-0, Zdr-1, Jm-0, Mc-0. Ha-0, Is-0, Sap-0 and Bg-2 did not show any convincing differences in plant growth parameters. In perception to *H. schachtii* these accessions transmit moderately more or similar number of females compared with Col-0 wild type. If we focus the population of male nematodes on these accessions, similar results were obtained except accession Bg-2 which gripped greater male and female numbers. However, variations in sex ratio among some lines were incredible, which ranged from 0.25 in Kondara (lowly susceptible) to 2.72 in Na-1

(highly susceptible). Similarly, the average number of males ranged from 14.3 in Kondara (lowly susceptible) to only 2 in Na-1 (highly susceptible). Other extreme lines also exhibit variable female to male sex ratio such as; 0.3 in UK-1 (lowly susceptible) to 1.6 in Kro-0 (highly susceptible) and the average number of males ranged 15 in Uk-1 (lowly susceptible) to 3.5 in Kro-0 (highly susceptible). Similarly, sex ratio variations ranged from 0.4 in van-0 (lowly susceptible) to 1.5 in Zdr-1 (highly susceptible) and the average number of males ranged from 18 in Van-0 (lowly susceptible) to 3 in Zdr-1 (highly susceptible). Therefore, considering these substantial deviancies in sex ratio, this trait was selected for further genetic characterization in response to cyst nematodes *H. Schachtii*.

4.7. SNPs in *AtS40-3* may lead to enhanced Arabidopsis susceptibility

Determining SNPs and amino acid changes with all these five candidate genes, we found remarkable polymorphism (SNPs) among lowly and highly susceptible accessions with AtS40-3 **(Fig. 3.17)**. Non-synonymous SNPs has been shown to play an important role in introducing amino acid changes in corresponding proteins. Therefore, the introduction of systematic SNPs in the coding region of *AtS40-3* in extreme accessions suggests that these changes might be additional variants for low susceptibility. This consistent polymorphism was not present with other four candidate genes (*PPR-proteins*, *XTH-29*, *GDSL-lipase* and *IWS2*) among extreme accessions. Polymorphism in the coding region of *AtS40-3* includes the substitution of Arganine to Lysine (19), Lysine to Asparagine (20), Asparagine to Tyrosine (28), Aspartic acid to Glutamic acid (39), Asparagine to Lysine (46), Isoleucine to Phenylalanine (53), Threonine to Lysine (63) and Histidine to Arginine (93).

4.8. AtS40-3 and PPR modulates sex ratio of nematodes

We selected two lowly susceptible (Xan-1 and Van-0) and highly susceptible (Kro-0 and Zdr-1) accessions to analyze the expressions of our selected candidate genes in root tissue without infection to cyst nematodes. We found that out of five candidate genes, expression of two genes (*AtS40-3* and *PPR-proteins*) was strongly reduced in lowly susceptible accessions (Xan-0 and Van-0) as compared with susceptible ones (Kro-0 and Zdr-1). This reduction in the expression (of *AtS40-3* and *PPR-protein*) strongly suggests for polymorphism in their promoter region, which might be the reason for variation in susceptibility of Arabidopsis to cyst nematodes. Indeed, analysis of DNA sequence among the selected lowly susceptible and highly susceptible accessions showed a consistent deletion of 28 nucleotides in the common putative promoter of *AtS40-3* and *PPR-proteins*. Although there were random nucleotide gaps

between both extreme bulks but consistent nucleotide gaps were present among lowly susceptible accessions positioned 10398529bp and it was missing among susceptible.

It is therefore, we proposed that expression of *AtS40-3* and *PPR-protein* is positively correlated with susceptibility of Arabidopsis to cyst nematodes. Notably, the same trend of reduced expression was observed when promoter from a lowly susceptible accession was used to derive the expression of GFP in the epidermis of *Nicotianna benthamiana*. In the support of this hypothesis; we found that loss-of-function mutants for *AtS40-3* showed a significant reduction in the susceptibility of plants to cyst nematodes. It showed significant decreased female numbers and increased male numbers compared with Col-0 control plants. However, in comparison to syncytia sizes with associated females, *AtS40-3* did not show any promising differences.

Determining whether these amino acid changes contribute to the susceptibility of Arabidopsis to cyst nematodes or not are beyond the reach of these studies; however, the fact that changes in expression of *AtS40-3* are positively coo-related with susceptibility to cyst nematodes, suggests that amino acids substitutions in *AtS40-3* in extreme accessions might not be a vital factor to control sex ratio in these accessions. Taken together, these findings supported our putative SNPs selection positioned at 10394611bp, 10394881bp and 10395213bp that these were the true positive allelic variants residing the reasons for susceptibility reduction in response to cyst nematodes *H. schachtii*.

4.9. Promoters involved in bi-directional transcriptional activities

The structural design of plant, growth and its interaction with different environmental conditions are controlled by the series of expression of genes (Chen. et al., 2010; Li. et al., 2011). As control device of gene expression, plant promoters are important in plant biotechnology and functional genomics research for their great application potential in genetic engineering (Cai et al., 2007; Yi et al., 2011). Researchers have been focused on analysis and cloning of uni-directional promoters such as constitutive promoters and inducible promoters (McElroy et al., 1990; Vijayan et al., 2015). However, it has been shown that bidirectional promoters show better applicability than unidirectional promoters in genetic improvement (Yang et al., 2013). It can derive the expression of two adjacent genes concurrently (Kumar et al., 2015) and has become a research focus in plants since recent years. With the development of genome sequencing and bioinformatics analyses in plants such as rice, Arabidopsis and Populus have revealed that the opposing genes regulated by bidirectional promoters have similar characters, like co-expression, functional association,

and conserved arrangement (Krom and Ramakrishna, 2008; Dhadi et al., 2009; Wang et al., 2009; Chen W. et al., 2010). Besides, bidirectional promoters exhibit similar structural characteristics including higher GC contents and less TATA boxes as in mammals (Dhadi et al., 2009).

Previous studies describe that some commonly shared promoters such as light-inducible and tissue-specific promoter between cab1 and cab 2, a tissue-specific common promoter between at5g06280 and at5g06290 and stress-inducible bi-directional promoter between at4g35987 and at4g35985 have been cloned successfully in many species (Bondino and Valle, 2009; Mitra et al., 2009; Banerjee et al., 2013). In this study, another tissue-specific promoter commonly transcribing genes at4g18975 (*PPR-protein*) and at4g18980 (senescence associated protein) showed differential expressions and considerably involved in regulating the male and female numbers of cyst nematode *H. schachtii*.

4.10. Promoter *cis*-acting elements involved in gene expressions regulations

Promoters can broadly be described as regions of DNA located upstream of the transcriptional start site (TSS) of a gene that serve as a binding site (BSs) for the RNA polymerase complex and other transcription factors (TFs). Structurally, a promoter is divided into proximal and distal regions.The proximal region comprises the region (-250 to +250 nucleotides) adjacent to transcription starting sites (TSS) (Butler and Kadonaga, 2002). The least continuous sequence region of the DNA that is necessary to correctly guide the initiation of transcription is called the core promoter and it includes the TSS with -35 to +35 nucleotides (Butler and Kadonaga, 2002). This region usually contains a conserved sequence (T/A or A/T), which is located at nearly 25-30bp from the TSS, known as TATA box. The proximal promoter elements called cis-elements are located 100 (CCAAT-box) and 200 bp (GC-box) above TSS (Griffiths et al., 2000). Other elements such as the initiator (Inr), the element recognized by transcription factor (TF) IIB, B recognition element (BRE) and downstream promoter element (DPE) are generally conserved (A/G) G (A/T) CGTG).

Previous studies reveal that transcription factors (TFs) and their corresponding *cis*-regulating elements (CREs) in promoter are key regulators of gene expression and play an imperative role in higher plants life cycle (Wei et al., 2004). These are programmed interactions between TFs and genomic DNA that bring a genome to its life and define many of its functional features (Grandori et al., 2000; Kohler et al., 2003). Arabidopsis contains more than 100 members including MYB, bHLH, MADS, and AP2/EREBP family of TFs (Riechmann and Ratcliffe, 2000; Hosoda et al., 2002; Heim et al., 2003; Parenicova et al., 2003; Toledo-Ortiz

et al., 2003). It has been therefore, observed commonly that mutations in the coding regions of genes are well conserved among organisms while polymorphism in non-coding sequences has a profound effect on phenotype by altering gene expression.

SNPs in upstream and downstream regions of genes play an important role in the regulation of gene expression because mostly deletions or substitutions are associated with regulatory sequences, e.g. promoters and terminators (Hirkawa et al., 2013) and therefore, might cause induction of phenotypic variations (Vidal et al., 2012; Shi et al., 2015; Shirasawa et al., 2016). Genome-wide functional SNPs are reported in many breeding crop varieties and lines, which influence different phenotypic variations (Kharabian-Masouleh et al. 2012; Kumar et al. 2014; Jang et al. 2015).

C*is*-acting elements have an evolutionary role and polymorphisms may generate the expression variance by regulating TFs binding schemes.The TATA-box is the chief *cis*-acting element in the promoter region comprising a short DNA sequence “3'-TATAAA-5'”. TATA-box are conserved sequences in most of the eukaryotic promoters located 25-35bp before the transcription sites of genes. It provides a platform to the transcription binding proteins (TBP) and recruitment of RNA-polymerase to transcribe the DNA complex.

GATA box or L-box containing GATA sequences mostly have been implicated in light and nitrate dependent control of transcriptions in plants. However, in Arabidopsis a number of GATA elements such as GATA-1, GATA-2, GATA-3 and GATA-4 (Teakle et al., 2002) have been detected with their interacting nuclear proteins. In a study, 28 loci have been found encoding putative GATA factors so far in Arabidopsis (Reyes et al., 2004). Tri-helix (helix-loop-helix-loop-helix) TFs was reported in the regulation of developmental processes (Li et al., 2008), in response to biotic and a-biotic stresses (Xi et al., 2012) and in treatment with phyto-hormones such as abscisic acid or salicylic acid (Fang et al., 2010).

In this study, *AtS40-3* promoter (bi-directional) analysis revealed that presence of some prominent transcription factor binding sequences including TATA box, GATA box, AT-Hook and Tri-helix etc.; however, these motifs were deleted in lowly susceptible accessions which led to the down expression of *AtS40-3* and *PPR-proteins*. We hypothesized that these omitted transcription regulators had strong indirect associations with the decline of cyst nematodes infection among lowly susceptible accessions. As well, the remarkable differences in male and female nematode numbers and in the light of this hypothesis further characterization of related genes paved their importance in the regulations of gene

expressions. However, the additional role of other binding co-factors in the bi-directional promoter of *AtS40-3* and *PPR-proteins* cannot be kept without being seen. Previous studies revealed that senescence-associated genes (SAG) are not only expressed during natural senescence but also in response to various stresses such as wounding, darkness, pathogen infections and in response to treatment of signaling hormones such as ethylene, jasmonate and abscisic acid (Gan and Amasino, 1997; Lim et al., 2007). *AtS40-3* encodes a nuclear targeted protein, which modulated senescence and is located at the 3975bp downstream to the SNP positioned at 10394881.

4.11. AtS40-3, a senescence associated protein in Arabidopsis

AtS40-3 gene in *Arabidopsis thaliana* belongs to a group of genes sharing the conserved DUF548 domain motif with unknown function till now. One member of this group in Barley HvS40 has been shown to display a role in regulation of leaf senescence. As far as the signaling and expression of *AtS40-3* with other senescence-associated genes (SAG) are concerned, it has been demonstrated that during natural senescence, the expression of *AtS40-3* coincides with that of *WRKY-53* transcription factor and senescence-associated gene-12 (*SAG12*). A detailed characterization via expression analysis and T-DNA mutants showed that *AtS40-3* regulates senescence, either by modulating the expression of *WRKY-53* or by activating *SAG12* independent of *WRKY53* (Miao et al., 2004). Our data showed that ats40-3 mutants displayed a significant in number of female nematodes and a significant increase in number of male nematodes, suggesting that the expression of *AtS40-3* is positively regulated relative to the sex ratio of cyst nematodes. Based on these data, we propose that *AtS40-3* expression positively regulates nematodes infection by delaying the arrival of senescence at syncytium thus ensuring a long lasting supply of nutrients to developing nematodes. Intriguingly, previous microarrays analysis with micro-aspirated syncytial content of *H. schachtii* showed that expression of *WRKY53* is strongly up-regulated at 5 and 15 dpi. Incontrast, *AtS40-3* does not show any significant change in expression (Szakastis et al., 2009). In a study mutants of the corresponding gene revealed that *AtS40-3* is involved in regulation of leaf senescence acting upstream of the central senescence regulatory gene *WRKY53* (Miao et al. 2004). However, these observations make it likely that *AtS40-3* acts independently of *WRKY53* to regulate infection of cyst nematode *H. schachtii*.

PPR are a large family of RNA-binding plant proteins which mediate several aspects of gene expression through processing, splicing, editing and translation of mRNAs (Manna, 2015). In addition to *AtS40-3*, our data also hints a positive role of a *PPR* gene (At4g18975) in

determining the sex ratio of cyst nematodes. However, the function of this PPR gene remained completely obscure till to date. Also in present work, it was not possible to produce loss-of-function mutant foe *PPR* gene, which precluded further characterization of this gene. Nonetheless, considering the formation of syncytium involves changes in expression of a large number of genes, we speculate that *PPR* gene positively regulate the nematode infection by influencing the expression of genes essential for syncytium formation and nematode development. However, further work is required to characterize the role of *PPR* in cyst nematodes infection.

4.12. XTH29 and GDSL-lipase modulate Arabidopsis susceptibility

Another candidate gene *GDSL-lipase* with locus at4g18970 neighboring to SNP positioned at 10394881 considered a promising susceptibility regulator in Arabidopsis and involved in modulating the sex ratio of cyst nematodes *H. schachtii*. In this study when knockout mutants of *GDSL-lipase* infected with cyst nematodes *H. schachtii*, a remarkable decrease in the average number of female and an increase in the average number of male nematodes were found compared with parental Col-0 wild type plants. Though the functional mechanism for this regulation in sex ratio is unclear but we can assume that this GDSL motif lipase plays an important role in regulating Arabidopsis susceptibility against cyst nematodes *H. schachtii.*

GDSL-lipases are hydrolytic enzymes of lipases and esterases families possessed conserved Serine, Aspirin and Histidine domains (Akoh et al., 2004) involved in the metabolism of lipid-derived molecules. Mostly lipases possess a GxSxG motif, whereas GDSL-lipases carry a GDSL motif GxSxxxxG with active site Ser located near to N-terminus (Brick et al., 1995). They are predicted extra-cellularly and function in multiple physiological processes including plant growth and defense reactions in Arabidopsis (Linget al., 2008).GDSL lipases play an important role in plant response to various biotic and a-biotic stresses.

Previous studies demonstrated that GDSL-lipase1 (*GLIP1*) an Arabidopsis GDSL-lipase, plays a promising role in eliciting resistance to necrotrophic fungus *Alternaria brassicicola* and it was assumed that it may be an important resistance mediator in association with ethylene signaling (Oh et al., 2005). Later on it was confirmed that GDSL-lipase (*GLIP1*) elicit systemic resistance in plants and functions independently of salicylic acid but solely depends on ethylene signaling to innate immunity (Kwon et al., 2009). It is localized in the cell wall and in response to stress they produce systemic signals in leaves and activate Ethylene receptor (*ERF1*) which in turn activates resistance to necrotrophs and biotrophs. Similarly, in Pepper CaGLIPS-1, encoding a GDSL-motif lipase showed its role in conferring

disease susceptibility against *Xanthomonas campestris* and a-biotic stress response (Kim *et al*., 2008 and Hong *et al*., 2008). It is proposed that GDSL lipase-1 (*GLIP1*) exhibit similar systemic resistance properties with previously described defense regulators; senescence associated gene 101 (*SAG101*) (Feys et al., 2005), enhanced disease susceptibility (EDS1) (Falk et al., 1999) and phytoalexin deficient 4 (*PAD4*) (Jirage et al., 1999).To extend the knowledge about kinetic energies of GLIPs in Arabidopsis another member of this family GDSL-lipase 2 (*GLIP2*) mutants were challenged with necrotrophic bacteria *Erwinia carotovora* and enhanced susceptibility was observed than wild type plants. A plant defense mechanism is solely depends on SA, JA and Ethylene signaling pathways (Schenk et al., 2000) and to extend the knowledge about kinetic energies of GLIPs in Arabidopsis, another member of this family GDSL-lipase 2 (*GLIP2*) was studied. It was proposed that *GLIP-2* plays an important role in regulating resistance against necrotrophic bacteria *Erwinia carotovora* via negative regulation of auxin signaling and its expression is controlled by SA, JA and ethylene in distinct ways (Lee et al., 2009). It is reported that Phytoalexin Deficient-4 (*PAD4*) in Arabidopsis encodes lipase like proteins are involved ineliciting plant immunity against soybean cyst nematodes; *H. glycine* and root knot nematodes; *M. incognita* and is important for mediating SA signaling mechanism (Youseef et al., 2013).

Two significant SNPs positioned at 10402226bp and 10403363bp hitting the non-synonymous regions of *XTH-29* gene. It is therefore, *XTH-29* considered an important susceptibility regulator against cyst nematodes *H. schachtii*. In this study, we proposed that the loss of function mutant of XTH-29 leads to an increase in the average number of female and total numbers of nematodes per plant. This increased preference of nematodes suggests the kinetic properties of *XTH-29* in plant-nematodes interactions. We suggest that absence of *XTH-29* enzyme and the lack of xyloglucan synthesis might affect the mechanical strength of cell walls which in turn facilitates nematode invasion and migration in Arabidopsis roots. This hypothesis is supported by the present results with an increased susceptibility to cyst nematodes *H. schachtii*.

4.13. Xyloglucans are involved in enhancing Arabidopsis susceptibility

Xyloglucans are hemicellulose found in cell walls of all the vascular plants where it is believed to play an important role in the structure and function of cell wall (Del-Bem at al., 2010). Primarily, they are linked with adjacent cellulose-microfibrils to develop a complex load-bearing structure providing mechanical strength to cell wall (Albersheim et al., 2010). In addition, Xyloglucans are the substrate of xyloglucan endotransglycosylase (XTH) enzymes

which has a key role in cell wall modification and expansion (Rose et al., 2002). However, XTH proteins catalyze the xyloglucan cross-links (XEH activity) through cleavage and synthesize new xyloglucans polymers (Smith and Fry 1991), thus support cell wall expansion without weakening. In a previous study, a knockout mutant of *XTH-33* in Arabidopsis when exposed to aphids "*Myzus persicae*" and a significant increased in number of aphids was observed in mutant plants compared with wild type which suggests the involvement of XTH genes in plant protection against pathogens (Divol et al., 2007). During differential expressions analysis of various genes in fungus inoculated jute species, Xyloglucan endotransglycosylase/hydrolase genes showed a steady rise in their expression upon fungus inoculation (Sharmin et al., 2012). The involvement of a large number of genes in cell wall modifications, synthesis and their response to various stresses still remained unclear.Taken together, we can speculate that this cell wall modifying enzyme *XTH29* is a positive regulator of Arabidopsis resistance to nematodes and plays a precise role in strengthening the cell wall to protect against biotic and a-biotic pressures.

Loss of function mutant of *IWS2* (transcription factor) was also brought under consideration and in infection to cyst nematodes did not show any convincing changing in male and female numbers in comparison with Col-0 control plants. It was distanced 11450bp from the most significant SNP (10394881bp). It showed that moving away upstream or downstream to significant genomic region, the marker (SNP)-trait signaling resolution weakens and genes residing loosen their effects. Genome-wide association studies (GWAS) has been emerged as an appealing tool to dissect the causal variants underlying various phenotypic traits in agriculture crops. Initially, it has vastly been conducted in human populations, where it is quite challenging to prove the functional participation of GWAS identified unknown genes.

Nevertheless, in Arabidopsis with the availability of high resolution sequences and appealing molecular studies of various traits, GWAS has become a feasible approach in number of species. However, in this study, GWAS has revealed new findings of novel genomic region (QTL) residing causal genes such as *AtS40-3*, *PPR-proteins*, *GDSL-lipase* and *Xyloglucan endotransglucosylae-29* (*XTH-29*) influencing the sex ratio variation of cyst nematodes *H. schachtii* in Arabidopsis.

5. Conclusion

In the present study, we reported the identification and confirmation of QTL significantly associated with susceptibility to cyst nematode *H. schachtii* in a diverse panel of 148 Arabidopsis accessions. Some extremely variable accessions presenting low and high susceptibility (with decreased and increased nematodes numbers) were identified, thus suggesting the importance and great potential of Arabidopsis as a novel and exotic genetic resource for cyst nematode management. We performed a genome-wide association study with three phenotypic traits; the average male nematodes/plant, the average female nematodes/plant and their subsequent female to male sex ratio/plant. An appropriate marker-trait analysis enabled us to identify several SNP markers significantly associated with QTL of female to male sex ratio at chromosome four. The availability and accessibility of the reference Arabidopsis genome sequence and gene annotation also facilitated the identification of candidate genes leading to the functional analysis. Therefore, various novel genetic variants such as; *GDSL-lipase* (at4g18970), *PPR-proteins* (at4g18975), *AtS40-3* (at4g18980), *XTH-29* (at4g18990) and *IWS2* (at4g19000) were selected regulating the variations in sex ratio of cyst nematode. However, by examining the expression and functional characterization we were then able to discriminate between genetic variants that *AtS40-3*, senescence associated nuclear targeted protein plays a general role in regulating the susceptibility to cyst nematode. Previous studies showed that biotrophic pathogens often delay senescence to keep the cells alive. Given the fact that nematodes are biotrophs, and syncytium serves as the sole source of nutrients for nematodes, it is plausible that regulation of senescence plays a role in the maintenance and functioning of syncytium. Indeed, senescence-like symptoms and upregulation of senescence-associated genes have been observed during resistance response to nematodes in a number of studies (Bleve-Zacheoaet al., 1998; Klink et al., 2009). Based on previous literature and our data, we propose that AtS40-3 expression positively regulates nematode infection by delaying syncytium senescence, thus ensuring an abundant supply of nutrients to nematodes, which favours the formation of females. On the other hand, knocking out *AtS40-3* leads to unfavourable conditions inside the syncytium, including the arrival of early senescence, which may support the development of more males. The results showed that GWAS can be employed as an effective strategy to identify DNA marker alleles associated with quantitative trait which can easily be used in marker-assisted selection (MAS) to facilitate the introduction of desirable alleles into breeding program.

6. References

Abdurakhmonov, I. Y., and Abdukarimov, A. (2008). Application of association mapping to understanding the genetic diversity of plant germplasm resources. International Journal of Plant Genomics, 2008, 1-18.

Akoh, C. C., Lee, G. C., Liaw, Y. C., Huang, T. H., and Shaw, J. F. (2004). GDSL family of serine esterases/lipases. Progress in Lipid Research, 43(6), 534-552.

Albersheim, P., Darvill, A., Roberts, K., Sederoff, R., and Staehelin, A. (2010). Plant Cell Walls. From chemistry to biology. Garland Science, New York, pp. 430.

Altshuler, D., Daly, M. J., and Lander, E. S. (2008). Genetic mapping in human disease. Science, 322, 881-888.

Apel, A., and Kämpfe, L. (1957). Beziehungen zwischen Wirt und Parasit im Infektionsverlauf von *Heterodera schachtii* Schmidt in kurzfristigen Topfversuchen. Nematologica, 2(2), 131-143.

Aranzana, M. J., Kim, S., Zhao, K., Bakker, E., Horton, M., Jakob, K., and Toomajian, C. (2005). Genome-wide association mapping in Arabidopsis identifies previously known flowering time and pathogen resistance genes. PLoS Genetics, 1(5), e60.

Asadi, M. (2007). Beet-Sugar Handbook. John Wiley and Sons Inc, New Jersey, pp. 866.

Atwell, S., Huang, Y. S., Vilhjálmsson, B. J., Willems, G., Horton, M., Li, Y., and Jiang, R. (2010). Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. Nature, 465, 627-631.

Bao, Y., Vuong, T., Meinhardt, C., Tiffin, P., Denny, R., Chen, S., and Young, N. D. (2014). Potential of association mapping and genomic selection to explore PI 88788 derived soybean cyst nematode resistance. The Plant Genome, 7(3), 1-13.

Bardgett, R. D., Cook, R., Yeates, G. W., and Denton, C. S. (1999). The influence of nematodes on below-ground processes in grassland ecosystems. Plant and Soil, 212(1), 23-33.

Barnes, R.D. (1980). Invertebrate zoology.4th edition. Saunders College, Philadelphia, pp1089.

Bastien, M., Sonah, H., and Belzile, F. (2014). Genome wide association mapping of resistance in soybean with a genotyping-by-sequencing approach. The Plant Genome, 7, 1-13.

Baxter, I., Brazelton, J. N., Yu, D., Huang, Y. S., Lahner, B., Yakubova, E., and Vitek, O. (2010). A coastal cline in sodium accumulation in *Arabidopsis thaliana* is driven by natural variation of the sodium transporter AtHKT1; 1. PLoS Genetics, 6(11).

Bello, A., López-Pérez, J. A., García-Álvarez, A., Sanz, R., and Lacasa, A. (2002). Biofumigation and nematode control in the Mediterranean region. Nematology, 4(2), 143.

Del Bem, L. E. V., and Vincentz, M. G. (2010). Evolution of xyloglucan-related genes in green plants. BMC Evolutionary Biology, 10(1), 341.

Bergelson, J., and Roux, F. (2010). Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. Nature Reviews Genetics, 11(12), 867-879.

Bio, J.F., and van Kan, J.A.L. (1988). The synthesis and possible functions of virus-induced proteins in plants. Microbiological sciences. 5, 47-52.

Bird, A. F., (1961). The ultrastructure and histochemistry of a nematode-induced giant cell. The Journal of Biophysical and Biochemical Cytology.11, 701-715.

Bird, D. (1992). Mechanisms of the Meloidogyne-host interaction. In Gommers F, Maas PWT (eds). Nematology: from Molecule to Ecosystem. European Society of Nematologists, Dundee, UK, pp. 51-59.

Bocca, S. N., Magioli, C., Mangeon, A., Junqueira, R. M., Cardeal, V., Margis, R., Sachetto-Martins, G. (2005). Survey of glycine-rich proteins (GRPs) in the *Eucalyptus* expressed sequence tag database (ForEST) Genetics and Molecular Biology, 28, 608-624.

Bohlmann, H., and Sobczak, M. (2014).The plant cell wall in the feeding sites of cyst nematodes. Frontiers in plant science, 5, 89.

Borevitz, J. O., and Nordborg, M. (2003). The impact of genomics on the study of natural variation in Arabidopsis. Plant Physiology, 132(2), 718-725.

Botstein, D., White, R. L., Skolnick, M., and Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American Journal of Human Genetics, 32(3), 314.

Brick, D. J., Brumlik, M. J., Buckley, J. T., Cao, J. X., Davies, P. C., Misra, S., and Upton, C. (1995). A new family of lipolytic plant enzymes with members in rice, Arabidopsis and maize. Febs Letters, 377(3), 475-480.

Bridgeman, M. R., and Kerry, B. R. (1980). The sex ratios of cyst-nematodes produced by adding single second-stage juveniles to host roots. Nematologica, 26(2), 209-213.

Brodie, B. B., and Murphy, W. S. (1975). Population dynamics of plant nematodes as affected by combinations of fallow and cropping sequence. Journal of Nematology, 7(1), 91.

Brown, R. H. (1987). Control strategies in low-value crops. In: Brown, R. H. and Kerry, B. R., eds. Principles and Practice of Nematode Control in Crops. Academic Press, Australia, pp. 351-388.

Buckler, E. S., Holland, J. B., Bradbury, P. J., Acharya, C. B., Brown, P. J., Browne, C., and Goodman, M. M. (2009). The genetic architecture of maize flowering time. Science, 325, 714-718.

Buechner, M., Hall, D. H., Bhatt, H., Hedgecock, E. M. (1999). Cystic canal mutants in *Caenorhabditis elegans* are defective in the apical membrane domain of the renal (excretory). Cell and Developmental Biology, 214: 227-241.

Butler, J. E. F., and Kadonaga, J. T. (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. Genes and Development, 6, 2583-2592.

Cai, M., Wei, J., Li, X., Xu, C., and Wang, S. (2007). A rice promoter containing both novel positive and negative cis-elements for regulation of green tissue-specific gene expression in transgenic plants. Plant biotechnology journal, 5(5), 664-674.

Cao, J., Schneeberger, K., Ossowski, S., Gu¨ nther, T., Bender, S., Fitz, J., Koenig, D., Lanz, C., Stegle, O., Lippert, C., et al. (2011). Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. Nature Genetics, 43, 956-963.

Chan, E. K., Rowe, H. C., Corwin, J. A., Joseph, B., and Kliebenstein, D. J. (2011). Combining genome-wide association mapping and transcriptional networks to identify novel genes controlling glucosinolates in *Arabidopsis thaliana*. PLoS Biology, 9(8), e1001125.

Chen, A. P., Zhong, N. Q., Qu, Z. L., Wang, F., Liu, N., and Xia, G. X. (2007). Root and vascular tissue-specific expression of glycine-rich protein AtGRP9 and its interaction with AtCAD5, a cinnamyl alcohol dehydrogenase, in *Arabidopsis thaliana*. Journal of plant research, 120(2), 337-343.

Chen, F., Gao, M. J., Miao, Y. S., Yuan, Y. X., Wang, M. Y., Li, Q., and He, Z. H. (2010). Plasma membrane localization and potential endocytosis of constitutively expressed XA21 proteins in transgenic rice. Molecular plant, 3(5), 917-926.

Chen, H., Xie, W., He, H., Yu, H., Chen, W., Li, J., Yu, R., Yao, Y., Zhang, W., He, Y., Tang, X., Zhou, F., Deng, X. W., and Zhang, Q. (2014). A high-density SNP genotyping array for rice biology and molecular breeding. Molecular Plant.7, 541-553.

Chen, L. Q., Hou, B. H., Lalonde, S., Takanaga, H., Hartung, M. L., Qu, X. Q., and Chermak, D. (2010). Sugar transporters for intercellular exchange and nutrition of pathogens. Nature, 468(7323), 527.

Chitwood, B. G., and Chitwood, M. B. (1950). General structure of nematodes. In: Chitwood, B.G., and Chitwood, M. B., eds. An introduction to nematology. University Park Press, Baltimore, pp. 7-27.

Choi, M. S., Kim, W., Lee, C., Oh, C. S. (2013). Harpins, multifunctional proteins secreted by gram-negative plant-pathogenic bacteria. Molecular Plant-Microbe Interactions, 26, 1115-1122.

Christie, J. R. (1929). Some observations on sex in the *Mermithidae*. Journal of Experimental Zoology Part A: Ecological Genetics and Physiology, 53(1), 59-76.

Chung, W. H., Jeong, N., Kim, J., Lee, W. K., Lee, Y. G., Lee, S. H., Yoon, W. C., Kim, J. H., Choi, I. Y., Choi, H. K., Moon, J. K., Kim, N., and Jeong, S. C. (2014). Population structure and domestication revealed by high-depth resequencing of Korean cultivated and wild soybean genomes. DNA Research, 21, 153-167.

Clark, R. M., Schweikert, G., Toomajian, C., Ossowski, S., Zeller, G., Shinn, P., and Chen, H. (2007). Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. Science, 317(5836), 338-342.

Clarke, A. J. (1968). The chemical composition of the cyst wall of the potato cyst-nematode, *Heterodera rostochiensis*. Biochemical Journal, 108(2), 221-224.

Cobb, N. A., Steiner, G., and Christie, J. R. (1927). When and how does sex arise? United state Department of Agriculture.Official Record 6, no. 23.

Collard, B. C. Y., Jahufer, M. Z. Z., Brouwer, J. B., and Pang, E. C. K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. Euphytica, 142(1-2), 169-196.

Collins, F. S., Guyer, M. S., and Chakravarti, A. (1997). Variations on a theme: cataloging human DNA sequence variation. Science, 278(5343), 1580-1581.

Condit, C.M., and Meagher, R.B. (1986). A gene encoding a novel glycine-rich structural protein of petunia. Nature, 323,178-181.

Cook, D. E., Lee, T. G., Guo, X., Melito, S., Wang, K., Bayless, A. M., and Diers, B. W. (2012). Copy number variation of multiple genes at Rhg1 mediates nematode resistance in soybean. Science, 338(6111), 1206-1209.

Cook, R., and Starr, J. L. (2006). Resistant cultivars. In: Perry, R.N., and Moens, M., eds. Plant Nematology. CAB International, UK, pp. 370-391.

Cooke, D. A. (1987).Beet cyst nematode (*Heterodera schachtii* Schmidt) and its control on sugar beet. Agricultural Zoology Reviews, 2, 135-184.

Cosgrove, D. J. (1997). Assembly and enlargement of the primary cell wall in plants. Annual review of cell and developmental biology, 13(1), 171-201.

Crosses, E. (2001). Review of statistical methods for QTL mapping in experimental crosses. Lab Animal, 30(7), 44-52.

Curtis, R. H. (2007). Plant parasitic nematode proteins and the host-parasite interaction. Briefings in Functional Genomics and Proteomics, 6(1), 50-58.

Curto, G. (2008). Sustainable methods for management of cyst nematodes. In: Ciancio A., Mukerji K.G., eds. Integrated Management and Biocontrol of Vegetable and Grain Crops Nematodes. Springer, Berlin, pp. 221-237.

d'Amato, F. (1984). Role of polyploidy in reproductive organs and tissues. In Embryology of angiosperms.Springer, Berlin, Heidelberg, pp. 519-566.

Davis, E. L. and Mitchum, M. G. (2005). Nematodes. Sophisticated parasites of legumes. Plant Physiology. 137 1182-1188.

Davis, E. L., Hussey, R. S., Baum, T. J., Bakker, J., Schots, A., Rosso, M. N., and Abad, P. (2000). Nematode parasitism genes. Annual Review of Phytopathology, 38(1), 365-396.

de Goede, R.G.M. and Bongers, T. (1994). Nematode community structure in relation to soil and vegetation characteristics. Applied Soil Ecology, 1, 29-44.

de Oliveira, D. E., Seurinck, J., Inzé, D., Van Montagu, M., and Botterman, J. (1990). Differential expression of five Arabidopsis genes encoding glycine-rich proteins. The Plant Cell, 2(5), 427-436.

Decraemer, W., and Hunt, D.J. (2006): Structure and classification. In: Perry, R.N. and Moens, M., eds. Plant Nematology, CAB International, UK, pp. 3-32.

deGoede, R. G., and Bongers, T. (1994). Nematode community structure in relation to soil and vegetation characteristics. Applied Soil Ecology, 1(1), 29-44.

Del Bem, L. E. V., and Vincentz, M. G. (2010). Evolution of xyloglucan-related genes in green plants. BMC Evolutionary Biology, 10(1), 341.

Den Ouden, H. (1960). A note on parthenogenesis and sex determination in *Heterodera rostochiensis* Woll. Nematologica, 5(3), 215-216.

Dhadi, S. R., Krom, N., and Ramakrishna, W. (2009). Genome-wide comparative analysis of putative bidirectional promoters from rice, Arabidopsis and Populus.Gene, 429(1), 65-73.

Dimkpa, S. O., Lahari, Z., Shrestha, R., Douglas, A., Gheysen, G., Price, A. H. (2016). A genome-wide association study of a global rice panel reveals resistance in *Oryza sativa* to root-knot nematodes. Journal of Experimental Botany, 67, 1191-1200.

Divol, F., Vilaine, F., Thibivilliers, S., Kusiak, C., Sauge, M. H., and Dinant, S. (2007). Involvement of the xyloglucan endotransglycosylase/hydrolases encoded by celery XTH1 and Arabidopsis XTH33 in the phloem response to aphids. Plant, Cell and Environment, 30(2), 187-201.

Djian-Caporalino, C., Fazari, A., Arguel, M. J., Vernie, T., VandeCasteele, C., Faure, I. and Abad, P. (2007). Root-knot nematode (*Meloidogyne spp*.) Mi resistance genes in pepper (*Capsicum annuum L*.) are clustered on the P9 chromosome. Theoretical and Applied Genetics, 114(3), 473-486.

Doliwa,U. (1956). Wissenschaftliche Zeitschrift der Universität. Rostock, pp. 5,133.

Donnelly, P. (2008). Progress and challenges in genome-wide association studies in humans. Nature, 456(7223), 728-731.

Dreyfuss, G., Matunis, M. J., Pinol-Roma, S., Burd, C. G. (1993). HnRNP proteins and the biogenesis of mRNA. Annual Review of Biochemistry, 62, 289-321.

Easlon, H. M., Nemali, K. S., Richards, J. H., Hanson, D. T., Juenger, T. E., and McKay, J. K. (2014). The physiological basis for genetic variation in water use efficiency and carbon isotope composition in *Arabidopsis thaliana*. Photosynthesis Research, 119(1-2), 119-129.

Edwards, D., Forster, J. W., Chagné, D., Batley, J. (2007). What are SNPs? In: Oraguzie, N. C., Rikkerink, E. H. A., Gardiner, S. E., and de Silva, H. N., eds. Association Mapping in Plants. Springer, New York, pp. 41-52.

Eisenback, J. D., and Triantaphyllou., H. H. (1991). Root-knot nematodes. In: Nickle, W. R., eds. *Meloidogyne* species and races. Manual of Agricultural Nematology. New York, pp. 191-274.

Ellenby, C. (1954). Environmental determination of the sex ratio of a plant parasitic nematode. Nature, 174(4439), 1016-1017.

Ellenby, C. (1963). The gold of the golden nematode of potatoes. Proceedings of XVl lnternational Congress of Zoology, Washington, 1(139).

Evans, K. and Rowe, J.A. (1998). Distribution and economic importance. In: Sharmam S.B. eds. The cyst nematodes. Dordrecht, The Netherlands, pp. 1-30.

Falk, A., Feys, B. J., Frost, L. N., Jones, J. D., Daniels, M. J., and Parker, J. E. (1999). EDS1, an essential component of R gene-mediated disease resistance in Arabidopsis has homology to eukaryotic lipases. Proceedings of the National Academy of Sciences, 96(6), 3292-3297.

Fernández, C., Pinochet, J., Esmenjaud, D., Salesses, G., and Felipe, A. (1994). Resistance among new Prunus rootstocks and selections to root-knot nematodes in Spain and France. Hort Science, 29(9), 1064-1067.

Ferris, H., Venette, R. C., Van der Meulen, H. R., and Lau, S. S. (1998). Nitrogen mineralization by bacterial-feeding nematodes: Verification and measurement. Plant and Soil, 203, 159-171.

Feys, B. J., Wiermer, M., Bhat, R. A., Moisan, L. J., Medina-Escobar, N., Neu, C., and Parker, J. E. (2005). Arabidopsis senescence-associated gene 101 stabilizes and signals within an enhanced disease susceptibility 1 complex in plant innate immunity. The Plant Cell, 17(9), 2601-2613.

Filiault, D. L., and Maloof, J. N. (2012). A genome-wide association study identifies variants underlying the *Arabidopsis thaliana* shade avoidance response. PLoS Genetics, 8(3).

Francisco, M., Joseph, B., Caligagan, H., Li, B., Corwin, J. A., Lin, C., and Kliebenstein, D. J. (2016). Genome wide association mapping in *Arabidopsis thaliana* identifies novel genes involved in linking allyl glucosinolate to altered biomass and defense. Frontiers in Plant Science, 7,1010.

Frandsen, G. I., Mundy, J., and Tzen, J. T. C. (2001). Oil bodies andtheir associated proteins, oleosin and caleosin.Physiologia Plantarum, 112, 301-307.

Frandsen, G. I., Mundy, J., andTzen, J. T. (2001). Oil bodies and their associated proteins, oleosin and caleosin. Physiologia Plantarum, 112(3), 301-307.

Frankham, R. (2005). Resolving the genetic paradox in invasive species. Heredity, 94(4), 385-385.

Gan, S., and Amasino, R. M. (1997). Making sense of senescence (molecular genetic regulation and manipulation of leaf senescence). Plant Physiology, 113(2), 313-319.

Gan, X., Stegle, O., Behr, J., Steffen, J.G., Drewe, P., Hildebrand, K.L., Lyngsoe, R., Schultheiss, S.J., Osborne, E.J., Sreedharan, V.T., et al. (2011). Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. Nature, 477, 419-423.

Gao, B., Allen, R., Maier, T., Davis, E. L., Baum, T. J., and Hussey, R. S. (2003). The parasitome of the phytonematode *Heterodera glycines*. Molecular Plant-Microbe Interactions, 16(8), 720-726.

Gelderman, H. (1975). Investigation on inheritance of quantitative characters in animals by gene markers. I. Methods.Theoretical and Applied Genetics, 46, 300-19.

Ghosh, S., Malhotra, P., Lalitha, P. V., Guha-Mukherjee, S., and Chauhan, V. S. (2002). Novel genetic mapping tools in plants: SNPs and LD-based approaches, Plant Science, 162 (3), 329-333.

Golinowski, W., Grundler, F.M.W., and Sobczak, M. (1996). Changes in the structure of *Arabidopsis thaliana* during female development of the plant-parasitic nematode *Heterodera schachtii*. Protoplasma, 194, 103-116.

Gòmez, J., Sanchez-Martíner, D., Stiefel, V., Rigau, J., Puigdomenech, P., and Pages, M. (1988). A gene induced by the plant hormone abscisic acid in response to water stress encodes a glycine-rich protein. Nature, 334, 262-264.

Grandori, C., Cowley, S. M., James, L. P., and Eisenman, R. N. (2000). The Myc/Max/Mad network and the transcriptional control of cell behavior. Annual Review of Cell and Developmental Biology, 16(1), 653-699.

Greco, N., Brandonisio, A., and De Marinis, G. (1982).Tolerance limit of the sugarbeet to *Heterodera schachtii*. Journal of Nematology, 14(2), 199.

Griffiths, A. J. F., Suzuki, D. T., Lewontin, R. C. and Gelbart, W. M. (2000). Regulation of gene transcription, in Modern Genetic Analysis, 3rd edition, (W.H. Freeman and Company), New York, USA.

Grundler, F. M. W., Betka, M., and Wyss, U. (1991). Influence of changes in the nurse cell system (syncytium) on sex determination and development of the cyst nematode *Heterodera schachtii*: total amounts of proteins and amino acids. Phytopathology, 81(1), 70-74.

Grundler, F. M. W., Sobczak, M., Golinowski, W. (1998). Formation of wall openings in root cells of Arabidopsis thaliana following infection by the plant-parasitic nematode *Heterodera schachtii*. European Journal Plant Pathology, 104, 545-551.

Gupta, P. K., Rustgi, S., and Kulwal, P. L. (2005). Linkage disequilibrium and association studies in higher plants: present status and future prospects. Plant Molecular Biology, 57(4), 461-485.

Halbrendt, J., S. A. Lewis, and Shipe, E. (1992). A technique for evaluating *Heterodera glycines* development in susceptible and resistant soybean. Journal of Nematology, 24, 84-91.

Haley, C. S., and Knott, S. A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity, 69(4), 315-324.

Hammes, U.Z., Schachtman, D.P., Berg, R.H., Nielsen, E., Koch, W., McIntyre, L.M., and Taylor, C.G. (2005). Nematode-induced changes of transporter gene expression in Arabidopsis roots. Molecular Plant Microbe Interaction, 18, 1247-1257.

Hammond-Kosack, K. E., and Jones, J. D. G. (1996). Resistance gene-dependent plant defense responses. Plant Cell, 8, 1773-1791.

Handoo, Z. A. (1998). Plant-parasitic nematodes. United State Department of Agriculture, Agriculture Research Services, Nematology Laboratory.

Heim, M. A., Jakoby, M., Werber, M., Martin, C., Weisshaar, B., and Bailey, P. C. (2003). The basic helix-loop-helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. Molecular Biology and Evolution, 20(5), 735-747.

Hindorff, L. A., Sethupathy, P., Junkins, H. A., Ramos, E. M., Mehta, J. P., Collins, F. S., and Manolio, T. A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proceedings of the National Academy of Sciences, 106(23), 9362-9367.

Hirakawa, H. I. D. E. K. I., Shirasawa, K. E. N. T. A., Ohyama, A. K. I. O., Fukuoka, H. I. R. O. Y. U. K. I., Aoki, K. O. H., Rothan, C. H. R. I. S. T. O. P. H. E., and Tabata, S. A. T. O. S. H. I. (2013). Genome-wide SNP genotyping to infer the effects on gene functions in tomato. DNA research, 20(3), 221-233.

Hodda, M. (2011). Phylum Nematoda Cobb 1932. Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness. Zootaxa, 3148, 63-95.

Hong, J. K., Choi, D. S., Kim, S. H., Yi, S. Y., Kim, Y. J., and Hwang, B. K. (2008). Distinct roles of the pepper pathogen-induced membrane protein gene CaPIMP1 in bacterial disease resistance and oomycete disease susceptibility. Planta, 228(3), 485-497.

Horton, M. W., Hancock, A. M., Huang, Y. S., Toomajian, C., Atwell, S., Auton, A., and Nordborg, M. (2012). Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. Nature Genetics, 44(2), 212-216.

Hosoda, K., Imamura, A., Katoh, E., Hatta, T., Tachiki, M., Yamada, H., and Yamazaki, T. (2002). Molecular structure of the GARP family of plant Myb-related DNA binding motifs of the Arabidopsis response regulators. The Plant Cell, 14(9), 2015-2029.

Huang, S. P. (1984). Cropping effects of marigolds, corn, and okra on population levels of *Meloidogyne javanica* and on carrot yields. Journal of Nematology, 16(4), 396.

Huang, X., and Han, B. (2014). Natural variations and genome-wide association studies in crop plants. Annual Review of Plant Biology, 65, 531-551.

Huang, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., Li, C., and Fan, D. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. Nature Genetics, 42(11), 961-967.

Hunt, D. J., Luc, M., and Manzanilla-Lopez, R. H. (2005). Identification, morphology and biology of plant parasitic nematodes. In: Luc, M. Sikora, R.A., and Bridge, J., 2nd eds. Plant parasific nematodes in subtropical and tropical agriculture. CAB International, UK, pp. 11-52.

Hunt, D., and Handoo, Z. (2009). Taxonomy, identification and principal species. In: R. N. Perry, M. Moens and J. L. Starr, eds. Root- knot nematodes. CAB International, London, pp. 55-88

Hussey, R. S. (1989). Disease-inducing secretions of plant-parasitic nematodes. Annual Review of Phytopathology, 27(1), 123-141.

Hussey, R. S., and Grundler, F. M. (1998). Nematode parasitism of plants. In: Perry, R. N., Wright, J. eds. Physiology and Biochemistry of Free-Living and Plant Parasitic Nematodes. CAB International Press, Oxford, pp. 213-243.

Hwang, E. Y., Song, Q., Jia, G., Specht, J. E., Hyten, D. L., Costa, J., and Cregan, P. B. (2014). A genome-wide association study of seed protein and oil content in soybean. BMC Genomics, 15(1), 1.

International HapMap 3 Consortium. (2010). Integrating common and rare genetic variation in diverse human populations. Nature, 467(7311), 52-58.

Jang, S. J., Sato, M., Sato, K., Jitsuyama, Y., Fujino, K., Mori, H., and Abe, J. (2015). A single-nucleotide polymorphism in an endo-1, 4-β-glucanase gene controls seed coat permeability in soybean. PLoS One, 10(6).

Jansen, R. C. (1993). Interval mapping of multiple quantitative trait loci. Genetics, 135(1), 205-211.

Jia, G., Huang, X., Zhi, H., Zhao, Y., Zhao, Q., Li, W., and Zhu, C. (2013). A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setariaitalica*). Nature Genetics, 45(8), 957-961.

Jirage, D., Tootle, T. L., Reuber, T. L., Frost, L. N., Feys, B. J., Parker, J. E., and Glazebrook, J. (1999). *Arabidopsis thaliana* PAD4 encodes a lipase-like gene that is important for salicylic acid signaling. Proceedings of the National Academy of Sciences, 96(23), 13583-13588.

Johnson, A. W. (1982). Managing nematode populations in crop production. Nematology in the southern region of the United States. Southern Cooperative Series Bulletin, 276, 193-203.

Johnson, A. W., and Fassuliotis, G. (1984). Nematode parasites of vegetable crops. Plant and insect nematodes, 323-372.

Johnson, R., and Viglierchio, D. (1969). Sugar beet nematode (*Heterodera schachtii*) reared on axenic Beta vulgaris root explants. Nematologica, 15,144-152.

Juenger, T. E. (2013). Natural variation and genetic constraints on drought tolerance. Current Opinion in Plant Biology, 16(3), 274-281.

Jun, T. H., Van, K., Kim, M. Y., Lee, S. H., and Walker, D. R. (2008). Association analysis using SSR markers to find QTL for seed protein content in soybean. Euphytica, 162(2), 179-191.

Kämpfe, L., and Kerstan, U. (1964). Die Beeinflussung des Geschlechtsverhältnisses in der Gattung *Heterodera* Schmidt. Nematologica, 10(3), 388-398.

Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman, D., Daly, M. J., and Eskin, E. (2008). Efficient control of population structure in model organism association mapping. Genetics, 178(3), 1709-1723.

Karimi, M., de Oliveira Manes, C. L., Van Montagu, M., andGheysen, G. (2002). Activation of a pollenin promoter upon nematode infection. Journal of nematology, 34(2), 75.

Karssen, G. (2002). The plant-parasitic nematode genus *MeloidogyneGoeldi*, 1892 (Tylenchida) in Europe. Brill, Leiden, The Netherlands, pp 157.

Kearsey, M. J., Pooni, H. S. (1996). The genetical analysis ofquantitative traits. London: Chapman and Hall, pp. 381.

Keller, B., Sauer, N., and Lamb, C.J. (1 988). Glycine-rich cell wall proteins in bean: Gene structure and association of the protein with the vascular system. The EMBO Journal,7, 3625-3633.

Kerry, B. R. (1982). The Decline of *Heterodera avenae* Populations1. EPPO Bulletin, 12(4), 491-496.

Kerry, B. R., Crump, D. H., and Mullen, L. A. (1982). Studies of the cereal cyst-nematode, *Heterodera avenae* under continuous cereals, 1975–1978. II. Fungal parasitism of nematode females and eggs. Annals of Applied Biology, 100(3), 489-499.

Kharabian-Masouleh, A., Waters, D. L. E., Reinke, R. F., Ward, R., and Henry, R. J. (2012). SNP in starch biosynthesis genes associated with nutritional and functional properties of rice. Science Reports, 2, 557.

Kharabian-Masouleh, A., Waters, D. L. E., Reinke, R. F., Ward, R., and Henry, R. J. (2012). SNP in starch biosynthesis genes associated with nutritional and functional properties of rice. Science Reports, 2, 557.

Kim, D. S., Kim, N. H., Hwang, B. K. (2015). Glycine-rich rna-binding protein1 interacts with receptor-like cytoplasmic protein kinase1 and suppresses cell death and defense responses in pepper (Capsicum annuum). New Phytology, 205: 786-800.

Kim, J. Y., Park, S. J., Jang, B., Jung, C.-H., Ahn, S. J., Goh, C.-H., et al. (2007). Functional characterization of a glycine-rich RNA-binding protein2 in *Arabidopsis thaliana* under abiotic stress conditions. Plant Journal. 50, 439-451.

Kim, Y. H., Kim, C. Y., Song, W. K., Park, D. S., Kwon, S. Y., Lee, H. S. and Kwak, S. S. (2008). Over expression of sweet-potato swpa4 peroxidase results in increased hydrogen peroxide production and enhances stress tolerance in tobacco. Planta, 227(4), 867-881.

Kingsley, P.D. and Palis, J. (1994). GRP2 proteins contain both CCHC zinc fingers and a cold shock domain. The Plant Cell, 6, 1522-1523.

Köhler, C., Hennig, L., Spillane, C., Pien, S., Gruissem, W., and Grossniklaus, U. (2003).The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1. Genes and Development, 17(12), 1540-1553.

Koliopanos, C. N., and Triantaphyllou, A. C. (1972). Effect of infection density on sex ratio of *Heterodera glycines*. Nematologica, 18,131-137.

Koncz, C., Chua, N. H., Schell, J. 1998. Methods in Arabidopsis Research. In: Martinez-Zapater, J. M., and Salinas, J., eds. Methods in Molecular Biology (Arabidopsis Protocols). Totowa, Humana, 82.

Koncz, C., Németh, K., Rédei, G.P. and Schell, J. (1992).T-DNA insertional mutagenesis in *Arabidopsis*. Plant Molecular Biology, 20, 963-976.

Koornneef, M., Alonso-Blanco, C., and Vreugdenhil, D. (2004). Naturally occurring genetic variation in *Arabidopsis thaliana*. Annual Review of Plant Biology, 55, 141-172.

Koornneef, M., and Meinke, D. (2010). The development of Arabidopsis as a model plant. The Plant Journal, 61(6), 909-921.

Korte, A., and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. Plant Methods, 9(1), 29.

Krom, N., Recla, J., and Ramakrishna, W. (2008). Analysis of genes associated with retrotransposons in the rice genome. Genetica, 134(3), 297-310.

Kruglyak, L. (1997). The use of a genetic map of biallelic markers in linkage studies. Nature Genetics, 17(1), 21-24.

Kumar, B., Abdel-Ghani, A. H., Pace, J., Reyes-Matamoros, J., Hochholdinger, F., and L? berstedt, T. (2014). Association analysis of single nucleotide polymorphisms in candidate genes with root traits in maize (*Zea mays L.*) seedlings. Plant Science,224, 9-19.

Kumar, S., AlAbed, D., Whitteck, J. T., Chen, W., Bennett, S., Asberry, A., and Gupta, M. (2015). A combinatorial bidirectional and bicistronic approach for coordinated multi-gene expression in corn. Plant molecular biology, 87(4-5), 341-353.

Kump, K. L., Bradbury, P. J., Wisser, R. J., Buckler, E. S., Belcher, A. R., Oropeza-Rosas, M. A., and Balint-Kurti, P. J. (2011). Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. Nature Genetics, 43(2), 163-168.

Kutach, A. K., and Kadonaga, J. T. (2000). The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters. Molecular and Cellular Biology, 20(13), 4754-4764.

Kwon, S. J., Kwon, S. I., Bae, M. S., Cho, E. J., and Park, O. K. (2007). Role of the methionine sulfoxide reductase MsrB3 in cold acclimation in Arabidopsis. Plant and Cell Physiology, 48(12), 1713-1723.

Lambshead, P. J. D. (1993). Recent developments in marine benthic biodiversity research. Oceanis, 19, 5-24.

Lander, E. S., and Botstein, D. (1986).Strategies for studying heterogeneous genetic traits in humans by using a linkage map of restriction fragment length polymorphisms. Proceedings of the National Academy of Sciences, 83(19), 7353-7357.

Lander, E. S., and Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics, 121(1), 185-199.

Lee, D. S., Kim, B. K., Kwon, S. J., Jin, H. C., and Park, O. K. (2009). Arabidopsis GDSL lipase 2 plays a role in pathogen defense via negative regulation of auxin signaling. Biochemical and Biophysical Research Communications, 379(4), 1038-1042.

Li, C., Wang, Y., Liu, L., Hu, Y., Zhang, F., Mergen, S., and Chu, C. (2011). A rice plastidial nucleotide sugar epimerase is involved in galactolipid biosynthesis and improves photosynthetic efficiency. PLoS Genetics, 7(7).

Li, X., Qin, G., Chen, Z., Gu, H., and Qu, L. J. (2008). A gain-of-function mutation of transcriptional factor PTL results in curly leaves, dwarfism and male sterility by affecting auxin homeostasis. Plant Molecular Biology, 66(3), 315-327.

Li, X., Wang, X., Zhang, S., Liu, D., Duan, Y., and Dong, W. (2011). Comparative profiling of the transcriptional response to soybean cyst nematode infection of soybean roots by deep sequencing. Chinese Science Bulletin, 56(18), 1904-1911.

Li, Y., Huang, Y., Bergelson, J., Nordborg, M., and Borevitz, J. O. (2010). Association mapping of local climate-sensitive quantitative trait loci in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences, 107(49), 21199-21204.

Lim, P. O., Kim, H. J., and Gil Nam, H. (2007). Leaf senescence. Annual Review of Plant Biology. 58, 115-136.

Linford, M. B. (1941). Parasitism of the root-knot nematode in leaves and stems. Phytopathology, 31(7), 634-648.

Ling, H. (2008). Sequence analysis of GDSL lipase gene family in *Arabidopsis thaliana*. Pakistan Journal of Biological Sciences, 11(5), 763.

Ling, H., Zhao, J., Zuo, K., Qiu, C., Yao, H., Qin, J., and Tang, K. (2006).Isolation and expression analysis of a GDSL-like lipase gene from Brassica napus L. Journal of Biochemistry and Molecular Biology, 39(3), 297.

Linthorst, H. J., Van Loon, L. C., Memelink, J., and Bol, J. F. (1990). Characterization of cDNA clones for a virus-inducible, glycine-rich protein from petunia. Plant molecular biology, 15(3), 521-523.

Linthorst, H.J.M., Meuwissen, R.L. J., Kauffmann, S., and Boi, J. F. (1989). Constitutive expression of pathogenesis-related proteins PR-1, GRP, and PR-S in tobacco has no effect on virus infection. Plant Cell, 1, 285-291.

Long, Q., Rabanal, F.A., Meng, D., Huber, C.D., Farlow, A., Platzer, A., Zhang, Q., Vilhja´ lmsson, B.J., Korte, A., Nizhynska, V., et al. (2013). Massive genomic variation and strong selection in *Arabidopsis thaliana* lines from Sweden. Nature Genetics, 45, 884-890.

Maggenti, A. R., (1981). General Nematology. Springer-Verlag, New York. Heidelberg, Berlin, pp. 372.

Maggenti, A. R., and Allen, M. W. (1960). The origin of the gelatinous matrix in *Meloidogyne*. The Proceedings of the Helminthological Society of Washington, 27, 4-10.

Manna, S. (2015). An overview of pentatricopeptide repeat proteins and their applications. Biochimie, 113, 93-99.

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J. and Cho, J. H. (2009). Finding the missing heritability of complex diseases. Nature, 461(7265), 747-753.

Mascarenhas, J. P. (1975). The biochemistry of angiosperm pollen development. Botanical Review, 3, 259-314.

Mauricio, R. (2001). Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. Nature Reviews Genetics, 2(5), 370-381.

McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P., and Hirschhorn, J. N. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nature Reviews Genetics, 9(5), 356-369.

McElroy, D., Zhang, W., Cao, J., and Wu, R. (1990). Isolation of an efficient actin promoter for use in rice transformation.The Plant Cell, 2(2), 163-171.

McGawley, E. C. (1986). Infuence of crop rotation on nematodes community structure. Journal of Nematology, 18(4), 620-620.

McKenry, M. V., and Roberts, P. A. (1985). Cooperative Extension University of California. Division of Agriculture and Natural Resources. Publications 4045. Phytonematology study guide.

McLaren, D. J. (1976). Nematode sense organs. Advances in Parasitology, 14, 195-265.

Meijón, M., Satbhai, S. B., Tsuchimatsu, T., and Busch, W. (2014). Genome-wide association study using cellular traits identifies a new regulator of root development in Arabidopsis. Nature Genetics, 46(1), 77-81.

Melton, T. A., Jacobsen, B. J., and Noel, G. R. (1986). Effects of temperature on development of *Heterodera glycines* on *Glycine max* and *Phaseolus vulgaris*. Journal of Nematology, 18(4), 468-474.

Meyerowitz, E. M., and Somerville, C. R. (1994). Arabidopsis. Cold Spring Harbor Laboratory Press, pp. 297-312.

Miao, Y., Laun, T., Zimmermann, P., and Zentgraf, U. (2004). Targets of the WRKY53 transcription factor and its role during leaf senescence in Arabidopsis. Plant Molecular Biology, 55(6), 853-867.

Mohan, M., Nair, S., Bhagwat, A., Krishna, T. G., Yano, M., Bhatia, C. R., and Sasaki, T. (1997). Genome mapping, molecular markers and marker-assisted selection in crop plants. Molecular Breeding, 3(2), 87-103.

Molina, A., Mena, M., Carbonero, P., and García-Olmedo, F. 1997. Differential expression of pathogen-responsive genes encoding two types of glycine-rich proteins in barley.Plant Molecular Biology, 33, 803-810.

Molz, E. (1920). Versuche zur Ermittlung des Einflusses äusserer Faktoren auf das Geschlechterverhältnis des Rübennematoden *Heterodera schachtii* Schmidt. Journal für Landwirtschaft, 54, 769-771/787-791.

Muller, J. (1985). Der Einfluss Wirtspflanze auf die Geschlechtsdeterminierung bei *Heterodera schachtii*. der Biologischen Bundesanstalt fur Land- und Forstwirtschaft, Institut für Nematologie, 46-63.

Muller, J., Rehbock, K., and Wyss, U. (1981). Growth of *Heterodera schachtii* with remarks on amounts of food consumed. Revue de Nématologie, 4(2), 227-234.

Mulvey, R. H. (1972). Identification of *Heterodera* cysts by terminal and cone top structures. Canadian Journal of Zoology, 50(10), 1277-1292.

Murphy, D. J., Hernández-Pinzón, I., and Patel, K. (2001). Role of lipid bodies and lipid-body proteins in seeds and other tissues. Journal of Plant Physiology, 158(4), 471-478.

Murray, M. G., and Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. Nucleic Acids Research, 8(19), 4321-4326.

Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. S., Zhang, Z., Costich, D. E., and Buckler, E. S. (2009). Association mapping: critical considerations shift from genotyping to experimental design. The Plant Cell, 21(8), 2194-2202.

Nelson, F. K., and Riddle, D. L. (1984). Functional study of the *Caenorhabditis elegans* secretory-excretory system using laser microsurgery. Journal of Experimental Zoology Part A: Ecological Genetics and Physiology, 231(1), 45-56.

Nicaise, V., Joe, A., Jeong, B., Korneli, C., Boutrot, F., Westedt, I., Staiger, D., Alfano, J. R., Zipfel, C. (2013). Pseudomonas HopU1 modulates plant immune receptor levels by blocking the interaction of their mRNAs with GRP7. The EMBO Journal, 32: 701-712.

Nicol, J. M., Turner, S. J., Coyne, D. L., Den Nijs, L., Hockland, S., and Maafi, Z. T. (2011). Current nematode threats to world agriculture. In Genomics and Molecular Genetics of Plant-nematode Interactions, Netherlands, pp. 21-43.

Nordborg, M., and Tavaré, S. (2002). Linkage disequilibrium: what history has to tell us. Trends in Genetics, 18(2), 83-90.

Nordborg, M., Borevitz, J. O., Bergelson, J., Berry, C. C., Chory, J., Hagenblad, J., and Stahl, E. A. (2002). The extent of linkage disequilibrium in *Arabidopsis thaliana*. Nature Genetics, 30(2), 190-193.

Nyle, C. B., and Weil, R. R. (2009). Elements of the Nature and Properties of Soils (3rd Edition). Prentice Hall. ISBN 9780135014332.

Oh, I. S., Park, A. R., Bae, M. S., Kwon, S. J., Kim, Y. S., Lee, J. E., and Park, O. K. (2005). Secretome analysis reveals an Arabidopsis lipase involved in defense against *Alternaria brassicicola*. The Plant Cell, 17(10), 2832-2847.

Orion, D., and Franck, A. (1990). An electron microscopy study of cell wall lysis by *Meloidogyne javanica* gelatinous matrix. Revue de Nematologie, 13, 105-108.

Palaisa, K. A., Morgante, M., Williams, M., and Rafalski, A. (2003). Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. The Plant Cell, 15(8), 1795-1806.

Pařenicová, L., de Folter, S., Kieffer, M., Horner, D. S., Favalli, C., Busscher, J., and Angenent, G. C. (2003). Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis. The Plant Cell Online, 15(7), 1538-1551.

Pariyar, S. R., Dababat, A. A., Sannemann, W., Erginbas-Orakci, G., Elashry, A., Siddique, S., and Grundler, F. M. (2016). Genome-wide association study in wheat identifies resistance to the cereal cyst nematode *Heterodera filipjevi*. Phytopathology, 106(10), 1128-1138.

Park, A. R, Cho, S. K., Yun, U.J., Jin, M. Y., Lee, S. H. and Sachetto-Martins, G. (2001). Interaction of the Arabidopsis receptor protein kinase Wak1 with a glycine-rich protein, AtGRP-3. Journal of Biological Chemistry, 276, 26688-26693.

Park, A. R., Somi, K. C., Yun, U. J., Jin, M. Y., Lee, S. H., Sachetto-Martins, G., and Park, O. K. (2001). Interaction of the Arabidopsis receptor protein kinase Wak1 with a glycine-rich protein, AtGRP-3. Journal of Biological Chemistry, 276, 26688-26693.

Park, J. H., Suh, M. C., Kim, T. H., Kim, M. C., Cho, S. H. (2008). Expression of glycine-rich protein genes, AtGRP5 and AtGRP23, induced by the cutin monomer 16-hydroxypalmitic acid in *Arabidopsis thaliana*. Plant Physiology and Biochemistry, 46: 1015-1018.

Pate, J. S., and Gunning, B. E. S. (1972). Transfer cells. Annual Review of Plant Physiology, 23,173-196.

Paterson, A. H., Lander, E. S., Hewitt, J. D., Peterson, S., Lincoln, S. E., and Tanksley, S. D. (1988). Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature, 335(6192), 721-726.

Pe´rier, R. C., Junier, T., and Bucher, P. 1998. The eukaryotic promoter database EPD. Nucleic Acids Research, 26, 353-357.

Perry, R. N., and Gaur, H. S. (1996). Host plant influences on the hatching of cyst nematodes. Fundamental and Applied Nematology, 19(6), 505-510.

Perry, R. N., Wright, D. J. The Physiology and Biochemistry of Free-Living and Plant-Parasitic Nematodes. CAB International, UK, pp.75-102.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Research, 29(9), e45.

Plomin, R., Haworth, C. M., and Davis, O. S. (2009). Common disorders are quantitative traits. Nature Reviews Genetics, 10(12), 872-878.

Porto, M. S., Pinheiro, M. P. N., Batista, V. G. L., dos Santos, R. C., de Albuquerque MeloFilho, P., and de Lima, L. M. (2014). Plant promoters: an approach of structure and function. Molecular Biotechnology, 56(1), 38-49.

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics, 38(8), 904-909.

Pullin, A. P. (2002). Conservation Biology. Cambridge University Press, UK.

Puthoff, D. P., Nettleton, D., Rodermel, S. R., and Baum, T. J. (2003). Arabidopsis gene expression changes during cyst nematode parasitism revealed by statistical analyses of microarray expression profiles. Plant Journal, 33, 911-921.

Rabiei, B. (2007). Linkage map of SSR markers and QTLs detection for heading date of Iranian rice cultivars, pp. 235-242.

Rabiei, B. (2010). Linkage map of SSR markers and QTLs detection for heading date of Iranian rice cultivars. Journal of Agricultural Science and Technology, 9, 235-42.

Rafalski, J. A. (2010). Association genetics in crop improvement. Current Opinion in Plant Biology, 174-180.

Reyes, J. C., Muro-Pastor, M. I., and Florencio, F. J. (2004). The GATA family of transcription factors in Arabidopsis and rice. Plant Physiology, 134(4), 1718-1732.

Riechmann, J. L., and Ratcliffe, O. J. (2000). A genomic perspective on plant transcription factors. Current Opinion in Plant Biology, 3(5), 423-434.

Risch, N., and Merikangas, K. (1996). The future of genetic studies of complex human diseases. Science, 273(5281), 1516-1517.

Rodiuc, N., Vieira, P., Banora, M.Y., and de Almeida Engler, J. (2014). On the track of transfer cell formation by specialized plant-parasitic nematodes.Frontiers in Plant Science, 5, 160.

Rohde, W., Rosch, K., Kroger, K., and Salamini, F. (1990). Nucleotide sequence of a *Hordeum vulgare* gene encoding a glycine-rich protein with homology to vertebrate cytokeratins.Plant Molecular Biology, 14, 1057-1059.

Ross, G. J. S., and Trudgill, D. L. (1969). The effect of population density on the sex ratio of *Heterodera rostochiensis*; a two dimensional model. Nematologica, 15(4), 601-607.

Ross, J. P. (1964). Effect of soil temperature on development of *Heterodera glycines* in soyabeanroots. Phytopathology, 54(1), 228-31.

Ruppert, E. E. B., Fox, R. D., Ruppert, R. S. E. E., Fox, R. S., and Barnes, R. D. (2004). Invertebrate zoology: a functional evolutionary approach, No. 592 RUPi.

Russell P. J. (1996). Genetics, 4th edn. Harper Collins College, New York.

Sachetto-Martins, G., Fernandes, L. D., Felix, D. B., and de Oliveira, D. E. (1995). Preferential transcriptional activity of a glycine-rich protein gene from *Arabidopsis thaliana* in protoderm-derived cells. International Journal of Plant Sciences, 156(4), 460-470.

Sachetto-Martins, G., Franco, L. O., and de Oliveira, D. E. (2000). Plant glycine-rich proteins: a family or just proteins with a common motif? Biochimicaet BiophysicaActa (BBA)-Gene Structure and Expression, 1492(1), 1-14.

Santos, J. R. P., Ndeve, A. D., Huynh, B. L., Matthews, W. C., and Roberts, P. A. (2018). QTL mapping and transcriptome analysis of cowpea reveals candidate genes for root-knot nematode resistance. PloS one, 13(1).

Sawant, S., Singh, P. K., Madanala, R., and Tuli, R. (2001). Designing of an artificial expression cassette for the high-level expression of transgenes in plants. TAG Theoretical and Applied Genetics, 102(4), 635-644.

Scheck, B., Neufeld, P., and Dwyer, J. (2000). Actual Innocence: Five Days to Execution and other Dispatches from Wrongly Convicted. New American Library, New York.

Schenk, P. M., Kazan, K., Manners, J. M., Anderson, J. P., Simpson, R. S., Wilson, I. W., and Maclean, D. J. (2003). Systemic gene expression in Arabidopsis during an incompatible interaction with *Alternaria brassicicola*. Plant Physiology, 132(2), 999-1010.

Schlang, J. (1991). Anbau resistenter Zwischenfrüchte zur biologischen Bekämpfung des Rübenzystennematoden. Zuckerrübe, 40, 476-488.

Schmidt, A., (1871). Ueber den Ruben-Nematodenzeitsdereft fur Rubenzucher. Industrie. 21: 1-19.

Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative CT method. Nature Protocols, 3(6), 1101-1108.

Scholl, R. L., May, S. T., and Ware, D. H. (2000). Seed and molecular resources for Arabidopsis. Plant Physiology, 124(4), 1477-1480.

Seren, Ü.,Vilhjálmsson, B. J., Horton, M. W., Meng, D., Forai, P., Huang, Y. S., and Nordborg, M. (2012). GWAPP: a web application for genome-wide association mapping in Arabidopsis. The Plant Cell, 24(12), 4793-4805.

Sharmin, S., Azam, M. S., Islam, M. S., Sajib, A. A., Mahmood, N., Hasan, A. M., and Khan, H. (2012). Xyloglucan endotransglycosylase/hydrolase genes from a susceptible and resistant jute species show opposite expression pattern following *Macrophomina phaseolina* infection. Communicative and Integrative Biology, 5(6), 598-606.

Shi, Z., Liu, S., Noe, J., Arelli, P., Meksem, K., and Li, Z. (2015). SNP identification and marker assay development for high-throughput selection of soybean cyst nematode resistance. BMC genomics, 16(1), 314.

Shindo, C., Bernasconi, G., and Hardtke, C. S. (2007). Natural genetic variation in Arabidopsis: tools, traits and prospects for evolutionary ecology. Annals of Botany, 99(6), 1043-1054.

Shirasawa, K., Hirakawa, H., Nunome, T., Tabata, S., and Isobe, S. (2016). Genome-wide survey of artificial mutations induced by ethyl methanesulfonate and gamma rays in tomato.Plant biotechnology journal, 14(1), 51-60.

Showalter, A. M., Zhou, J., Rumeau, D., Worst, S. G., and Varner, J. E. (1991). Tomato extensin and extensin-like cDNAs: Structure and expression in response to wounding. Plant Molecular Biology, 16, 547-565.

Siddiqi, M. R. (2000). Tylenchida: parasites of plants and insects, 2nd edition, CAB International, UK, pp-833.

Siddique, S., and Grundler, F.M.W. (2015). Metabolism in nematodes: In Plant-nematode interactions. Elsevier, Editors: Carolina Escobar and Carmen Fenoll, pp. 119-138.

Sijmons, P. C., Atkinson, H. J., Wyss, U. (1994). Parasitic strategies of root nematodes and associated host-cell responses. Annual Reviewof Phytopathology, 32, 235-259.

Sijmons, P. C., Grundler, F. M., Mende, N., Burrows, P. R., and Wyss, U. (1991). *Arabidopsis thaliana* as a new model host for plant-parasitic nematodes. The Plant Journal, 1(2), 245-254.

Slatkin, M. 2008. Linkage disequilibrium-understanding the evolutionary past and mapping the medical future. Nature Reviews Genetics, 9(6), 477-485.

Smith, R. C., and Fry, S. C. (1991). Endotransglycosylation of xyloglucans in plant cell suspension cultures. Biochemical Journal, 279(2), 529-535.

Sobczak, M., Golinowski, W., Grundler, F. M. W. (1997).Changes in the structure of *Arabidopsis thaliana* roots induced during development of males of the plant parasitic nematode *Heterodera schachtii*. European Journal of Plant Pathology, 103, 113-124.

Soller, M., Brody, T., and Genizi, A. (1976). On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. Theoretical and Applied Genetics, 47(1), 35-39.

Sonah, H., Bastien, M., Iquira, E., Tardivel, A., Légaré, G., Boyle, B., and Belzile, F. (2013).An improved genotyping by sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. PloS One, 8(1), e54603.

Starr, J. L. and Mercer, C.F. (2009). Development of resistant varieties. In: Perry, R.N., Moens, M. and Starr, J. L., eds. Root knot nematodes. CAB International, UK, pp. 326-337

Steele, A. E. (1971). Orientation and development of *Heterodera schachtii* larvae on tomato and sugar beet roots. Journal of Nematology, 3,424-426.

Stepek, G., Buttle, D. J., Duce, I. R., and Behnke, J. M. (2006). Human gastrointestinal nematode infections: Are new control methods required? International Journal of Experimental Pathology, 87(5), 325-341.

Stirling, G. R. (1991). Biological control of plant parasitic nematodes: progress, problems and prospects. CAB International. UK, pp-212.

Stuber, C. W., Lincoln, S. E., Wolff, D. W., Helentjaris, T., and Lander, E. S. (1992). Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics, 132(3), 823-839.

Symonds, V. V., Godoy, A. V., Alconada, T., Botto, J. F., Juenger, T. E., Casal, J. J., and Lloyd, A. M. (2005). Mapping quantitative trait loci in multiple populations of *Arabidopsis thaliana* identifies natural allelic variation for trichome density. Genetics, 169(3), 1649-1658.

Szakasits, D., Heinen, P., Wieczorek, K., Hofmann, J., Wagner, F., Kreil, D.P., Sykacek, P., Grundler, F.M.W., and Bohlmann H. (2009). The transcriptome of syncytia induced by the cyst nematode *Heterodera schachtii* in Arabidopsis roots. Plant Journal, 57, 771-784.

Tanksley, S. D. (1993). Mapping polygenes.Annual review of genetics, 27(1), 205-233.

Tao, T. Y., Ouellet, T., Dadej, K., Miller, S. S., Johnson, D. A., Singh, J. (2006). Characterization of a novel glycine-rich protein from the cell wall of maize silk tissues. Plant Cell Reports, 25: 848-858.

Teakle, G. R., Manfield, I. W., Graham, J. F., and Gilmartin, P. M. (2002). *Arabidopsis thaliana* GATA factors: organisation, expression and DNA-binding characteristics. Plant Molecular Biology, 50(1), 43-56.

Tian, F., Bradbury, P. J., Brown, P. J., Hung, H., Sun, Q., Flint-Garcia, S., and Buckler, E. S. (2011). Genome-wide association study of leaf architecture in the maize nested association mapping population. Nature Genetics, 43(2), 159-162.

Toledo-Ortiz, G., Huq, E., and Quail, P. H. (2003). The Arabidopsis basic/helix-loop-helix transcription factor family. The Plant Cell, 15(8), 1749-1770.

Triantaphyllou, A. C. (1973). Environmental sex differentiation of nematodes in relation to pest management. Annual Review of Phytopathology, 11(1), 441-462.

Triantaphyllou, A. C., and Hirschmann, H. (1960). Post infection development of *Meloidogyne incognita* Chitwood 1949 (Nematoda-*Heteroderidae*). In Annales de l'Institut Phytopathologique Benaki, 3(1), 1-11.

Triantaphyllou, A. C., and Hirschmann, H. (1964). Reproduction in plant and soil nematodes. Annual Review of Phytopathology, 57-80.

Trudgill, D. L., Blok, V. C., Bala, G., Daudi, A., Davies, K. G., Gowen, S. R., and Netscher, C. (2000). The importance of tropical root-knot nematodes (*Meloidogyne spp*.) and factors affecting the utility of Pasteuriapenetrans as a biocontrol agent. Nematology, 2(8), 823-845.

Tyler, J. (1933). Reproduction without males in aseptic root cultures of the root-knot nematode. California Agriculture, 7(10), 373-388.

Ueki, S., and Citovsky, V. (2002). The systemic movement of a tobamovirus is inhibited by a cadmium-ion-induced glycine-rich protein. Nature cell biology, 4(7), 478-486.

Vidal, R. O., Nascimento, L. C. D., Maurício Costa Mondego, J., Amarante Guimarães Pereira, G., and Falsarella Carazzolle, M. (2012). Identification of SNPs in RNA-seq data of two cultivars of Glycine max (soybean) differing in drought resistance. Genetics and molecular biology, 35(1), 331-334.

Vijayan, J., Devanna, B. N., Singh, N. K., and Sharma, T. R. (2015). Cloning and functional validation of early inducible *Magnaporthe oryzae* responsive CYP76M7 promoter from rice. Frontiers in plant science, 6, 371.

Vuong, T. D., Sonah, H., Meinhardt, C. G., Deshmukh, R., Kadam, S., Nelson, R. L., and Nguyen, H. T. (2015). Genetic architecture of cyst nematode resistance revealed by genome-wide association study in soybean. BMC Genomics, 16(1), 593.

Walker, L. R., and Del Moral, R. (2003). Primary succession and ecosystem rehabilitation. Cambridge University Press, New York.

Wang, J., McClean, P. E., Lee, R., Goos, R. J., and Helms, T. (2008). Association mapping of iron deficiency chlorosis loci in soybean (Glycine max L. Merr.) advanced breeding lines. Theoretical and Applied Genetics, 116(6), 777-787.

Wang, Q., Wan, L., Li, D., Zhu, L., Qian, M., and Deng, M. (2009). Searching for bidirectional promoters in *Arabidopsis thaliana*. Bmc Bioinformatics, 10(1), S29.

Wang, X., Meyers, D., Yan, Y., Baum, T., Smant, G., Hussey, R., and Davis, E. (1999). In planta localization of a β-1, 4-endoglucanase secreted by *Heterodera glycines*. Molecular Plant-Microbe Interactions, 12(1), 64-67.

Wei, G.-H., Liu, D.-P, and Liang, C. C. (2004). Charting gene regulatory networks: strategies, challenges and perspectives. Biochemical Journal, 381(1), 1-12.

Weinstein, P. P. (1952). Regulation of water balance as a function of the excretory system of the filariform larvae of Nippostrongylusmuris and Ancylostomacaninum. Experimental Parasitology, 1(4), 363-376.

Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A., and Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research, 18(22), 6531-6535.

Williamson, V. M., Hussey, R. S. (1996). Nematode pathogenesis and resistance in plants. Plant Cell, 8:1735-1745.

Wrather, J. A., Anderson, T. R., Arsyad, D. M., Tan, Y., Ploper, L. D., Porta-Puglia, A., and Yorinori, J. T. (2001). Soybean disease loss estimates for the top ten soybean-producing counries in 1998. Canadian Journal of Plant Pathology, 23(2), 115-121.

Wyss, U. (1997). Root parasitic nematodes: an overview. In: Fenoll, C, Grundler, F. M. W., and Ohl, S.A., eds. Cellular and molecular aspects of plant-nematode interactions. Kluwer Academic, Netherlands, pp. 5-22.

Wyss, U., and Grundler, F. M. W. (1992). Feeding behavior of sedentary plant parasitic nematodes. Netherlands Journal of Plant Pathology, 98(2), 165-173.

Xi, J., Qiu, Y., Du, L., and Poovaiah, B. W. (2012). Plant-specific trihelix transcription factor *AtGT2L* interacts with calcium/calmodulin and responds to cold and salt stresses. Plant Science, 185, 274-280.

Xia, Y., Li, R., Ning, Z., Bai, G., Siddique, K. H., Yan, G., Baum, M., Varshney, R. K., and Guo, P. (2013). Single nucleotide polymorphisms in HSP17.8 and their association with agronomic traits in barley.PLoS One, 8 (2), e56816.

Yang, L., Lin, C., Jin, C., Yang, J. C., Tanasa, B., Li, W., and Evans, C. P. (2013). lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. Nature, 500(7464), 598.

Yeates, G.W., Bongers, T., de Goede, R. G. M., Freckman, D.W., Georgieva, S. S. (1993a). Feeding habits in soil nematode families and genera - an outline for ecologists. Journal of Nematology, 25, 315-331.

Yi, N., Oh, S. J., Kim, Y. S., Jang, H. J., Park, S. H., Jeong, J. S., and Kim, J. K. (2011). Analysis of the Wsi18, a stress-inducible promoter that is active in the whole grain of transgenic rice. Transgenic research, 20(1), 153-163.

Yokoyama, R., and Nishitani, K. (2006). Identification and characterization of *Arabidopsis thaliana* genes involved in xylem secondary cell walls. Journal of plant research, 119(3), 189-194.

Young, K. S. (1996). Internet Addiction: The Emergence of a New Clinical Disorder. Cyber Psychology and Behavior, 1, 237-244.

Youssef, R. M., MacDonald, M. H., Brewer, E. P., Bauchan, G. R., Kim, K. H., and Matthews, B. F. (2013). Ectopic expression of AtPAD4 broadens resistance of soybean to soybean cyst and root-knot nematodes. BMC Plant Biology, 13(1), 67.

Yu, J., Holland, J. B., McMullen, M. D., and Buckler, E. S. (2008). Genetic design and statistical power of nested association mapping in maize. Genetics, 178(1), 539-551.

Zeng, Z. B. (1993). Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. Proceedings of the National Academy of Sciences, 90(23), 10972-10976.

Zeng, Z. B. (1994). Precision mapping of quantitative trait loci. Genetics, 136(4), 1457-1468.

Zhang, H., Li, C., Davis, E. L., Wang, J., Griffin, J. D., Kofsky, J., and Song, B. H. (2016). Genome-wide association study of resistance to soybean cyst nematode (*Heterodera glycines*) HG Type 2.5. 7 in wild soybean (*Glycine soja*). Frontiers in plant science, 7, 1214.

Zhang, Z. Q. (2011). Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness. Zootaxa, 3148, 1-237.

Zhu, C., Gore, M., Buckler, E. S., and Yu, J. (2008). Status and prospects of association mapping in plants. The Plant Genome, 1(1), 5-20.

Zou, C., Sun, K., Mackaluso, J. D., Seddon, A. E., Jin, R., Thomashow, M. F., and Shiu, S. H. (2011). Cis-regulatory code of stress-responsive transcription in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences, 108(36), 14992-14997.

Zunke, U. (1991). Observations on the invasion and endo-parasitic behavior of the root lesion nematode *Pratylenchus penetrans*. Journal of Nematology, 22, 309-320.

Zunke, U., and Eisenback, J. D. (1998). Morphology and ultrastructure. In: Sharma S.B., eds. The Cyst nematodes. London, UK, pp. 31-56.

7. Annex

Table S1: % of phenotypic data for the average no. of female to male sex ratio of nematodes per plant (Section 3.4: from chapter 3. Phenotypic histogram for the average female to male sex ratio of nematodes/plant)

Codes	Accessions	%sex ratio	Codes	Accessions	%sex ratio
6929	Kondara	0.255814	7255	Mh-0	0.789474
7351	Ty-0	0.339181	6909	Col-0	0.793302
8389	Ta-0	0.351724	6981	Ws-2	0.794872
7378	Uk-1	0.365385	6987	Ak-1	0.8
7143	Gel-1	0.37234	9230	Del-10	0.805825
7169	Hh-0	0.392157	9169	Kastel-1	0.813433
7075	Cit-0	0.4	5752	Lan-1	0.81893
7199	Kl-5	0.403101	8296	Gd-1	0.819355
6977	Van-0	0.409396	8374	Rsch-4	0.821053
7355	Tiv-1	0.441558	9806	Star-8	0.836066
6970	Ts-1	0.447005	6992	Ang-0	0.836066
7514	RRS-7	0.449541	6975	Uod-1	0.852941
6915	Ei-2	0.462687	5836	B002-3	0.865248
7282	or-0	0.464088	6951	Pu2-23	0.866667
7250	Me-0	0.464455	7460	Da(1)-12	0.869159
9065	xan-1	0.465517	9800	Agu-1	0.87234
6926	Kin-0	0.469697	8215	Fei-0	0.874074
7244	Mnz-0	0.474026	7520	Lp2-2	0.876543
6956	Pu2-7	0.476471	7258	Nw-0	0.878049
7306	Pog-0	0.497908	9230	Del-0	0.88
6939	Mt-0	0.517647	8310	Hs-0	0.885246
7430	Nc-1	0.520325	8334	Lu-1	0.885246
7165	Hn-0	0.52514	6905	Bur-0	0.886228
7333	sei-0	0.529762	8420	Kelstrbch-4	0.888889
7384	Ven-1	0.531532	7069	Cha-0	0.907895
7126	Es-0	0.559441	6900	Bil-5	0.913043
6968	Tamm-2	0.567073	6958	Ra-0	0.94
7031	Bsch-0	0.567708	7320	Rou-0	0.941176
8388	Stw-0	0.571895	2274	SLSP-30	0.94152
6922	Gu-o	0.573034	7352	Te-0	0.953642
7223	Li-2:1	0.573427	8213	Pro-0	0.964912
7524	Rmx-Ao2	0.578125	6961	Se-0	0.978723
6990	Amel-1	0.58	7283	Ors-1	0.984848
7404	Wc-1	0.58209	6929	Kondara	1
6898	An-1	0.584615	8233	Dem-4	1.004695
7330	sapporo-1	0.585859	7280	Old-1	1.017241
7250	Me-0	0.590643	7337	Si-0	1.028169
7000	Aa-0	0.598394	9118	Bak-2	1.055866
7300	Pla-0	0.602041	6936	Lz-0	1.058824
7013	Bd-0	0.604278	7121	En-T	1.085502
7323	Rubezhnoe1	0.608434	9169	kastel-1	1.098901
6919	Ga-0	0.611111	9127	Yeg-1	1.102362
9343	dja-1	0.612245	7523	Pna-17	1.12069
6979	Wei-0	0.617021	8325	Lip-0	1.121951
7382	Utrecht	0.61745	6978	Wa-1	1.129032
7244	Mnz-0	0.619048	9122	Bak-7	1.12987
7268	Np-0	0.621302	9179	Ayu-Dag-3	1.130045

5805	UKID.101	0.628743	6967	Sq-0	1.131579
7062	Ca-0	0.636364	6910	Ct-1	1.147783
7014	Ba-1	0.639456	7147	Gie-0	1.153846
6904	Br-0	0.641791	7276	Ob-0	1.15625
8395	Tu-0	0.643216	7192	Kil-0	1.176471
6994	Ann-1	0.651316	6983	Yo-0	1.190476
7430	Nc-1	0.662651	6897	Ag-0	1.195652
7354	Ting-1	0.683908	8366	Rd-0	1.2
7106	Dr-0	0.684211	6982	Wt-5	1.208333
7002	Baa-1	0.693252	8245	Seattle-0	1.227273
8271	Bu-0	0.699029	7479	PHW-10	1.231293
6940	Oy-0	0.713816	6709	Bg-2	1.261905
7372	Tscha-1	0.720721	8378	Sap-0	1.265306
7344	Sg-1	0.72093	9805	Sha	1.3
6938	Ms-0	0.735294	8312	Is-0	1.309091
7166	Hey-1	0.740741	7411	Wl-0	1.351351
7015	Bla-1	0.747253	9806	vie-0	1.363636
7071	Chat-1	0.75	8300	Gr-1	1.4
7343	sp-0	0.75	7163	Ha-0	1.514286
8354	Per-1	0.75	9806	Tuescha-9	1.553571
2290	Ste-3	0.758621	7252	Mc-0	1.574468
7472	S96	0.758929	8329	Lm-2	1.6
6931	Kz-9	0.772727	6984	Zdr-1	1.608696
7424	Jl-3	0.775	8313	Jm-0	1.641026
9805	Lag2-2	0.776316	7206	Kro-0	1.65
7524	Rmx-A02	0.777778	8343	Na-1	2.857143
8314	Ka-0	0.782609			
6945	Nok-3	0.782609			

Table S2: % of phenotypic data for the average no. of male nematodes per plant (Section 3.3 : from chapter 3. Phenotypic histogram for the average no. of male nematodes/plant)

Codes	Accessions	% males	Codes	Accessions	% males
8343	Na-1	38.18182	8213	Pro-0	114
6945	Nok-3	46	7330	sapporo-1	115.1163
6984	Zdr-1	51.11111	7280	Old-1	116
7106	Dr-0	54.28571	7523	Pna-17	116
6978	Wa-1	56.36364	6987	Ak-1	116.6667
7524	Rmx-A02	60	6961	Se-0	117.5
6929	Kondara	61.66667	7337	Si-0	118.3333
8388	Stw-0	63.25581	6940	Oy-0	118.6787
7352	Te-0	68.01802	7472	S96	119.2465
6967	Sq-0	69.09091	7166	Hey-1	120
9118	Bak-2	69.87988	7223	Li-2:1	120.9302
8366	Rd-0	70	6898	An-1	120.9302
8334	Lu-1	71.44144	8310	Hs-0	122
7479	PHW-10	71.73423	7411	Wl-0	123.3333
9805	Sha	72.72727	8354	Per-1	123.6364
7206	Kro-0	72.72727	9343	dja-1	124.3129
9806	vie-0	73.33333	9805	Lag2-2	126.6667
8245	Seattle-0	73.33333	7069	Cha-0	126.6667
6931	Kz-9	73.33333	9122	Bak-7	128.3333
8325	Lip-0	74.54545	6994	Ann-1	128.5412
6938	Ms-0	75.55556	7372	Tscha-1	129.0698
7015	Bla-1	76.9556	7384	Ven-1	129.0698
8313	Jm-0	78	6981	Ws-2	130
7252	Mc-0	78.33333	7075	Cit-0	130.2326
6910	Ct-1	79.24925	7300	Pla-0	130.2326
7276	Ob-0	80	6897	Ag-0	131.4286
8314	Ka-0	80.23256	7126	Es-0	133.0233
8378	Sap-0	81.66667	6929	Kondara	133.3333
8233	Dem-4	83.15315	7199	Kl-5	133.3333
6983	Yo-0	84	9800	Agu-1	134.2857
7283	Ors-1	85.88589	7520	Lp2-2	135
9230	Del-10	86.16474	7244	Mnz-0	139.5349
6982	Wt-5	87.27273	7268	Np-0	142.9175
7163	Ha-0	87.5	7344	Sg-1	143.3333
8329	Lm-2	87.5	6956	Pu2-7	143.7632
8215	Fei-0	87.83784	7378	Uk-1	145.1163
5805	UKID.101	88.90254	7354	Ting-1	147.1459
6904	Br-0	89.03654	7165	Hn-0	151.3742
6900	Bil-5	89.78979	7002	Baa-1	151.6279
8420	Kelstrbch-4	90	9169	kastel-1	151.6667
7258	Nw-0	91.11111	7255	Mh-0	152
8312	Is-0	91.66667	9806	Star-8	152.5
6922	Gu-o	91.98966	6968	Tamm-2	152.5581
6709	Bg-2	92.22973	6926	Kin-0	153.4884
6990	Amel-1	93.02326	7323	Rubezhnoe1	154.4186
9806	Tuescha-9	93.33333	6979	Wei-0	156.6667
7147	Gie-0	94.54545	7143	Gel-1	158.9852
6905	Bur-0	97.79279	7250	Me-0	159.0698
9127	Yeg-1	99.15916	7244	Mnz-0	159.1731
6939	Mt-0	99.54955	9065	xan-1	165.7143

6958	Ra-0	100	6919	Ga-0	167.4419
6909	Col-0	100	8395	Tu-0	168.2875
8300	Gr-1	100	8389	Ta-0	168.6047
6909	col-0	100	7424	Jl-3	169.1332
2274	SLSP-30	100.1351	7404	Wc-1	169.9789
5752	Lan-1	101.6409	7169	Hh-0	172.5159
6992	Ang-0	101.6667	6977	Van-0	173.2558
2290	Ste-3	101.8919	7382	Utrecht	173.2558
7320	Rou-0	102	7013	Bd-0	173.9535
7355	Tiv-1	102.3256	8271	Bu-0	174.2072
7121	En-T	105.015	7062	Ca-0	175.4153
7343	sp-0	106.6667	7351	Ty-0	176.7442
7524	Rmx-Ao2	108.2452	8296	Gd-1	180.2326
9179	Ayu-Dag-3	108.8213	6970	Ts-1	183.5095
5836	B002-3	110.0901	7514	RRS-7	184.3552
7430	Nc-1	110.299	7250	Me-0	196.2791
7460	Da(1)-12	110.5943	7031	Bsch-0	198.4496
9230	Del-0	111.1111	7306	Pog-0	202.1142
8374	Rsch-4	111.2613	6915	Ei-2	207.7519
6951	Pu2-23	112.5	7282	or-0	210.4651
9169	Kastel-1	113.3192	7000	Aa-0	210.5708
6975	Uod-1	113.3333	7333	sei-0	223.2558
7071	Chat-1	113.3333	7014	Ba-1	248.6258
7192	Kil-0	113.3333	7430	Nc-1	254.2636
6936	Lz-0	113.3333			

Table S3: % of phenotypic data for the average no. of female nematodes per plant (Section 3.2 : from chapter 3. Phenotypic histogram for the average no. of female nematodes/plant)

Codes	Accessions	% females	Codes	Accessions	% females
6945	Nok-3	34.83871	9800	Agu-1	113.3641
7106	Dr-0	35.9447	7280	Old-1	114.1935
7524	Rmx-A02	45.16129	7524	Rmx-Ao2	114.5068
6938	Ms-0	53.76344	7520	Lp2-2	114.5161
6931	Kz-9	54.83871	8314	Ka-0	114.8936
6939	Mt-0	58.66667	7255	Mh-0	116.129
6929	Kondara	59.67742	7206	Kro-0	116.129
6978	Wa-1	61.58358	8312	Is-0	116.129
6929	Kondara	62.41135	6936	Lz-0	116.129
5805	UKID.101	63.63636	7337	Si-0	117.7419
8388	Stw-0	66.19385	7252	Mc-0	119.3548
8334	Lu-1	72	9806	Star-8	123.3871
7352	Te-0	73.84615	7330	sapporo-1	123.4043
9065	xan-1	74.65438	7169	Hh-0	123.7911
6967	Sq-0	75.65982	8313	Jm-0	123.871
7343	sp-0	77.41935	9127	Yeg-1	124.4444
7258	Nw-0	77.41935	6956	Pu2-7	125.3385
8420	Kelstrbch-4	77.41935	7384	Ven-1	125.5319
9230	Del-10	79.04762	7523	Pna-17	125.8065
6984	Zdr-1	79.56989	7223	Li-2:1	126.8859
8325	Lip-0	80.93842	7163	Ha-0	128.2258
8366	Rd-0	81.29032	7192	Kil-0	129.0323
7071	Chat-1	82.25806	6898	An-1	129.3617
6992	Ang-0	82.25806	7121	En-T	129.7778
7355	Tiv-1	82.67477	6977	Van-0	129.7872
9118	Bak-2	84	6926	Kin-0	131.9149
7166	Hey-1	86.02151	6709	Bg-2	132.5
8245	Seattle-0	87.09677	7430	Nc-1	133.7386
8215	Fei-0	87.40741	8329	Lm-2	135.4839
2290	Ste-3	88	8300	Gr-1	135.4839
7276	Ob-0	89.51613	7126	Es-0	136.1702
8354	Per-1	89.73607	6909	col-0	137.7445
6987	Ak-1	90.32258	7244	Mnz-0	138.0615
6958	Ra-0	90.96774	9343	dja-1	139.265
9805	Sha	91.4956	9179	Ayu-Dag-3	140
7320	Rou-0	92.90323	9122	Bak-7	140.3226
6900	Bil-5	93.33333	9806	Tuescha-9	140.3226
6975	Uod-1	93.54839	7300	Pla-0	143.465
6979	Wei-0	93.54839	7165	Hn-0	145.4545
6951	Pu2-23	94.35484	6970	Ts-1	150.0967
9230	Del-0	94.62366	7514	RRS-7	151.6441
5752	Lan-1	94.7619	6897	Ag-0	152.0737
8233	Dem-4	95.11111	6994	Ann-1	153.1915
9805	Lag2-2	95.16129	7244	Mnz-0	158.0547
7075	Cit-0	95.31915	6968	Tamm-2	158.2979
7283	Ors-1	96.2963	7411	Wl-0	161.2903
6940	Oy-0	96.44444	9169	kastel-1	161.2903
6922	Gu-o	96.4539	7268	Np-0	162.4758
6983	Yo-0	96.77419	7250	Me-0	166.8085
9806	vie-0	96.77419	9169	Kastel-1	168.6654

7378	Uk-1	97.02128	7372	Tscha-1	170.2128
7199	Kl-5	98.34515	7323	Rubezhnoe1	171.9149
6905	Bur-0	98.66667	7250	Me-0	171.9149
6990	Amel-1	98.7234	7460	Da(1)-12	175.8865
6909	Col-0	100	6915	Ei-2	175.8865
6981	Ws-2	100	7282	or-0	178.7234
7344	Sg-1	100	7404	Wc-1	181.0445
8378	Sap-0	100	7354	Ting-1	184.1393
7479	PHW-10	100.5556	7306	Pog-0	184.1393
6982	Wt-5	102.0528	6919	Ga-0	187.234
7472	S96	103.0303	7013	Bd-0	192.3404
6910	Ct-1	103.5556	7002	Baa-1	192.3404
8374	Rsch-4	104	7382	Utrecht	195.7447
8310	Hs-0	104.5161	8395	Tu-0	198.0658
6904	Br-0	104.5593	7062	Ca-0	204.2553
7015	Bla-1	105.2224	7031	Bsch-0	206.1466
8343	Na-1	105.5718	7333	sei-0	216.4134
7147	Gie-0	105.5718	8271	Bu-0	222.824
8213	Pro-0	106.4516	7000	Aa-0	230.5609
2274	SLSP-30	107.3333	7424	Jl-3	239.8453
7143	Gel-1	108.3172	7430	Nc-1	242.0804
5836	B002-3	108.4444	8296	Gd-1	270.2128
8389	Ta-0	108.5106	7014	Ba-1	290.9091
7351	Ty-0	109.6927			
6961	Se-0	111.2903			
7069	Cha-0	111.2903			

AtS40-3 DNA sequences among lowly susceptible accessions

Kondara	ATGTCAGAAGAATTTCAAGAATCTGAGGTTATTTTTTCCGATGAGTCTTTCACAAGGAAAGATAACAAGATCAGTCACAACAACGAAAACTACGAAAGAAAGTCG ACGGAGAAGGAGAAAATTTCATCTCCGGTGAGAATTCCGTCAAGAACTACTATCCGGTATACGGAAGAGGAGGGAGAAATGACACCGCCACATGTCATAATCGA AAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACACCGTAACACCGTTCTTAGGATGACCGGCTTTTTGGA AGTTTA-
Ta-0	ATGTCAGAAGAATTTCAAGAATCCGAGGTTATATTTTCCGATGAGTCTTTCACAAAGAACGATAACAAGATCAGTCATAACTACGAAAACTACGAAAGAAAGTCG ACGGAGAAGGAGAAAATTTCATCTCCGGTGAAAATTCCGTCAAGAACTACTTTCCGGTATACGGAAGAGGAGGGAGAAATGAAACCGCCACATGTCATAATCGA AAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACGCCGTAACACCGTTCTTAGGATGACTGGTTTTTTGGA AGTCTAA
Ty-0	ATGTCAGAAGAATTTCAAGAATCCGAGGTTATATTTTCCGATGAGTCTTTCACAAAGAACGATAACAAGATCAGTCATAACTACGAAAACTACGAAAGAAAGTCG ACGGAGAAGGAGAAAATTTCATCTCCGGTGAAAATTCCGTCAAGAACTACTTTCCGGTATACGGAAGAGGAGGGAGAAATGAAACCGCCACATGTCATAATCGA AAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACGCCGTAACACCGTTCTTAGGATGACTGGTTTTTTGGA AGTCTAA
Uk-1	ATGTCAGAAGAATTTCAAGAATCCGAGGTTATATTTTCCGATGAGTCTTTCACAAAGAACGATAACAAGATCAGTCATAACTACGAAAACTACGAAAGAAAGTCG ACGGAGAAGGAGAAAATTTCATCTCCGGTGAAAATTCCGTCAAGAACTACTTTCCGGTATACGGAAGAGGAGGGAGAAATGAAACCGCCACATGTCATAATCGA AAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACGCCGTAACACCGTTCTTAGGATGACTGGTTTTTTGGA AGTCTAA
Gel-1	ATGTCAGAAGAATTTCAAGAATCCGAGGTTATATTTTCCGATGAGTCTTTCACAAAGAACGATAACAAGATCAGTCATAACTACGAAAACTACGAAAGAAAGTCG ACGGAGAAGGAGAAAATTTCATCTCCGGTGAAAATTCCGTCAAGAACTACTTTCCGGTATACGGAAGAGGAGGGAGAAATGAAACCGCCACATGTCATAATCGA AAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACGCCGTAACACCGTTCTTAGGATGACTGGTTTTTTGGA AGTCTAA
Van-0	ATGTCAGAAGAATTTCAAGAATCCGAGGTTATATTTTCCGATGAGTCTTTCACAAAGAACGATAACAAGATCAGTCATAACTACGAAAACTACGAAAGAAAGTCG ACGGAGAAGGAGAAAATTTCATCTCCGGTGAAAATTCCGTCAAGAACTACTTTCCGGTATACGGAAGAGGAGGGAGAAATGAAACCGCCACATGTCATAATCGA AAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACGCCGTAACACCGTTCTTAGGATGACTGGTTTTTTGGA AGTCTAA
Ts-1	ATGTCAGAAGAATTTCAAGAATCCGAGGTTATATTTTCCGATGAGTCTTTCACAAAGAACGATAACAAGATCAGTCATAACTACGAAAACTACGAAAGAAAGTCG ACGGAGAAGGAGAAAATTTCATCTCCGGTGAAAATTCCGTCAAGAACTACTTTCCGGTATACGGAAGAGGAGGGAGAAATGAAACCGCCACATGTCATAATCGA AAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACGCCGTAACACCGTTCTTAGGATGACTGGTTTTTTGGA AGTCTAA
RRS-7	ATGTCAGAAGAATTTCAAGAATCCGAGGTTATATTTTCCGATGAGTCTTTCACAAAGAACGATAACAAGATCAGTCATAACTACGAAAACTACGAAAGAAAGTCG ACGGAGAAGGAGAAAATTTCATCTCCGGTGAAAATTCCGTCAAGAACTACTTTCCGGTATACGGAAGAGGAGGGAGAAATGAAACCGCCACATGTCATAATCGA AAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACGCCGTAACACCGTTCTTAGGATGACTGGTTTTTTGGA AGTCTAA
Ei-2	ATGTCAGAAGAATTTCAAGAATCCGAGGTTATATTTTCCGATGAGTCTTTCACAAAGAACGATAACAAGATCAGTCATAACTACGAAAACTACGAAAGAAAGTCG ACGGAGAAGGAGAAAATTTCATCTCCGGTGAAAATTCCGTCAAGAACTACTTTCCGGTATACGGAAGAGGAGGGAGAAATGAAACCGCCACATGTCATAATCGA AAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACGCCGTAACACCGTTCTTAGGATGACTGGTTTTTTGGA AGTCTAA
Xan-1	ATGTCAGAAGAATTTCAAGAATCCGAGGTTATATTTTCCGATGAGTCTTTCACAAAGAACGATAACAAGATCAGTCATAACTACGAAAACTACGAAAGAAAGTCG ACGGAGAAGGAGAAAATTTCATCTCCGGTGAAAATTCCGTCAAGAACTACTTTCCGGTATACGGAAGAGGAGGGAGAAATGAAACCGCCACATGTCATAATCGA AAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACGCCGTAACACCGTTCTTAGGATGACTGGTTTTTTGGA AGTCTAA

Figure S1: DNA sequences of *AtS40-3* among lowly susceptible accessions (Kondara, Ta-0, Ty-0, Uk-1, Gel-1, Van-0, Ts-1, RRS-7, Ei-2 and Xan-1) with shaded nucleotides indicating SNPs in coding and non-coding regions of AtS40-3 gene.

AtS40-3 DNA sequences among highly susceptible accessions

Zdr-1	ATGTCAGAAGAATTTCAAGAATCTGAGGTTATATTTTCCGATGAGTCTTTCACAAGGAAAGATAACAAGATCAGTCACAACAACGAAAACTACGAAAGAAA GTCGACGGAGAAGGATAAAATTTCATCTCCGGTGAGAATTCCGTCAAGAACTACTATCCGGTATACGGAAGAGGAGGGAGAAATGACACCGCCACATGTCA TAATCGAAAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACACCGTAACACCGTTCTTAGGATGACCG GCTTTTTGGAAGTTTAA
Kro-0	ATGTCAGAAGAATTTCAAGAATCTGAGGTTATATTTTCCGATGAGTCTTTCACAAGGAAAGATAACAAGATCAGTCACAACAACGAAAACTACGAAAGAAA GTCGACGGAGAAGGATAAAATTTCATCTCCGGTGAGAATTCCGTCAAGAACTACTATCCGGTATACGGAAGAGGAGGGAGAAATGACACCGCCACATGTCA TAATCGAAAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACACCGTAACACCGTTCTTAGGATGACCG GCTTTTTGGAAGTTTAA
Jm-0	ATGTCAGAAGAATTTCAAGAATCTGAGGTTATATTTTCCGATGAGTCTTTCACAAGGAAAGATAACAAGATCAGTCACAACAACGAAAACTACGAAAGAAA GTCGACGGAGAAGGATAAAATTTCATCTCCGGTGAGAATTCCGTCAAGAACTACTATCCGGTATACGGAAGAGGAGGGAGAAATGACACCGCCACATGTCA TAATCGAAAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACACCGTAACACCGTTCTTAGGATGACCG GCTTTTTGGAAGTTTAA
Mc-0	ATGTCAGAAGAATTTCAAGAATCTGAGGTTATATTTTCCGATGAGTCTTTCACAAGGAAAGATAACAAGATCAGTCACAACAACGAAAACTACGAAAGAAA GTCGACGGAGAAGGATAAAATTTCATCTCCGGTGAGAATTCCGTCAAGAACTACTATCCGGTATACGGAAGAGGAGGGAGAAATGACACCGCCACATGTCA TAATCGAAAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACACCGTAACACCGTTCTTAGGATGACCG GCTTTTTGGAAGTTTAA
Ha-0	ATGTCAGAAGAATTTCAAGAATCTGAGGTTATATTTTCCGATGAGTCTTTCACAAGGAAAGATAACAAGATCAGTCACAACAACGAAAACTACGAAAGAAA GTCGACGGAGAAGGATAAAATTTCATCTCCGGTGAGAATTCCGTCAAGAACTACTATCCGGTATACGGAAGAGGAGGGAGAAATGACACCGCCACATGTCA TAATCGAAAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACACCGTAACACCGTTCTTAGGATGACCG GCTTTTTGGAAGTTTAA
Gr-1	ATGTCAGAAGAATTTCAAGAATCTGAGGTTATATTTTCCGATGAGTCTTTCACAAGGAAAGATAACAAGATCAGTCACAACAACGAAAACTACGAAAGAAA GTCGACGGAGAAGGATAAAATTTCATCTCCGGTGAGAATTCCGTCAAGAACTACTATCCGGTATACGGAAGAGGAGGGAGAAATGACACCGCCACATGTCA TAATCGAAAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACACCGTAACACCGTTCTTAGGATGACCG GCTTTTTGGAAGTTTAA
Sap-0	ATGTCAGAAGAATTTCAAGAATCTGAGGTTATATTTTCCGATGAGTCTTTCACAAGGAAAGATAACAAGATCAGTCACAACAACGAAAACTACGAAAGAAA GTCGACGGAGAAGGATAAAATTTCATCTCCGGTGAGAATTCCGTCAAGAACTACTATCCGGTATACGGAAGAGGAGGGAGAAATGACACCGCCACATGTCA TAATCGAAAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACACCGTAACACCGTTCTTAGGATGACCG GCTTTTTGGAAGTTTAA
Bg-2	ATGTCAGAAGAATTTCAAGAATCTGAGGTTATATTTTCCGATGAGTCTTTCACAAGGAAAGATAACAAGATCAGTCACAACAACGAAAACTACGAAAGAAA GTCGACGGAGAAGGATAAAATTTCATCTCCGGTGAGAATTCCGTCAAGAACTACTATCCGGTATACGGAAGAGGAGGGAGAAATGACACCGCCACATGTCA TAATCGAAAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACACCGTAACACCGTTCTTAGGATGACCG GCTTTTTGGAAGTTTAA
Is-0	ATGTCAGAAGAATTTCAAGAATCTGAGGTTATATTTTCCGATGAGTCTTTCACAAGGAAAGATAACAAGATCAGTCACAACAACGAAAACTACGAAAGAAA GTCGACGGAGAAGGATAAAATTTCATCTCCGGTGAGAATTCCGTCAAGAACTACTATCCGGTATACGGAAGAGGAGGGAGAAATGACACCGCCACATGTCA TAATCGAAAAACGAAGAACGGAGGCGCAAATGGCGTTTTCTTTTTGTACCCTTAAAGGAAGAGACTTGAGTCGACACCGTAACACCGTTCTTAGGATGACCG GCTTTTTGGAAGTTTAA

Figure S2: DNA sequences of *AtS40-3* among highly susceptible accessions (Zdr-1, Kro-0, Jm-0, Mc-0, Ha-0, Gr-1, Sap-0, Bg-2 and Is-0).

AtS40-3 protein sequences among lowly susceptible accessions

Kondara	MSEEFQESEVIFSDESFTRKDNKISHNNENYERKSTEKEKISSPVRIPSRTTIRYTEEEGEMTPPHVIIEKRRTEAQMAFSFCTLKGRDLSRHRNTVLRMTGFLEVZ
Ta-0	MSEEFQESEVIFSDESFTKNDNKISHNYENYERKSTEKEKISSPVKIPSRTTFRYTEEEGEMKPPHVIIEKRRTEAQMAFSFCTLKGRDLSRRRNTVLRMTGFLEVX
Ty-0	MSEEFQESEVIFSDESFTKNDNKISHNYENYERKSTEKEKISSPVKIPSRTTFRYTEEEGEMKPPHVIIEKRRTEAQMAFSFCTLKGRDLSRRRNTVLRMTGFLEVX
Uk-1	MSEEFQESEVIFSDESFTKNDNKISHNYENYERKSTEKEKISSPVKIPSRTTFRYTEEEGEMKPPHVIIEKRRTEAQMAFSFCTLKGRDLSRRRNTVLRMTGFLEVX
Gel-1	MSEEFQESEVIFSDESFTKNDNKISHNYENYERKSTEKEKISSPVKIPSRTTFRYTEEEGEMKPPHVIIEKRRTEAQMAFSFCTLKGRDLSRRRNTVLRMTGFLEVX
Van-0	MSEEFQESEVIFSDESFTKNDNKISHNYENYERKSTEKEKISSPVKIPSRTTFRYTEEEGEMKPPHVIIEKRRTEAQMAFSFCTLKGRDLSRRRNTVLRMTGFLEVX
Ts-1	MSEEFQESEVIFSDESFTKNDNKISHNYENYERKSTEKEKISSPVKIPSRTTFRYTEEEGEMKPPHVIIEKRRTEAQMAFSFCTLKGRDLSRRRNTVLRMTGFLEVX
RRS-7	MSEEFQESEVIFSDESFTKNDNKISHNYENYERKSTEKEKISSPVKIPSRTTFRYTEEEGEMKPPHVIIEKRRTEAQMAFSFCTLKGRDLSRRRNTVLRMTGFLEVX
Ei-2	MSEEFQESEVIFSDESFTKNDNKISHNYENYERKSTEKEKISSPVKIPSRTTFRYTEEEGEMKPPHVIIEKRRTEAQMAFSFCTLKGRDLSRRRNTVLRMTGFLEVX
Xan-1	MSEEFQESEVIFSDESFTKNDNKISHNYENYERKSTEKEKISSPVKIPSRTTFRYTEEEGEMKPPHVIIEKRRTEAQMAFSFCTLKGRDLSRRRNTVLRMTGFLEVX

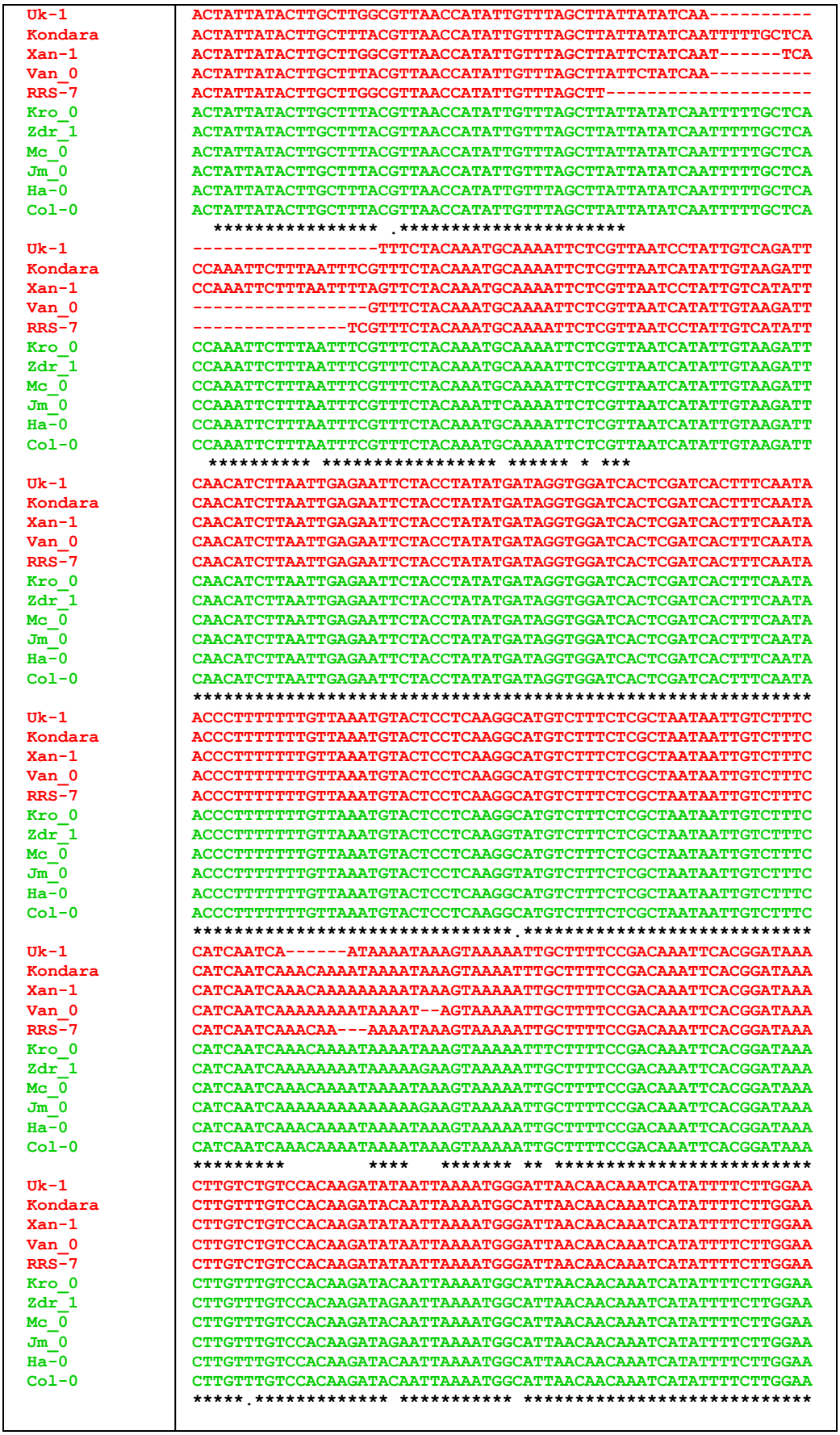
Figure S3: Protein sequences of *AtS40-3* among lowly susceptible accessions (Kondara, Ta-0, Ty-0, Uk-1, Gel-1, Van-0, Ts-1, RRS-7, Ei-2 and Xan-1) with shaded nucleotides indicating substituted amino acids with AtS40-3 gene.

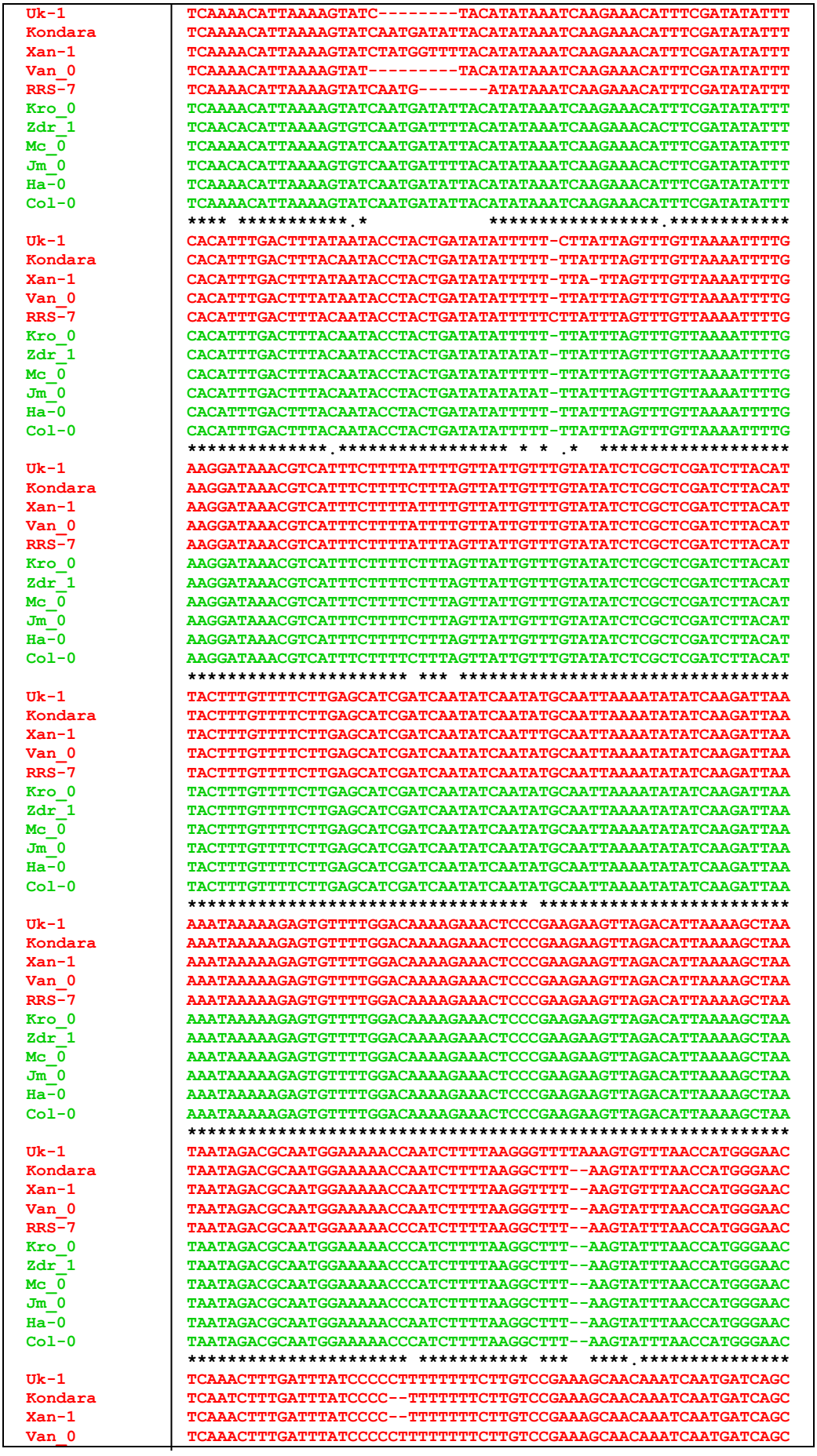
AtS40-3 protein sequences among highly susceptible accessions

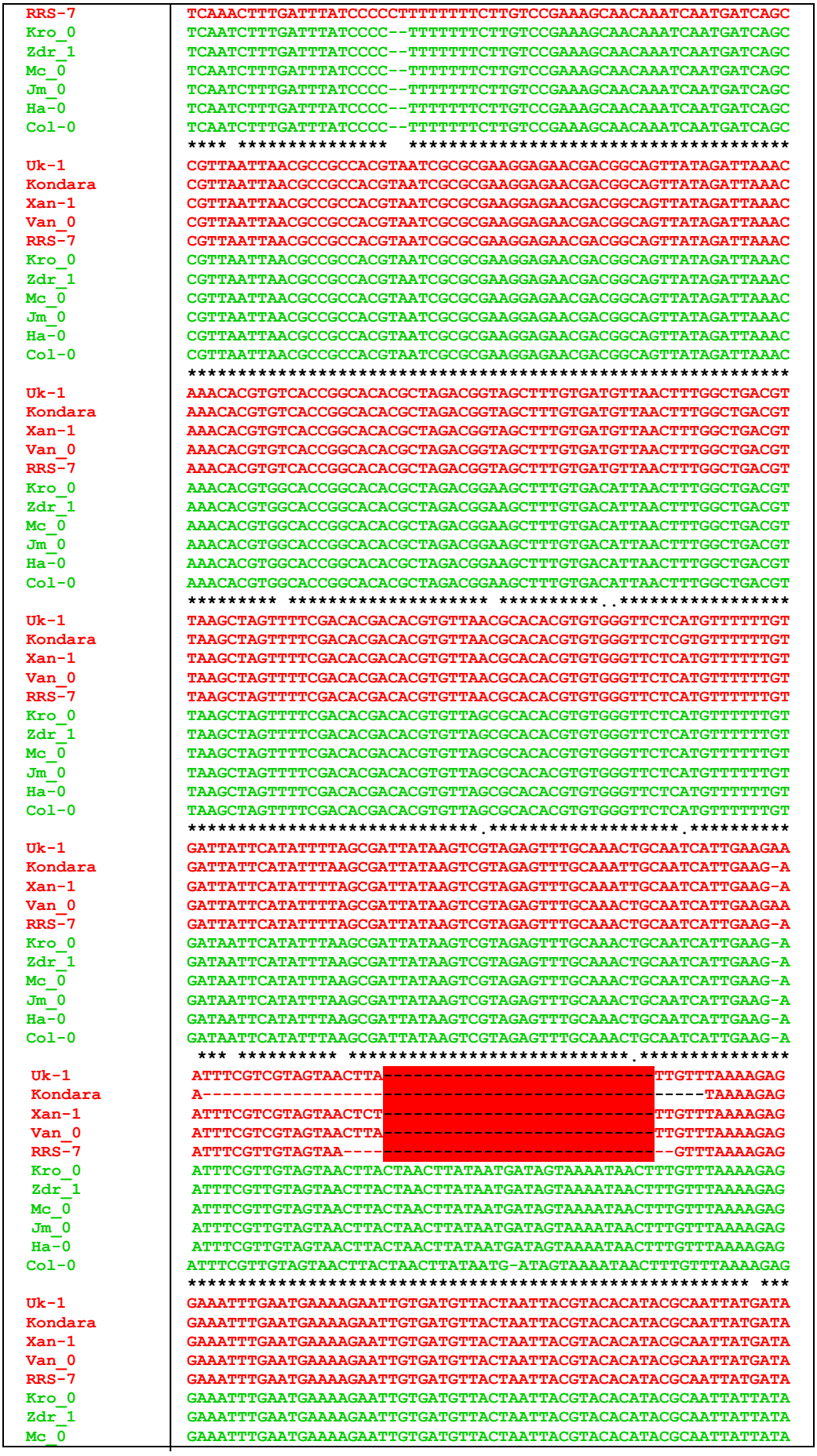
Zdr-1	MSEEFQESEVIFSDESFTRKDNKISHNNENYERKSTEKDKISSPVRIPSRTTIRYTEEEGEMTPPHVIIEKRRTEAQMAFSFCTLKGRDLSRHRNTVLRMTGFLEVX
Kro-0	MSEEFQESEVIFSDESFTRKDNKISHNNENYERKSTEKDKISSPVRIPSRTTIRYTEEEGEMTPPHVIIEKRRTEAQMAFSFCTLKGRDLSRHRNTVLRMTGFLEVX
Jm-0	MSEEFQESEVIFSDESFTRKDNKISHNNENYERKSTEKDKISSPVRIPSRTTIRYTEEEGEMTPPHVIIEKRRTEAQMAFSFCTLKGRDLSRHRNTVLRMTGFLEVX
Mc-0	MSEEFQESEVIFSDESFTRKDNKISHNNENYERKSTEKDKISSPVRIPSRTTIRYTEEEGEMTPPHVIIEKRRTEAQMAFSFCTLKGRDLSRHRNTVLRMTGFLEVX
Ha-0	MSEEFQESEVIFSDESFTRKDNKISHNNENYERKSTEKDKISSPVRIPSRTTIRYTEEEGEMTPPHVIIEKRRTEAQMAFSFCTLKGRDLSRHRNTVLRMTGFLEVX
Gr-1	MSEEFQESEVIFSDESFTRKDNKISHNNENYERKSTEKDKISSPVRIPSRTTIRYTEEEGEMTPPHVIIEKRRTEAQMAFSFCTLKGRDLSRHRNTVLRMTGFLEVX
Sap-0	MSEEFQESEVIFSDESFTRKDNKISHNNENYERKSTEKDKISSPVRIPSRTTIRYTEEEGEMTPPHVIIEKRRTEAQMAFSFCTLKGRDLSRHRNTVLRMTGFLEVX
Bg-2	MSEEFQESEVIFSDESFTRKDNKISHNNENYERKSTEKDKISSPVRIPSRTTIRYTEEEGEMTPPHVIIEKRRTEAQMAFSFCTLKGRDLSRHRNTVLRMTGFLEVX
Is-0	MSEEFQESEVIFSDESFTRKDNKISHNNENYERKSTEKDKISSPVRIPSRTTIRYTEEEGEMTPPHVIIEKRRTEAQMAFSFCTLKGRDLSRHRNTVLRMTGFLEVX

Figure S4: Protein sequences of *AtS40-3* among highly susceptible accessions (Kondara, Ta-0, Ty-0, Uk-1, Gel-1, Van-0, Ts-1, RRS-7, Ei-2 and Xan-1) indicating no polymorphisms.

AtS40-3 promoter sequence alignment among lowly and highly susceptible accessions




Uk-1 TCAAAACATTAAAAGTATC--------TACATATAAATCAAGAAACATTTCGATATATTT
Kondara TCAAAACATTAAAAGTATCAATGATATTACATATAAATCAAGAAACATTTCGATATATTT
Xan-1 TCAAAACATTAAAAGTATCTATGGTTTTACATATAAATCAAGAAACATTTCGATATATTT
Van_0 TCAAAACATTAAAAGTAT---------TACATATAAATCAAGAAACATTTCGATATATTT
RRS-7 TCAAAACATTAAAAGTATCAATG-------ATATAAATCAAGAAACATTTCGATATATTT
Kro_0 TCAAAACATTAAAAGTATCAATGATATTACATATAAATCAAGAAACATTTCGATATATTT
Zdr_1 TCAACACATTAAAAGTGTCAATGATTTTACATATAAATCAAGAAACACTTCGATATATTT
Mc_0 TCAAAACATTAAAAGTATCAATGATATTACATATAAATCAAGAAACATTTCGATATATTT
Jm_0 TCAACACATTAAAAGTGTCAATGATTTTACATATAAATCAAGAAACACTTCGATATATTT
Ha-0 TCAAAACATTAAAAGTATCAATGATATTACATATAAATCAAGAAACATTTCGATATATTT
Col-0 TCAAAACATTAAAAGTATCAATGATATTACATATAAATCAAGAAACATTTCGATATATTT
**** **********.* *****************.************
Uk-1 CACATTTGACTTTATAATACCTACTGATATATTTTT-CTTATTAGTTTGTTAAAATTTTG
Kondara CACATTTGACTTTACAATACCTACTGATATATTTTT-TTATTTAGTTTGTTAAAATTTTG
Xan-1 CACATTTGACTTTATAATACCTACTGATATATTTTT-TTA-TTAGTTTGTTAAAATTTTG
Van_0 CACATTTGACTTTATAATACCTACTGATATATTTTT-TTATTTAGTTTGTTAAAATTTTG
RRS-7 CACATTTGACTTTACAATACCTACTGATATATTTTTCTTATTTAGTTTGTTAAAATTTTG
Kro_0 CACATTTGACTTTACAATACCTACTGATATATTTTT-TTATTTAGTTTGTTAAAATTTTG
Zdr_1 CACATTTGACTTTACAATACCTACTGATATATATAT-TTATTTAGTTTGTTAAAATTTTG
Mc_0 CACATTTGACTTTACAATACCTACTGATATATTTTT-TTATTTAGTTTGTTAAAATTTTG
Jm_0 CACATTTGACTTTACAATACCTACTGATATATATAT-TTATTTAGTTTGTTAAAATTTTG
Ha-0 CACATTTGACTTTACAATACCTACTGATATATTTTT-TTATTTAGTTTGTTAAAATTTTG
Col-0 CACATTTGACTTTACAATACCTACTGATATATTTTT-TTATTTAGTTTGTTAAAATTTTG
**************.***************** * * .* ******************
Uk-1 AAGGATAAACGTCATTTCTTTTATTTTGTTATTGTTTGTATATCTCGCTCGATCTTACAT
Kondara AAGGATAAACGTCATTTCTTTTCTTTAGTTATTGTTTGTATATCTCGCTCGATCTTACAT
Xan-1 AAGGATAAACGTCATTTCTTTTATTTTGTTATTGTTTGTATATCTCGCTCGATCTTACAT
Van_0 AAGGATAAACGTCATTTCTTTTATTTTGTTATTGTTTGTATATCTCGCTCGATCTTACAT
RRS-7 AAGGATAAACGTCATTTCTTTTATTTAGTTATTGTTTGTATATCTCGCTCGATCTTACAT
Kro_0 AAGGATAAACGTCATTTCTTTTCTTTAGTTATTGTTTGTATATCTCGCTCGATCTTACAT
Zdr_1 AAGGATAAACGTCATTTCTTTTCTTTAGTTATTGTTTGTATATCTCGCTCGATCTTACAT
Mc_0 AAGGATAAACGTCATTTCTTTTCTTTAGTTATTGTTTGTATATCTCGCTCGATCTTACAT
Jm_0 AAGGATAAACGTCATTTCTTTTCTTTAGTTATTGTTTGTATATCTCGCTCGATCTTACAT
Ha-0 AAGGATAAACGTCATTTCTTTTCTTTAGTTATTGTTTGTATATCTCGCTCGATCTTACAT
Col-0 AAGGATAAACGTCATTTCTTTTCTTTAGTTATTGTTTGTATATCTCGCTCGATCTTACAT
********************** *** ********************************
Uk-1 TACTTTGTTTTCTTGAGCATCGATCAATATCAATATGCAATTAAAATATATCAAGATTAA
Kondara TACTTTGTTTTCTTGAGCATCGATCAATATCAATATGCAATTAAAATATATCAAGATTAA
Xan-1 TACTTTGTTTTCTTGAGCATCGATCAATATCAATTTGCAATTAAAATATATCAAGATTAA
Van_0 TACTTTGTTTTCTTGAGCATCGATCAATATCAATATGCAATTAAAATATATCAAGATTAA
RRS-7 TACTTTGTTTTCTTGAGCATCGATCAATATCAATATGCAATTAAAATATATCAAGATTAA
Kro_0 TACTTTGTTTTCTTGAGCATCGATCAATATCAATATGCAATTAAAATATATCAAGATTAA
Zdr_1 TACTTTGTTTTCTTGAGCATCGATCAATATCAATATGCAATTAAAATATATCAAGATTAA
Mc_0 TACTTTGTTTTCTTGAGCATCGATCAATATCAATATGCAATTAAAATATATCAAGATTAA
Jm_0 TACTTTGTTTTCTTGAGCATCGATCAATATCAATATGCAATTAAAATATATCAAGATTAA
Ha-0 TACTTTGTTTTCTTGAGCATCGATCAATATCAATATGCAATTAAAATATATCAAGATTAA
Col-0 TACTTTGTTTTCTTGAGCATCGATCAATATCAATATGCAATTAAAATATATCAAGATTAA
********************************** ************************
Uk-1 AAATAAAAAGAGTGTTTTGGACAAAAGAAACTCCCGAAGAAGTTAGACATTAAAAGCTAA
Kondara AAATAAAAAGAGTGTTTTGGACAAAAGAAACTCCCGAAGAAGTTAGACATTAAAAGCTAA
Xan-1 AAATAAAAAGAGTGTTTTGGACAAAAGAAACTCCCGAAGAAGTTAGACATTAAAAGCTAA
Van_0 AAATAAAAAGAGTGTTTTGGACAAAAGAAACTCCCGAAGAAGTTAGACATTAAAAGCTAA
RRS-7 AAATAAAAAGAGTGTTTTGGACAAAAGAAACTCCCGAAGAAGTTAGACATTAAAAGCTAA
Kro_0 AAATAAAAAGAGTGTTTTGGACAAAAGAAACTCCCGAAGAAGTTAGACATTAAAAGCTAA
Zdr_1 AAATAAAAAGAGTGTTTTGGACAAAAGAAACTCCCGAAGAAGTTAGACATTAAAAGCTAA
Mc_0 AAATAAAAAGAGTGTTTTGGACAAAAGAAACTCCCGAAGAAGTTAGACATTAAAAGCTAA
Jm_0 AAATAAAAAGAGTGTTTTGGACAAAAGAAACTCCCGAAGAAGTTAGACATTAAAAGCTAA
Ha-0 AAATAAAAAGAGTGTTTTGGACAAAAGAAACTCCCGAAGAAGTTAGACATTAAAAGCTAA
Col-0 AAATAAAAAGAGTGTTTTGGACAAAAGAAACTCCCGAAGAAGTTAGACATTAAAAGCTAA
****
Uk-1 TAATAGACGCAATGGAAAAACCAATCTTTTAAGGGTTTTAAAGTGTTTAACCATGGGAAC
Kondara TAATAGACGCAATGGAAAAACCAATCTTTTAAGGCTTT--AAGTATTTAACCATGGGAAC
Xan-1 TAATAGACGCAATGGAAAAACCAATCTTTTAAGGTTTT--AAGTGTTTAACCATGGGAAC
Van_0 TAATAGACGCAATGGAAAAACCAATCTTTTAAGGGTTT--AAGTATTTAACCATGGGAAC
RRS-7 TAATAGACGCAATGGAAAAACCCATCTTTTAAGGCTTT--AAGTATTTAACCATGGGAAC
Kro_0 TAATAGACGCAATGGAAAAACCCATCTTTTAAGGCTTT--AAGTATTTAACCATGGGAAC
Zdr_1 TAATAGACGCAATGGAAAAACCCATCTTTTAAGGCTTT--AAGTATTTAACCATGGGAAC
Mc_0 TAATAGACGCAATGGAAAAACCCATCTTTTAAGGCTTT--AAGTATTTAACCATGGGAAC
Jm_0 TAATAGACGCAATGGAAAAACCCATCTTTTAAGGCTTT--AAGTATTTAACCATGGGAAC
Ha-0 TAATAGACGCAATGGAAAAACCAATCTTTTAAGGCTTT--AAGTATTTAACCATGGGAAC
Col-0 TAATAGACGCAATGGAAAAACCCATCTTTTAAGGCTTT--AAGTATTTAACCATGGGAAC
********************** ********** *** ****.***************
Uk-1 TCAAACTTTGATTTATCCCCCTTTTTTTTCTTGTCCGAAAGCAACAAATCAATGATCAGC
Kondara TCAATCTTTGATTTATCCCC--TTTTTTTCTTGTCCGAAAGCAACAAATCAATGATCAGC
Xan-1 TCAAACTTTGATTTATCCCC--TTTTTTTCTTGTCCGAAAGCAACAAATCAATGATCAGC
Van_0 TCAAACTTTGATTTATCCCCCTTTTTTTTCTTGTCCGAAAGCAACAAATCAATGATCAGC


RRS-7 TCAAACTTTGATTTATCCCCCTTTTTTTTCTTGTCCGAAAGCAACAAATCAATGATCAGC
Kro_0 TCAATCTTTGATTTATCCCC--TTTTTTTCTTGTCCGAAAGCAACAAATCAATGATCAGC
Zdr_1 TCAATCTTTGATTTATCCCC--TTTTTTTCTTGTCCGAAAGCAACAAATCAATGATCAGC
Mc_0 TCAATCTTTGATTTATCCCC--TTTTTTTCTTGTCCGAAAGCAACAAATCAATGATCAGC
Jm_0 TCAATCTTTGATTTATCCCC--TTTTTTTCTTGTCCGAAAGCAACAAATCAATGATCAGC
Ha-0 TCAATCTTTGATTTATCCCC--TTTTTTTCTTGTCCGAAAGCAACAAATCAATGATCAGC
Col-0 TCAATCTTTGATTTATCCCC--TTTTTTTCTTGTCCGAAAGCAACAAATCAATGATCAGC
**** *************** ************************************
Uk-1 CGTTAATTAACGCCGCCACGTAATCGCGCGAAGGAGAACGACGGCAGTTATAGATTAAAC
Kondara CGTTAATTAACGCCGCCACGTAATCGCGCGAAGGAGAACGACGGCAGTTATAGATTAAAC
Xan-1 CGTTAATTAACGCCGCCACGTAATCGCGCGAAGGAGAACGACGGCAGTTATAGATTAAAC
Van_0 CGTTAATTAACGCCGCCACGTAATCGCGCGAAGGAGAACGACGGCAGTTATAGATTAAAC
RRS-7 CGTTAATTAACGCCGCCACGTAATCGCGCGAAGGAGAACGACGGCAGTTATAGATTAAAC
Kro_0 CGTTAATTAACGCCGCCACGTAATCGCGCGAAGGAGAACGACGGCAGTTATAGATTAAAC
Zdr_1 CGTTAATTAACGCCGCCACGTAATCGCGCGAAGGAGAACGACGGCAGTTATAGATTAAAC
Mc_0 CGTTAATTAACGCCGCCACGTAATCGCGCGAAGGAGAACGACGGCAGTTATAGATTAAAC
Jm_0 CGTTAATTAACGCCGCCACGTAATCGCGCGAAGGAGAACGACGGCAGTTATAGATTAAAC
Ha-0 CGTTAATTAACGCCGCCACGTAATCGCGCGAAGGAGAACGACGGCAGTTATAGATTAAAC
Col-0 CGTTAATTAACGCCGCCACGTAATCGCGCGAAGGAGAACGACGGCAGTTATAGATTAAAC
****
Uk-1 AAACACGTGTCACCGGCACACGCTAGACGGTAGCTTTGTGATGTTAACTTTGGCTGACGT
Kondara AAACACGTGTCACCGGCACACGCTAGACGGTAGCTTTGTGATGTTAACTTTGGCTGACGT
Xan-1 AAACACGTGTCACCGGCACACGCTAGACGGTAGCTTTGTGATGTTAACTTTGGCTGACGT
Van_0 AAACACGTGTCACCGGCACACGCTAGACGGTAGCTTTGTGATGTTAACTTTGGCTGACGT
RRS-7 AAACACGTGTCACCGGCACACGCTAGACGGTAGCTTTGTGATGTTAACTTTGGCTGACGT
Kro_0 AAACACGTGGCACCGGCACACGCTAGACGGAAGCTTTGTGACATTAACTTTGGCTGACGT
Zdr_1 AAACACGTGGCACCGGCACACGCTAGACGGAAGCTTTGTGACATTAACTTTGGCTGACGT
Mc_0 AAACACGTGGCACCGGCACACGCTAGACGGAAGCTTTGTGACATTAACTTTGGCTGACGT
Jm_0 AAACACGTGGCACCGGCACACGCTAGACGGAAGCTTTGTGACATTAACTTTGGCTGACGT
Ha-0 AAACACGTGGCACCGGCACACGCTAGACGGAAGCTTTGTGACATTAACTTTGGCTGACGT
Col-0 AAACACGTGGCACCGGCACACGCTAGACGGAAGCTTTGTGACATTAACTTTGGCTGACGT
********* ******************** **********..**********
Uk-1 TAAGCTAGTTTTCGACACGACACGTGTTAACGCACACGTGTGGGTTCTCATGTTTTTTGT
Kondara TAAGCTAGTTTTCGACACGACACGTGTTAACGCACACGTGTGGGTTCTCGTGTTTTTTGT
Xan-1 TAAGCTAGTTTTCGACACGACACGTGTTAACGCACACGTGTGGGTTCTCATGTTTTTTGT
Van_0 TAAGCTAGTTTTCGACACGACACGTGTTAACGCACACGTGTGGGTTCTCATGTTTTTTGT
RRS-7 TAAGCTAGTTTTCGACACGACACGTGTTAACGCACACGTGTGGGTTCTCATGTTTTTTGT
Kro_0 TAAGCTAGTTTTCGACACGACACGTGTTAGCGCACACGTGTGGGTTCTCATGTTTTTTGT
Zdr_1 TAAGCTAGTTTTCGACACGACACGTGTTAGCGCACACGTGTGGGTTCTCATGTTTTTTGT
Mc_0 TAAGCTAGTTTTCGACACGACACGTGTTAGCGCACACGTGTGGGTTCTCATGTTTTTTGT
Jm_0 TAAGCTAGTTTTCGACACGACACGTGTTAGCGCACACGTGTGGGTTCTCATGTTTTTTGT
Ha-0 TAAGCTAGTTTTCGACACGACACGTGTTAGCGCACACGTGTGGGTTCTCATGTTTTTTGT
Col-0 TAAGCTAGTTTTCGACACGACACGTGTTAGCGCACACGTGTGGGTTCTCATGTTTTTTGT
****************************.******************.**********
Uk-1 GATTATTCATATTTTAGCGATTATAAGTCGTAGAGTTTGCAAACTGCAATCATTGAAGAA
Kondara GATTATTCATATTTAAGCGATTATAAGTCGTAGAGTTTGCAAATTGCAATCATTGAAG-A
Xan-1 GATTATTCATATTTTAGCGATTATAAGTCGTAGAGTTTGCAAATTGCAATCATTGAAG-A
Van_0 GATTATTCATATTTTAGCGATTATAAGTCGTAGAGTTTGCAAACTGCAATCATTGAAGAA
RRS-7 GATTATTCATATTTTAGCGATTATAAGTCGTAGAGTTTGCAAACTGCAATCATTGAAG-A
Kro_0 GATAATTCATATTTAAGCGATTATAAGTCGTAGAGTTTGCAAACTGCAATCATTGAAG-A
Zdr_1 GATAATTCATATTTAAGCGATTATAAGTCGTAGAGTTTGCAAACTGCAATCATTGAAG-A
Mc_0 GATAATTCATATTTAAGCGATTATAAGTCGTAGAGTTTGCAAACTGCAATCATTGAAG-A
Jm_0 GATAATTCATATTTAAGCGATTATAAGTCGTAGAGTTTGCAAACTGCAATCATTGAAG-A
Ha-0 GATAATTCATATTTAAGCGATTATAAGTCGTAGAGTTTGCAAACTGCAATCATTGAAG-A
Col-0 GATAATTCATATTTAAGCGATTATAAGTCGTAGAGTTTGCAAACTGCAATCATTGAAG-A
*** ********** ****************************.***************
Uk-1 ATTTCGTCGTAGTAACTTA----------TTGTTTAAAAGAG
Kondara A----------TAAAAGAG
Xan-1 ATTTCGTCGTAGTAACTCT----------TTGTTTAAAAGAG
Van_0 ATTTCGTCGTAGTAACTTA----------TTGTTTAAAAGAG
RRS-7 ATTTCGTTGTAGTAA----------GTTTAAAAGAG
Kro_0 ATTTCGTTGTAGTAACTTACTAACTTATAATGATAGTAAAATAACTTTGTTTAAAAGAG
Zdr_1 ATTTCGTTGTAGTAACTTACTAACTTATAATGATAGTAAAATAACTTTGTTTAAAAGAG
Mc_0 ATTTCGTTGTAGTAACTTACTAACTTATAATGATAGTAAAATAACTTTGTTTAAAAGAG
Jm_0 ATTTCGTTGTAGTAACTTACTAACTTATAATGATAGTAAAATAACTTTGTTTAAAAGAG
Ha-0 ATTTCGTTGTAGTAACTTACTAACTTATAATGATAGTAAAATAACTTTGTTTAAAAGAG
Col-0 ATTTCGTTGTAGTAACTTACTAACTTATAATG-ATAGTAAAATAACTTTGTTTAAAAGAG
*** ***
Uk-1 GAAATTTGAATGAAAAGAATTGTGATGTTACTAATTACGTACACATACGCAATTATGATA
Kondara GAAATTTGAATGAAAAGAATTGTGATGTTACTAATTACGTACACATACGCAATTATGATA
Xan-1 GAAATTTGAATGAAAAGAATTGTGATGTTACTAATTACGTACACATACGCAATTATGATA
Van_0 GAAATTTGAATGAAAAGAATTGTGATGTTACTAATTACGTACACATACGCAATTATGATA
RRS-7 GAAATTTGAATGAAAAGAATTGTGATGTTACTAATTACGTACACATACGCAATTATGATA
Kro_0 GAAATTTGAATGAAAAGAATTGTGATGTTACTAATTACGTACACATACGCAATTATTATA
Zdr_1 GAAATTTGAATGAAAAGAATTGTGATGTTACTAATTACGTACACATACGCAATTATTATA
Mc_0 GAAATTTGAATGAAAAGAATTGTGATGTTACTAATTACGTACACATACGCAATTATTATA

```
Jm_0        GAAATTTGAATGAAAAGAATTGTGATGTTACTAATTACGTACACATACGCAATTATTATA
Ha-0        GAAATTTGAATGAAAAGAATTGTGATGTTACTAATTACGTACACATACGCAATTATTATA
Col-0       GAAATTTGAATGAAAAGAATTGTGATGTTACTAATTACGTACACATACGCAATTATTATA
            ********************************************************* ***

Uk-1        AGGCAAAATGATGTGAAAGTGTATATAAAGACTTGTGTAGTTCACGATAATCACATCCTT
Kondara     AGGCAAAATGATGTGAAAGTGTATATAAAGACTTGTGTAGTTCACGATAATCACATCCTT
Xan-1       AGGCAAAATGATGTGAAAGTGTATATAAAGACTTGTGTAGTTCACGATAATCACATCCTT
Van_0       AGGCAAAATGATGTGAAAGTGTATATAAAGACTTGTGTAGTTCACGATAATCACATCCTT
RRS-7       AGGCAAAATGATGTGAAAGTGTATATAAAGACTTGTGTAGTTCACGATAATCACATCCTT
Kro_0       AGGCAAAATGATGTGAAAGTGTAAATAAAGACTTGTGTAGTTCACGATAATCACATCCTT
Zdr_1       AGGCAAAATGATGTGAAAGTGTAAATAAAGACTTGTGTAGTTCACGATAATCACATCCTT
Mc_0        AGGCAAAATGATGTGAAAGTGTAAATAAAGACTTGTGTAGTTCACGATAATCACATCCTT
Jm_0        AGGCAAAATGATGTGAAAGTGTAAATAAAGACTTGTGTAGTTCACGATAATCACATCCTT
Ha-0        AGGCAAAATGATGTGAAAGTGTAAATAAAGACTTGTGTAGTTCACGATAATCACATCCTT
Col-0       AGGCAAAATGATGTGAAAGTGTAAATAAAGACTTGTGTAGTTCACGATAATCACATCCTT
            ********************** ************************************

Uk-1        AATGATTAATTAATAATGAGATCGAATCCCTTGACTTCACTCAAACACCTTCATGATGAT
Kondara     AATGATTAATTAATAATGAGATCGAATCCCTTGACTTCACTCAAACACCTTCACGATGAT
Xan-1       AACGATTAATTAATAATGAGATCGAATCCCTTGACTTCACTCAAACACCTTCACGATGAT
Van_0       AATGATTAATTAATAATGAGATCGAATCCCTTGACTTCACTCAAACACCTTCACGATGAT
RRS-7       AATGATTAATTAATAATGAGATCGAATCCCTTGACTTCACTCAAACACCTTCACGATGAT
Kro_0       AATGATTAATTAATAATGAGATCGAATCCCTTGACTTCACTCAAACACCTTCACGATGAT
Zdr_1       AATGATTAATTAATAATGAGATCGAATCCCTTGACTTCACTCAAACACCTTCACGATGAT
Mc_0        AATGATTAATTAATAATGAGATCGAATCCCTTGACTTCACTCAAACACCTTCACGATGAT
Jm_0        AATGATTAATTAATAATGAGATCGAATCCCTTGACTTCACTCAAACACCTTCACGATGAT
Ha-0        AATGATTAATTAATAATGAGATCGAATCCCTTGACTTCACTCAAACACCTTCACGATGAT
Col-0       AATGATTAATTAATAATGAGATCGAATCCCTTGACTTCACTCAAACACCTTCACGATGAT
            **.************************************************.******

Uk-1        GTCATCTGATGTAACAAAAGATATTAGTTAATACTTACACGTTTATATAC----------
Kondara     GTCATCTGATGTAACAAAAGATATTAGTTAATACTTACACGTTTCATACATATATG-AA
Xan-1       GTCATCTGATGTAACAAAAGATATTAGTTAATACTTACACGTTTACATACATATATGTAA
Van_0       GTCATCTGATGTAACAAAAGATATTAGTTAATACTTACACGTTTATATACATATATGTAA
RRS-7       GTCATCTGATGTAACAAAAGATATTAGTTAATACTTACACGTTTACATACATATATGTAA
Kro_0       GTCATCTGATGTAACAAAAGATATTAGTTAATACTTACACGTTTATATACATATATGTAA
Zdr_1       GTCATCTGATGTAACAAAAGATATTAGTTAATACTTACACGTTTATATACATATATGTAA
Mc_0        GTCATCTGATGTAACAAAAGATATTAGTTAATACTTACACGTTTATATACATATATGTAA
Jm_0        GTCATCTGATGTAACAAAAGATATTAGTTAATACTTACACGTTTATATACATATATGTAA
Ha-0        GTCATCTGATGTAACAAAAGATATTAGTTAATACTTACACGTTTATATACATATATGTAA
Col-0       GTCATCTGATGTAACAAAAGATATTAGTTAATACTTACACGTTTATATACATATATGTAA
            ******************************************* .****

Uk-1        ------------------------------------------------------------
Kondara     CTTATATGTGTATATATACGTGTAAAACATGAGATTCTCTTTTGCGATGTCTTCATCTAT
Xan-1       -TTATATGTGTATATATACGTGTAAAACATGAGATTCTCTTTTGCGATGTCTTCATCTAT
Van_0       CTTATATGTGTATATATACGTGTAAAACATGAGATTCTCTTTTGCGATGTCTTCATCTAT
RRS-7       CTTATATGTGTATATATACGTGTAAAACATGAGATTCTCTTTTGCGATGTCTTCATCTAT
Kro_0       CTTATATGTGTATATATACGTGTAAAACATGAGATTCTCTTTTGCGATGTCTTCATCTAT
Zdr_1       CTTATATGTGTATATATACGTGTAAAACATGAGATTCTCTTTTGCGATGTCTTCATCTAT
Mc_0        CTTATATGTGTATATATACGTGTAAAACATGAGATTCTCTTTTGCGATGTCTTCATCTAT
Jm_0        CTTATATGTGTATATATACGTGTAAAACATGAGATTCTCTTTTGCGATGTCTTCATCTAT
Ha-0        CTTATATGTGTATATATACGTGTAAAACATGAGATTCTCTTTTGCGATGTCTTCATCTAT
Col-0       CTTATATGTGTATATATACGTGTAAAACATGAGATTCTCTTTTGCGATGTCTTCATCTAT
            ************************.***.  **************.****.*********

Uk-1        ATAAATTCCACTTTTTTCCTTTAGCTGTGGATCTTCTTTTTTCTACCGTATTCTATATTT
Kondara     ATAAATTCCACTTTTTTCCTTTAGCTGTAGATCTTCTTTTTTCTACCGTACTCTATATTT
Xan-1       ATAAATTCCACTTTTTTCCTTTAGTTGTG--TCTTCTTTTTTCTATCGTATTCTATATTT
Van_0       ATAAATTCCACTTTTTTCCTTTAGCTGTGGATCTTCTTTTTTCTACCGTATTCTATATTT
RRS-7       ATAAATTCCACTTTTTTCCTTTAGCTGTGGATCTTCTTTTTTCTACCGTATTCTATATTT
Kro_0       ATAAATTCCACTTTTTTCCTTTAGCTGTGGATCTTCTTTTTTCTACCGTACTCTATATTT
Zdr_1       ATAAATTCCACTTTTTTCCTTTAGCTGTGGATCTTCTTTTTTCTACCGTACTCTATATTT
Mc_0        ATAAATTCCACTTTTTTCCTTTAGCTGTGGATCTTCTTTTTTCTACCGTACTCTATATTT
Jm_0        ATAAATTCCACTTTTTTCCTTTAGCTGTGGATCTTCTTTTTTCTACCGTACTCTATATTT
Ha-0        ATAAATTCCACTTTTTTCCTTTAGCTGTGGATCTTCTTTTTTCTACCGTACTCTATATTT
Col-0       ATAAATTCCACTTTTTTCCTTTAGCTGTGGATCTTCTTTTTTCTACCGTACTCTATATTT
            ************************.***.  **************.****.*********

Uk-1        TTCCTTACTTTCGTGGCAAAAACTTTACGATCTTTCAATG
Kondara     TTCCTTACTTTCGTGGTAAAAACTTTACGATCTTTCAATG
Xan-1       TTCCTTACTTTCGTGGCAAAAACTTTACGATCTTTCAATG
Van_0       TTCCTTACTTTCGTGGCAAAAACTTTACGATCTTTCAATG
RRS-7       TTCCTTACTTTCGTGGCAAAAACTTTACGATCTTTCAATG
Kro_0       TTCCTTACTTTCTAGGCAAAAACTTTACGATCTTTCAATG
Zdr_1       TTCCTTACTTTCTAGGCAAAAACTTTACGATCTTTCAATG
Mc_0        TTCCTTACTTTCTAGGCAAAAACTTTACGATCTTTCAATG
Jm_0        TTCCTTACTTTCTAGGCAAAAACTTTACGATCTTTCAATG
Ha-0        TTCCTTACTTTCTAGGCAAAAACTTTACGATCTTTCAATG
Col-0       TTCCTTACTTTCTAGGCAAAAACTTTACGATCTTTCAATG
```

Figure S5: Sequence alignment of shared promoter of *AtS40-3*and *PPR-proteins* in four lowly susceptible (Kondara, Uk-1, Xan-1 and Van-0) with nucleotide deletions positioned at 10398529-10398563 (red shaded) and four highly susceptible (Kro-0, Zdr-1, Jm-0 and Mc-0) accessions without nucleotide deletions.

Vectors used for gateway cloning

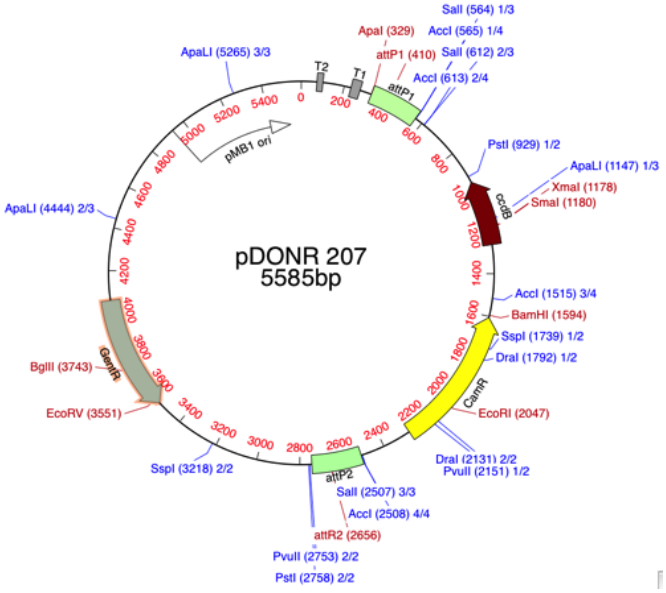


Figure S6: Vector map for the gateway donor vector pDONR 207 (*www.arabidopsis.org*).

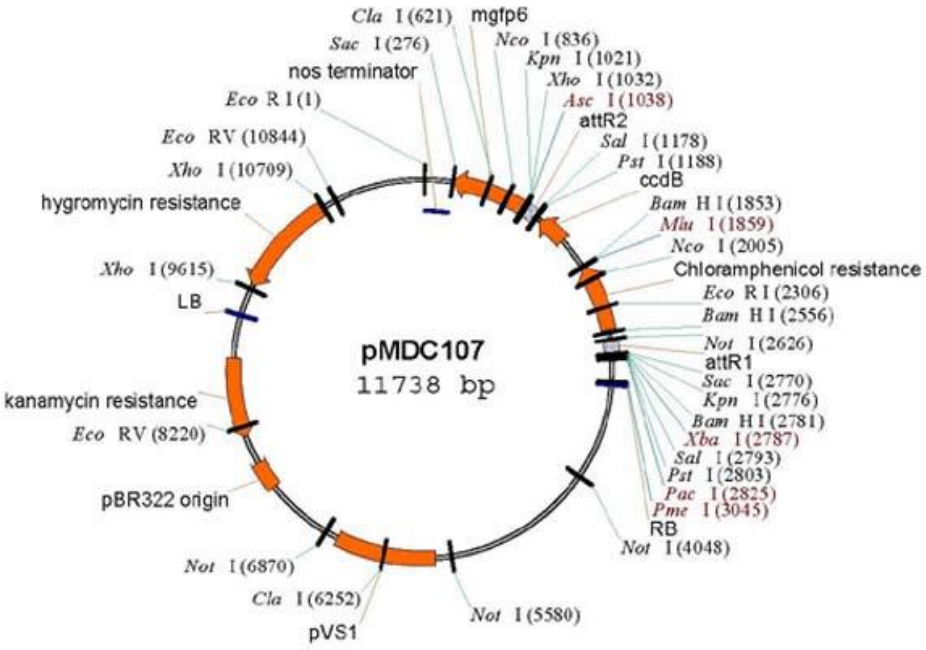


Figure S7: Vector map for the gateway destination vector pMDC 107 (*www.arabidopsis.org*).

8. Acknowledgement

Foremost, I would like to express my special gratitude and thanks to my supervisor Prof. Dr. Florian M.W. Grundler for his acceptance and providing an opportunity to work in his lab. I am grateful for his financial support during the entire period of my practical work. I am also thankful for his moral support in many personal issues and always playing a positive role to make me improved.

Besides my supervisor, I would like to thank to Dr. Shahid Siddique and Dr. Ali Naz for their valuable and scientific role in my research work. I am also grateful to my past and present lab mates for the stimulating discussions including Syed Jehangir Shah, Muhammad Shahzad Anjum, Shree Pariyar, Zoran Radakovic, Marion Hütton, Badou Mendy and other members for their practical and moral help during my entire research work. In addition, I would like to appreciate Dr. Adeel Mustafa for his help and our technical staff including Stefan Neumann and Gisela Sichtermann for maintenance of *Heterodera* and *Meloidogyne* cultur.

Last but not the least; I would like to express my gratitude towards my parents for the encouragement which helped me in completion of this work. Especial thank to my beloved and supportive wife, Ayesha Basharat, who has always been by my side when times I needed her most.