

Aus dem Institut für Pflanzenbau und Pflanzenzüchtung
der Christian-Albrechts-Universität zu Kiel

**Cloning of the beet cyst nematode resistance gene *Hs4* from the
wild beet *Patellifolia procumbens***

Dissertation

Zur Erlangung des Doktorgrades
der Agrar- und Ernährungswissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

M.Sc. Avneesh Kumar

aus Meerut, India

Kiel, 2020

Dekan: Prof. Karl H. Mühling

1. Berichterstatter: Prof. Dr. Christian Jung

2. Berichterstatter: Prof. Dr. Daguang Cai

Tag der mündlichen Prüfung: 17.02.2021

Table of Contents

1	General introduction.....	1
1.1	The sugar beet crop	1
1.2	Classification of the genus <i>Beta</i> and <i>Patellifolia</i>	1
1.3	<i>Beta</i> genetics and genomics.....	4
1.4	Plant parasitic nematodes	5
1.5	The beet cyst nematode	6
1.6	Plant-nematode interaction	7
1.7	Genetic resources for beet cyst nematode resistance and tolerance.....	12
1.8	Mapping and characterization of the translocation segment from <i>P. procumbens</i> to sugar beet	12
1.9	Resistance genes and candidate genes from the <i>P. procumbens</i> translocation...	13
1.10	Plant proteases and their role in plant defense mechanisms	15
1.11	Aims, expectations, and objectives.....	19
2	Materials and Methods	20
2.1	Plant materials and growth conditions	20
2.2	Nematode resistance tests	20
2.3	DNA isolation	21
2.4	RNA isolation	21
2.5	PCR and RT-qPCR	21
2.6	DNA sequencing	22
2.7	DNA sequence libraries.....	22
2.8	Bioinformatic analysis.....	23
2.9	CRISPR-Cas mediated gene knockout and overexpression studies	23
2.10	Agrobacterium mediated hairy root transformation	25
2.11	Microscopic studies	25
3	A complete sequence of the <i>P. procumbens</i> translocation in sugar beet	26
3.1	Genome assembly and translocation characterization	26
3.2	Identifying the translocation breakpoint at the end of super scaffold 2	31
4	<i>Hs4</i> candidate gene identification	34
4.1	Identification of the critical regions	34
4.2	Searching for ORFs within the critical regions	36
4.3	Candidate gene identification and characterization	36
4.4	<i>In-silico</i> analysis of ORF1	39
4.5	Transcriptional analysis of ORF1	41
5	Functional analysis of the ORF1.....	42

Table of Contents

5.1	<i>ORF1</i> knockout mutants produced by CRISPR-Cas9 mutagenesis	42
5.2	<i>ORF1</i> overexpression mutants.....	44
5.3	Infection tests with transgenic hairy roots.....	45
6	Closing Discussion.....	47
6.1	A complete sequence of <i>P. procumbens</i> translocation	47
6.2	Identification of the translocation breakpoint.....	48
6.3	Identification of candidate genes	50
6.4	Functional analysis of the candidate gene.....	52
6.5	Possible mechanism of rhomboid-like protease-mediated resistance.....	53
7	Summary	55
8	Zusammenfassung.....	56
9	Appendix	58
10	References	61
11	Supplementary data on CD/DVD	78
12	Curriculum Vitae and Publications.....	79
12.1	Curriculum vitae.....	79
13	Declaration of own contribution.....	80
14	Acknowledgments.....	81

List of Tables:

Table 1: The species of genus Beta and Patellifolia (adapted from (Kadereit et al.) 2006).	3
Table 2: Cloned nematode resistance genes.	11
Table 3: Wild beet chromosomes carrying genes for beet cyst nematode resistance or tolerance.....	13
Table 4: List of proposed candidate genes/ORF's from resistant sugar beet translocation line. RGA: Resistance Gene Analogue; NBS-LRR: Nucleotide Binding Site-Leucine Rich Repeats; MAP3K: Mitogen Activated Protein 3 kinase.....	15
Table 5: Plant proteases with known functions in plant defense response.....	18
Table 6: Seed code information for all the seed material used in this project.	20
Table 7: Whole genome sequencing/transcriptome sequencing data used in this study. PE: Paired-end, MP: Mate-pair, RNA-Seq: Transcriptome sequencing. The <i>P. procumbens</i> sequences were kindly provided by Dr. Chris Richards, (ARS USDA, USA).....	22
Table 8: Whole genome assembly statistics for TR520.	27
Table 9: Scaffolds from the translocation region (TR520), their size and the number of genes/ORFs present in each. The presence of a sequence was concluded based on the read mapping coverage.	31
Table 10: Scaffolds with high sequence homology between TR520 and TR363. These sequences were named 'critical regions', because they were not present in susceptible TR659.	34
Table 11: Genes present in the critical region. Gene ID is based on the most similar hit on the UniProt database. UniProt description is the description of Gene ID on the UniProt database. ORF number is given based on the position of ORF in the critical region. Previous ID, ID used in previous studies for the similar ORF.	36
Table 12: Likelihood of ORF1 sub-cellular localization as calculated by DeepLoc. DeepLoc predicts the subcellular localization of eukaryotic proteins.	40
Table 13: Summary statistics of hairy roots obtained from knock-out and overexpression experiments.	45

List of Figures:

Figure 1: Plant morphology of (A) sugar beet distant wild relative *Patellifolia procumbens* and (B) sugar beet (*Beta vulgaris* spp. *vulgaris*). 2

Figure 2: Life cycle of the beet cyst nematode (BCN) *Heterodera schachtii*. The life cycle of BCN can be divided into seven stages: egg stage, four larval stages, and two adult stages. The second stage juveniles (J2 larvae) are the infectious form. J2 larvae differentiate into J3 males or J3 females depending on the nutrition uptake. Males feed and leave the root in search of females, as J4 males. J3 females continue to feed on multinucleate cells, develop into J4 females, and get filled with eggs after copulation with males. A single cyst can have up to 200 eggs and can survive in the soil for up to 10 years. 7

Figure 3: Schematic representation of different stages of plant defense response along with some well-studied resistance genes. Recognition of effectors by RLK/RLP such as *Cf-2*, *Gpa2*, *Hero-4*. Signal integration via salicylic acid (SA), jasmonic acid (JA), ethylene (ET), abscisic acid (ABA), reactive oxygen species (ROS), reactive oxygen-nitrogen species (RONS) signaling. The last stage, defense response, includes programmed cell death (PCD), cell wall modification, RNAi, ROS, Phytoalexins. NBS: Nucleotide-binding site domain; TM: transmembrane domain; LRR: Leucine-rich repeat. 10

Figure 4: Sub-cloning of AtU6-26(P)-sgRNA cassette into p201G vector. The p201G plasmid carries Cas9 and eGFP encoding genes, each driven by a 35SCaMV promoter and an I-PpoI restriction site to clone the AtU6-26 (P)-sgRNA cassette. The cassette was amplified from the pChimera vector using rbd_gRNA_F and rbd_gRNA_R primers with I-PpoI restriction site overhangs and ligated into the p201G vector. 24

Figure 5: Cloning of *ORF1* cDNA into the overexpression vector plasmid pBin35SRed. The *ORF1* gene was cloned into pBin35SRed plasmid carrying the RFP encoding gene driven by the CsVMV (Cassava vein mosaic virus) promoter. *ORF1* cDNA was amplified using F_OEX and R_OEX primers with EcoR1 and Xba1 restriction sites overhangs. The PCR amplicon was ligated to the pBin35SRed vector behind the 35S promoter at the EcoR1/Xba1 restriction sites. 24

Figure 6: Localization of previously published *P. procumbens* specific repetitive elements on translocation specific scaffolds. pTS4.1 (Sau3A II) (red vertical bars), pTS3 (Sau3A III) (purple vertical bars), pAp4 (green vertical bars) and pAp22 (black vertical bars). Scale size is 50 kb. Each horizontal bar represents a scaffold..... 28

Figure 7: Comparison of physical map of the *P. procumbens* translocation based on (A) this study. and (B) previous work of Jäger (2012). (A) Physical map of the *P. procumbens* translocation from the resistant sugar beet line TR520 based on WGS data, *P. procumbens* specific repetitive sequences and YAC and BAC clones, in comparison to translocation lines TR363 (resistant), TR659 and TR320 (both susceptible). Green bar depicts sugar beet, red depicts the *P. procumbens* translocation. The dotted lines depict the sequences only present in resistant lines. Four sequences each for YAC and BAC clones are shown as black horizontal bars. Blue labels are published *P. procumbens* specific repeats. Individual scaffolds originating from WGS are shown as purple bars along with the name of each scaffold. The gaps between the red bars for TR320, TR659,

and TR363 represent absent translocation regions. The breakpoint of the translocation is marked with a black inverted triangle. Translocation line TR520 possesses a translocation segment of size 3.2 Mb composed of two super scaffolds bearing sizes 1.109 kb and 2.121 kb. **(B)** A sequence-based physical map of the *P. procumbens* translocation of line TR520. The map was aligned with molecular markers (CAU numbers) derived from the bacterial artificial chromosome (BAC), and yeast artificial chromosome (YAC) ends and scaffold sequences obtained by WGS sequencing (see Figure 1). A minimal tiling path of YACs and BACs and the scaffolds from WGS sequencing are integrated into this map. ‘Sections’ denote translocation-specific sequences present or absent on different translocation lines identified by marker analysis. The term ‘super scaffold’ (SS) is used for assembled contigs and scaffolds of different genome assemblies of *P. procumbens* and the translocation lines TR520 and TR363 of the WGS data. ‘Super contigs’ are the largest types of sequence assemblies on the physical map and incorporate BACs and super scaffolds of the WGS sequencing, whereas the contigs of the physical map presented by Capistrano (2009) are named ‘BAC contigs’ in the following (compare Figure 1). Scaffold mapping is shown in Figure 14. Shaded areas highlight the different sequence resources mapped to the translocation: YACs, BACs and WGS scaffolds. Regions in common between both resistant translocations and absent from the susceptible translocation are highlighted in yellow. 29

Figure 8: Gel electrophoresis with PCR products using a primer combination flanking translocation breakpoint. Genotyping the translocation breakpoint with two flanking markers (primer combination H208/H203). The primers were tested in four different genotypes, *P. procumbens*, translocation line TR520, translocation line TR363 and the sugar beet line 093161. A comparatively smaller band was observed in TR363 and no amplification was observed in *P. procumbens* and 093161. 1 % gel, 30 min., 90 V..... 33

Figure 9: Schematic representation of the translocation breakpoint in TR520. The dotted blue box depicts the position of the breakpoint. Black arrow indicated the position of the translocation at the end of chromosome 9 of translocation line TR520. Black arrow indicates the presence of translocation segment at the end of one of the homologous chromosome 9 of sugar beet. Orange arrows indicate the position of translocation breakpoint. Blue arrows indicate the positions of the primers used to amplify the region. 33

Figure 10: Identification of the critical regions. The ideogram shows sequence coverage of reads from TR363 (red) and TR659 (blue) mapped to the translocation region of TR520. The regions on scaffolds 11953, 33615, and 12163, where sequence reads from TR363 (red) are present while sequence reads from TR659 (blue) are absent, were considered to be the regions of interest. Critical regions are the regions within the black lines, where there is continuous presence of solid red vertical bars but no blue bars. Each horizontal bar represents one scaffold/contig of the translocation region. The green boxes within the bars represent the ORFs found in each scaffold. Scale bar = 20 kb. 35

Figure 11: The unrooted phylogenetic tree of ORF1 with top nine BLASTP hits from UniProt database. The two closest homologs (60% identity) to ORF1 are SOVF_040520, a rhomboid-domain encoding gene from *Spinacia oleracea* (A0A0K9RQE5), and BVRB_2g031850, which is a rhomboid-domain encoding gene from *Beta vulgaris subsp. vulgaris* (A0A0J8FQU9). The phylogenetic tree was constructed using Clustal Omega and edited in iTOL. Branch lengths are written above the branches. 39

Figure 12: *In-silico* prediction of the *ORF1* gene structure and protein domains. (A) Structural analysis of *ORF1*. *ORF1* contains 5-exons (grey boxes) and 5-introns (black line). The 5' and 3'-UTRs are depicted as red boxes. ATG: Translation start site, TGA: Translation stop site. (B) *ORF1* coding sequence. *ORF1* coding sequence is a linear combination of 5 exons covering 633 bp. (C) *ORF1* polypeptide sequence. The signal peptide is represented by the blue box. Rhomboid family domain is represented by an orange box. Five transmembrane domains are represented by black stripes, and two active sites of *ORF1* protein are shown by red bars (S106 and H181).40

Figure 13: Sequence alignment of *ORF1* protein with its closest homologue in the sugar beet genome (A0A0J8FQU9). *ORF1* is 210 aa long, while A0A0J8FQU9 is 307 aa long. Two active sites of *ORF1* protein are shown by red asterisks (S106 and H181). The signal peptide is marked by the green line. The rhomboid family domain is marked by the blue line. Brown line indicates the insertions41

Figure 14: Expression analysis of *ORF1* gene in infected and non-infected roots. Quantitative expression analysis of *ORF1* gene in response to nematode infection in roots from the resistant translocation line 3, 6, 9 and 12 dpi. Gene expression was quantified relative to *BvGAPDH*. Error bars are defined by the SEM of three biological replicates. No statistical significance was observed between the stages. Statistical significance was analyzed applying ANOVA.....41

Figure 15: Detection of GFP in hairy root clones under fluorescence microscope. (A) Bright field, dark field and merged images of three weeks old hairy roots from control hairy root clone, showing no green fluorescence (B) Bright field, dark field and merged images of three weeks old hairy roots from NEMATA expressing the GFP gene from the p201G vector. Scale bar = 1000 μ m.43

Figure 16: Results from the knockout experiments. (A) The sequences and positions of the two target sites in exon 1 of *ORF1* are highlighted in orange. (B) Four independent CRISPR-Cas induced mutant alleles. *hs4_1* is a 9 bp deletion. *hs4_2* and *hs4_3* are single base pair insertion. *hs4_4* is a 5 bp deletion. Red letter indicates insertions; Red hyphens indicate deletion.43

Figure 17: Detection of RFP in hairy root clones under the fluorescence microscope. (A) Bright field, dark field, and merged images of hairy roots from control hairy root clone, showing no red fluorescence (B) Bright field, dark field and merged images of 12 days old hairy roots from the susceptible line 093161 expressing the DsRed gene and carrying the *ORF1*-overexpression cassette from the pBin35SRed vector. Scale bar = 1000 μ m...44

Figure 18: Infection tests of CRISPR-Cas mutated hairy root clones. Box-plot is showing the number of cysts 4-weeks after infection in resistant (NEMATA), susceptible (093161), and 4 CRISPR-Cas mutated hairy root clones (*hs4_1*, *hs4_2*, *hs4_3*, *hs4_4*). Each box shows an interquartile range; horizontal line inside each box represents the median number of J4 females; vertical lines under the box show quartile 1; vertical lines above the box show quartile 3. Each dot represents an individual number of J4 females in the sample.....46

Figure 19: Number of cysts is negatively correlated with expression levels of *ORF1*. A bar-plot showing the expression levels of *ORF1* (blue bars, left y-axis) in ten different overexpression hairy root clones, and average number of cysts of five replicates 4-weeks

after infection (yellow bars, right y-axis). Clone OEX3 showed no expression of *ORF1* and showed the highest number of cysts among all RFP-positive hairy root clones. Hairy root clones OEX1, OEX6, OEX7, and OEX10 had no developing J4 females. Error bars for the relative expression represent SEM of three biological replicates, and error bars for infection tests represent SEM of cysts in five replicates. 46

List of abbreviations:

%	Percent
µl	Microliter
µm	Micrometer
ABA	Abscisic acid
<i>Avr</i>	<i>Avirulence</i>
AFLP	Amplified Fragment Length Polymorphism
<i>B. vulgaris</i>	<i>Beta vulgaris</i>
<i>B. nana</i>	<i>Beta nana</i>
<i>B. patula</i>	<i>Beta patula</i>
BAC	Bacterial Artificial Chromosome
BCN	Beet Cyst Nematode
BWA	Burrows-Wheeler Aligner
Cas9	CRISPR associated protein 9
CCN	Cereal Cyst Nematode
<i>CDC2a, CDKA;1</i>	<i>Cell cycle marker for division competence</i>
<i>CERK1</i>	<i>Chitin Elicitor Receptor Kinase</i>
CN	Cyst Nematode
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
<i>CYC1At; CYCB1;1</i>	<i>Cell cycle marker for G2-phase</i>
CYSTM	Cysteine-rich TM module stress tolerance
DAMP	Damage-associated molecular patterns
DNA	Deoxyribonucleic acid
EST	Expressed Sequence Tags
ET	Ethylene
ETI	Effector-triggered immunity
FISH	Fluorescence <i>in situ</i> hybridization
GFP	Green fluorescence protein
HR	Hypersensitive Response
IAEA	International Atomic Energy Agency
IGV	Integrative Genome Viewer
ISC	Initial Syncytial Cell
ISR	Induced systemic resistance
JA	Jasmonic acid
J1/2/3/4	Juvenile 1/2/3/4

List of abbreviations

KASP	Kompetitive allele specific PCR
Kbp	Kilo base pairs
MAL	Monosomic addition line
MAMP	Molecular-associated molecular patterns
Mbp	Mega base pairs
MTI	MAMP-triggered immunity
NBS-LRR	Nucleotide-Binding Leucine Rich Repeats
NGS	Next Generation Sequencing
ORF	Open Reading Frame
<i>P. procumbens</i>	<i>Patellifolia procumbens</i>
PCD	Programmed cell death
PCN	Potato Cyst Nematode
PCR	Polymerase Chain Reaction
<i>PIN</i>	<i>PIN-FORMED 1</i>
PPN	Plant-parasitic nematodes
rDNA	Ribosomal DNA
RGA	Resistance Gene Analogue
RLP	Receptor-Like Protein
RNA	Ribonucleic acid
RNAi	RNA interference
RONS	Reactive oxygen-nitrogen species
ROS	Reactive oxygen species
RAPD	Random Amplification of Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
RFP	Red Fluorescence Protein
RKN	Root Knot Nematode
RLK	Receptor-Like Kinase
SA	Salicylic acid
SAR	Systemic acquired resistance
SCN	Soybean Cyst Nematode
SNP	Single Nucleotide Polymorphism
SSR	Simple Sequence Repeat
SSU	Small Subunit
STAR	Spliced Transcripts Alignment to a Reference
TDF	Transcript-derived fragments

List of abbreviations

WGS	Whole Genome Sequencing
XTH	Xyloglucan endo-transglycosylase/hydrolase
YAC	Yeast Artificial Chromosome

1 General introduction

1.1 The sugar beet crop

Sugar beet (*Beta vulgaris* ssp. *vulgaris*) is a biennial dicotyledonous plant and a member of the family Chenopodiaceae (Enrico Biancardi, 2012; Enrico Biancardi, 2005). Sugar beet is the sixth most cultivated crop in the world in terms of tonnage and stands eighth among all the cultivated crops in terms of calories consumed (Ross-Ibarra et al., 2007). Sugar beets are primarily cultivated for sugar production. At small but progressive scale, sugar beet is also used in energy production (Maung and Gustafson, 2011; Rodríguez et al., 2010). It is the second most important sugar crop, after sugarcane, and supplies about 20% of the world's sugar (Finkenstadt, 2014).

Sugar in the form of sucrose is produced in leaves and translocated to other plant tissues (Ruan, 2014). In the first period of the sugar beet growing season, where the plants grow vegetatively, the sugar produced in the leaves is stored in the roots. However, in the second year of the growing season, as plants switch to the reproductive growth phase after vernalization, the sugar stored in the roots is translocated to the reproductive organs, causing a major reduction in sugar yields. Therefore, and to reduce losses in sugar yield, the roots of sugar beet plants are harvested at the end of the first year that is the vegetative phase of its cultivation period. Sugar beet roots constitute of ~75% water, ~25% soluble solids (~75% as sucrose) and ~5% insoluble solids (Hoffmann, 2010; Hoffmann and Märlander, 2005). The top five sugar beet-producing countries Russia, France, the United States, Germany, and Turkey, accounted for approximately 167 Million tons of sugar beet in the year 2019 (FAO 2019).

The cultivation of beets dates far back. Vegetative parts of beet plants, including roots and leaves were majorly consumed as vegetables by the Romans and the Greeks. Sugar beet is the most recent crop type among all the distinct cultivated lineages/crop types of *Beta vulgaris* spp. *vulgaris*. Sugar beet originated in the middle of the eighteenth century in German Silesia. All sugar beets present today originated from the “Weiße Schlesische Rübe”. It is assumed that nowadays sugar beet, is a product of mass selection over time from a cross between fodder beet and chard (Fischer, 1989).

1.2 Classification of the genus *Beta* and *Patellifolia*

Plant breeding is a deliberate manipulation of plant species to introduce superior genotypes with desirable traits such as higher yields, higher homogeneity, and resistance against biotic and abiotic stresses. Plant breeding has saved humans and animals from major food crises so far. However, these deliberate actions of intensive breeding of many years have unintentionally led to the narrowing of the genetic diversity in many crop species. Wild species/relatives of crops still possess a broad genetic base as compared to their respective cultivated species. Therefore, wild relatives of crops can serve as a potential source of broadening the genetic pool. Intensive breeding in sugar beet for the last 200 years has also led to a narrow genetic base in the cultivated lineages.

Species of genus *Beta* are distributed among four crop types: leaf beets (such as Swiss chard), garden beets (such as beetroot), fodder beets (including mangolds), and sugar beets. Evolutionary studies suggested that the *Beta* species had differentiated from its close relatives around six million years ago (Romeiras et al., 2016). The genus *Beta* comprises the sections: *Beta* and *Corollinae*. Section *Beta* is the most widely distributed because it encompasses all the cultivated forms, in addition to its close wild relatives. Most diverse naturally occurring species are centered in Europe and Western Asia,

however wild forms extend as far as China to the east and California, the USA to the west. Previously, the genus *Beta* included the section *Procumbentes*, which is now classified as a separate genus, *Patellifolia* (Kadereit et al., 2006). The new classification was based on the characterization of nuclear ribosomal DNA (rDNA) units, RAPD (Random Amplification of Polymorphic DNA), and RFLP (Restriction Fragment Length Polymorphism) markers (Mita et al., 1991; Santoni and Bervillé, 1992; Shen et al., 1998). Strong evidence from *in-situ* hybridization of three repetitive DNA families, specific to *Procumbentes* (Schmidt and Heslop-Harrison, 1996), suggested a distant relationship between *Procumbentes* and *Beta* species. The difference in the phenotype of *P. procumbens* and a commercial sugar beet hybrid can be seen in Figure 1. *P. procumbens* is a herbaceous, annual, or perennial with an erect/procumbent growth habit. The leaf blade is heart-shaped or hastate and is marked by a profusely fibrous root system. In contrast, a robust single elongated taproot system rich in sugar content is characteristic of sugar beet. The leaves are elongated, triangular in shape, and are arranged as spirals originating from the crown (Enrico Biancardi, 2005).



Figure 1: Plant morphology of (A) sugar beet distant wild relative *Patellifolia procumbens* and (B) sugar beet (*Beta vulgaris* spp. *vulgaris*).

Based on genetic relatedness, Harlan and de Wet (1971) categorized a genus into primary, secondary and tertiary gene pools. Primary gene pools comprise all the cultivated forms and closely related wild species. The species of primary gene pools are easy to cross and produce fertile progenies. Secondary gene pools encompass the species that can be crossed with primary gene pool members, but can lead to limited seed set, insufficient chromosome pairing and partially fertile hybrids. Whereas, in tertiary gene pools the seed set after crossing with the primary gene pool members is more or less possible, but seedlings of the hybrids show a range of physiological abnormalities such as lack of root formation.

The relatedness between the species of genus *Beta* and *Patellifolia* has been well studied. Due to the weak reproductive barriers among the species of section *Beta*, which includes *Beta vulgaris* ssp. *vulgaris*, ssp. *maritima*, ssp. *adanensis*, *B. patula*, and *B. macrocarpa*,

were grouped into the primary gene pool (Kadereit et al., 2006). Moreover, the species of the primary gene pool are genetically close to sugar beet, and thus homologous recombination is possible. This close genetic relatedness made it possible for breeders to successfully exploit the variation in the primary gene pool to improve sugar beet (Frese, 2010; Van Geyt et al., 1990). In fact, it is possible to find hybrids of cultivated beets and wild forms, such as *B. vulgaris* spp. *maritima*, in the geographical areas where they grow together (Arnaud et al., 2003; Bartsch et al., 1999).

Table 1: The species of genus *Beta* and *Patellifolia* (adapted from Kadereit et al. 2006).

Primary gene pool	Genus <i>Beta</i>
	I. Section <i>Beta</i>
	<i>B. vulgaris</i> L.
	spp. <i>vulgaris</i> <ul style="list-style-type: none"> • Cultivar group Leaf beet • Cultivar group Garden beet • Cultivar group Fodder beet • Cultivar group Sugar beet
	spp. <i>maritima</i> (L.) Arcang.
	spp. <i>adanensis</i> (Pamuk) Ford-Lloyd & J. T. Will.
	<i>B. macrocarpa</i> Guss.
	<i>B. patula</i> Ait.
	Secondary gene pool
	<i>Base species</i>
	<i>B. corolliflora</i> Zosimovich ex Buttler
	<i>B. macrorhiza</i> Steven
	<i>B. lomatogona</i> Fisch & Meyer
	<i>B. nana</i> Boiesier & Heldreich
	<i>Hybrid species</i>
	<i>B. trigyna</i> Waldst. & Kit.
Tertiary gene pool	Genus <i>Patellifolia</i>
	<i>Patellifolia procumbens</i> A. J. Scott et al.
	<i>P. patellaris</i> (C. Sm.) A. J. Scott et al
	<i>P. webbiana</i> (Moq) A. J. Scott et al

The species of section *Corollinae*, for instance, *B. corolliflora*, *B. macrorhiza*, and *B. lomatogona* are placed in the secondary gene pool. After experimental crosses of species of section *Corollinae* with sugar beet, limited seed set and few fertile hybrids with insufficient chromosome pairing were observed (Dalke, 1977). High homology between

the *B. corolliflora*-specific satellite sequences pBC1279 and pBC1944 with *B. nana*-specific repeat sequence pRN1 was found (Gao et al., 2000). *B. nana*, a species that also comes under section *Corollinae*, was, therefore, placed in the secondary gene pool (Kadereit et al., 2006).

Studies have shown that the distant wild relatives of sugar beet, *Patellifolia procumbens*, *P. webbiana*, and *P. patellaris*, come under the tertiary gene pool. RFLP markers (Mita et al., 1991), DNA fingerprinting (Jung et al., 1993), and SSR markers (Nachtigall et al., 2016) showed that *P. webbiana* composes a geographically isolated population of *P. procumbens* adapted to a specific habitat. Despite the development and fertility distortions that result from crossing sugar beet and *Patellifolia* species, the distantly related wild species *P. procumbens* contributed economically important traits to the sugar beet breeding pool, such as resistance against a range of biotic stresses, like beet cyst nematodes (BCN), *Polymyxa betae* which carries the beet necrotic yellow vein virus and partial resistance to *Cercospora* leaf spot (Enrico Biancardi, 2012). The species of genus *Beta* and *Patellifolia* are summarized in Table 1.

1.3 *Beta* genetics and genomics

The sugar beet genome has been well studied even long before the next generation sequencing (NGS) based genome assembly was introduced. For gene discovery and positional cloning, several genetic and physical maps based on DNA molecular markers have been constructed. For instance, Barzen et al. (1995) constructed an RFLP-based linkage map of the nine chromosomes of sugar beet. The map was constructed using a segregating population from a cross of two sugar beet plants. The plants were heterozygous for several agronomically important traits such as rhizomania resistance, monogermity, and hypocotyl color. Pillen et al. (1993) constructed a linkage map encompassing 177 segregating markers (2 morphological traits, 7 isozymes, 168 RFLP markers) on nine linkage groups. Using the RFLP-based linkage map Pillen et al. (1993) mapped the restorer gene *X* that was terminally located on linkage group III. The deviated markers from the expected segregation ratios suggested that the distorted segregation was due to linkage with four different lethal genes.

Hohmann et al. (2003) constructed a BAC library for sugar beet and a physical map of the chromosomal region harboring the bolting gene *B*. The BAC libraries were later used to develop molecular markers for the gene *B* (Gaafar et al., 2005). In the era of whole-genome sequencing and assemblies, physical maps have proven to be of great value, especially in the case of highly repetitive genomes such as that of sugar beet. The developed physical maps were employed to assign scaffolds/contigs to their respective chromosomes and to find the correct order of scaffolds in the genome assembly of sugar beet (Dohm et al., 2014b).

Considering their frequency, even distribution, and ease of analysis, SNP (Single nucleotide polymorphism) markers are highly favored for genotyping and genetic diversity analyses. Based on the previously developed RFLP-based linkage maps, candidate genes mapped earlier and ~10,000 unique sugar beet EST (Expressed Sequence Tag) sequences. Schneider et al. (2007) developed an SNP-based map derived from 16 *B. vulgaris* lines to identify polymorphisms from 315 EST- and 43 non-coding RFLP-derived loci. To further delineate the genetic structure of sugar beet, Stevanato et al. (2014) developed a panel of 192 SNPs that could effectively differentiate between different sugar beet genotypes. The SNPs were developed based on marker array technology, QuantStudio 12K Flex system. In a recent study, Galewski and McGrath (2020) used SNP data from WGS to assess the shared genetic variation, demographic

history, and crop type differentiation within the four cultivated lineages of *B. vulgaris ssp. vulgaris*. The SNP data was generated from 23 different beet cultivars and breeding lines with 25 individuals each. Galewski and McGrath (2020) showed that the studied crop types displayed extensive shared genetic variation, which suggested a shared demographic history within each crop type.

The sugar beet genome consists of 7 metacentric and 2 sub-centric mitotic chromosomes with proper chromosome nomenclature (Schondelmaier and Jung, 1997). The average physical chromosome length is 2.5 μm . The estimated genome size of sugar beet is 758 Mb (Arumuganathan and Earle, 1991) with a repeat content of 63% (Flavell et al., 1974). Repetitive sequences have been localized along sugar beet chromosomes using different techniques, such as fluorescence in situ hybridization (FISH) (Dechyeva and Schmidt, 2006; Paesold et al., 2012), Southern blotting (Schmidt and Metzloff, 1991), and sequencing-based assay of PCR products (Dechyeva and Schmidt, 2006; Paesold et al., 2012; Schmidt and Metzloff, 1991; Staginnus et al., 2001). Telomeric sequences of *A. thaliana* (TTTAGGG)_n repeats were cloned by Richards and Ausubel (1988) and were later localized at the ends of chromosome arms of sugar beet (Schmidt et al., 1998).

The first assembled sugar beet genome of the doubled haploid line KWS2320, comprised 567 Mb, of which 85% could be assigned to chromosomes. The genome contained 27,421 protein-coding genes (Dohm et al., 2014b). Recently, another sugar beet genome assembly of the inbred line EL10 has been published (Funk et al., 2018; McGrath et al., 2020). The EL10.1 genome assembly is 540 Mb in size with 24,255 gene models. EL10.1 genome assembly is reported to be more contiguous compared to the published sugar beet reference genome, however, it is smaller and contains comparatively a smaller number of gene models predicted.

Recently, the genomes of the two important wild beet species from section Beta, *B. patula* and *B. vulgaris ssp. maritima*, have also been sequenced (Rodríguez del Río et al., 2019). The genomes of the two species were compared with that of their cultivated relative sugar beet, and large syntenic regions between the three genomes were identified. The genome of *P. procumbens* is not assembled yet, but genetic mapping of some repetitive sequences has been performed. Sau3A satellites I and II have been isolated and physical mapping of the DNA families has been carried out using FISH (Schmidt and Heslop-Harrison, 1996).

1.4 Plant parasitic nematodes

Plant-parasitic nematodes belong to the phylum Nematoda. The evolutionary history of the members of phylum Nematoda goes back for an estimated one billion years, which explains their enormous diversity (Wang et al., 1999). Through small subunit sequence phylogeny analysis, it has been shown that nematodes evolved several times to parasitize animals and plants (Blaxter et al., 1998). Nematodes are soft-bodied multicellular organisms. Similar to insects, nematodes need to molt between several juvenile stages. Generally, nematodes undergo four molting phases in their life cycle. Nematodes have evolved to live in almost every conceivable place on earth (Yeates, 2010) and feed on bacteria, fungi, other nematodes, animals, and plants (Blanc et al., 2006; Decraemer and Hunt, 2006; Okada et al., 2005; Yeates et al., 1993). At least 15% of the described species can manage to feed on plants. Today, plant-parasitic nematodes are recognized as one of the major agricultural pathogens causing global losses of around USD 157 billion annually (Abad et al., 2008). Plant-parasitic nematodes can be classified into different categories depending on their feeding patterns. They can feed through lesions, lance, needles, sting, and stunt (Kumar et al., 2020). Root-knot nematode (RKN) and cyst

nematodes are the two most economically damaging nematodes. The most notable cyst nematodes are the beet cyst nematode (BCN) (*Heterodera schachtii*), soybean cyst nematode (*Heterodera glycines*), potato cyst nematodes (*Globodera pallida* and *G. rostochiensis*), and cereal cyst nematodes (*Heterodera avenae* and *H. filipjevi*) (Jones et al., 2013).

1.5 The beet cyst nematode

The beet cyst nematode (*Heterodera schachtii* Schmidt), is a main pest of sugar beet (*Beta vulgaris* L.). It was first discovered in 1859 and was named by Schmidt (1871). It has a broad host range, including more than 80% of species from different plant families, such as *Amaranthaceae* and *Brassicaceae*. The BCN is present in all the sugar beet growing areas in the world, therefore, it has high agronomical relevance (McCarter, 2009). Stunted growth and wilting of aerial parts of the sugar beet plants are the characteristic symptoms that can be identified as big circular patches in an aerial view of the field.

H. schachtii completes its life cycle in seven stages: an egg stage, four larval stages (J1, J2, J3, and J4), and two adult stages (Figure 2). Until the J2 stage, there is no differentiation between male and female nematodes. First stage juveniles (J1) develop inside the eggshell, while J2 larvae hatch and actively search for the root tissue to feed (Raski, 1950). Juvenile cyst nematodes enter the elongation zones of host plant roots after emerging from soil-borne cysts (Sijmons et al., 1991). Larvae use thrusts of their stylet and collection of cell wall degrading enzymes to hydrolyze the cellulose in root tissues (Vanholme et al., 2009). Mechanisms and signals attracting J2 larvae to the host roots are not fully resolved yet. Once in the roots, they migrate towards the vascular cylinder to initiate the formation of a multinucleated cell, called initial syncytial cell (ISC) from a procambial or pericycle cell (Grundler et al., 1997). Nematodes complete their life cycle feeding at the same place (Hussey, 1989). J3 stage is where males and females are distinguished and, thereafter, follow different life cycles. Sex determination is mainly controlled by nutrient availability (Böckenhoff and Grundler, 2009). J3 males have longer ingestion periods and consume less food when compared to J3 females. J3 males leave the feeding site to mate females. Therefore, they do not cause severe damages to the root tissue (Wyss et al., 1992). After completing their life cycle in the root, the females attach to the roots and form cysts with several hundred eggs. Depending on environmental factors, nematodes spend 30 days in the roots to complete their life cycle. Hatching can be delayed until adequate environmental cues are present. This dormancy phase can last up to ten years.

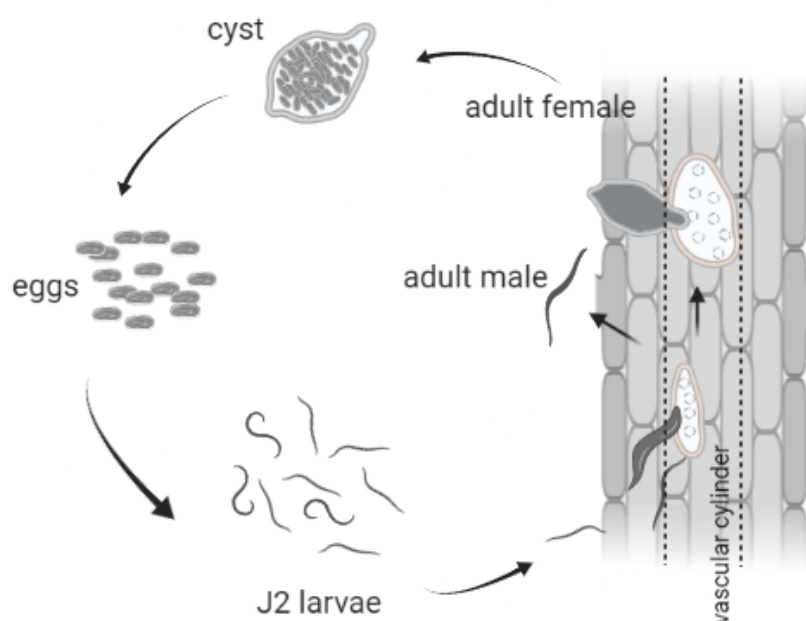


Figure 2: Life cycle of the beet cyst nematode (BCN) *Heterodera schachtii*. The life cycle of BCN can be divided into seven stages: egg stage, four larval stages, and two adult stages. The second stage juveniles (J2 larvae) are the infectious form. J2 larvae differentiate into J3 males or J3 females depending on the nutrition uptake. Males feed and leave the root in search of females, as J4 males. J3 females continue to feed on multinucleate cells, develop into J4 females, and get filled with eggs after copulation with males. A single cyst can have up to 200 eggs and can survive in the soil for up to 10 years.

1.6 Plant-nematode interaction

Plant-parasitic nematodes manipulate the molecular machinery of the plant cell to induce a feeding site inside the host plant (McK Bird, 1996). One of the very early hypotheses was that auxin, a phytohormone that tightly regulates plant development and growth, could be involved during nematode infection (Goverse et al., 2000). Goverse et al. (2000) provided a line of evidence for the implication of auxin in nematode feeding site formation in *Arabidopsis*. The study showed that auxin is locally accumulated in the developing syncytia. In addition, auxin-insensitive tomato and *Arabidopsis pin1* mutants showed significantly reduced nematode infection.

The formation of a multinucleate feeding cell requires the dissolution of cell walls and fusion of protoplasts of neighboring cells (Golinowski et al., 1996; Sobczak et al., 1997). Feeding sites are the sole nutrient source for nematodes; hence they need a continuous flow of nutrients from neighboring cells (Siddique and Grundler, 2015). This flow of nutrients into the feeding cell increases the pressure inside the feeding cell. To protect the feeding cell from bursting, nematodes induce cell wall thickening of the feeding cell, likely by inducing xyloglucan endo-transglycosylase/hydrolase (XTH) genes callose or lignin deposition (Grunder et al., 1998). Niebel et al. (1996) showed that cell cycle marker for division competence *CDKA1* (*CYCLIN-DEPENDENT KINASE A-1*) and cell cycle marker for the G2-phase *CYCB11* (*CYCLIN-B1-1*) were transcriptionally activated in the very early phase of ISC development (De Almeida Engler et al., 2004).

Transcriptome analysis of syncytia induced by BCN in *Arabidopsis* roots showed that the syncytia are transcriptionally different from the roots and other organs of the *Arabidopsis* plant (Szakasits et al., 2009). Studies have shown that polar auxin transporter genes *PIN-FORMED 1* (*PIN1*) were found to be expressed during syncytium initiation and

expansion (Grunewald et al., 2009). PIN proteins are asymmetrically localized plasma-membrane proteins. The movement of these proteins between the plasma membrane and the endosomal parts of the plant cell allows dynamic changes in their localization and redirects auxin to other parts of the cell (Geldner et al., 2001).

Using electron microscopy, comparative analysis of root histology, anatomy, and ultrastructure of nematode-induced syncytia in roots of susceptible and resistant sugar beet was performed. The analysis revealed pronounced structural and ultrastructural differences of syncytia formed in roots of resistant plants that were smaller and less hypertrophied than syncytia in susceptible plants. In addition, vacuoles in resistant plants were larger but less in number, and only rough endoplasmic reticulum was prevalent in syncytia as compared to susceptible plants (Holtmann et al., 2000). In another study by Sobczak et al. (1997), the structural changes induced during the development of males of the plant-parasitic nematode *H. schachtii* in *Arabidopsis thaliana* roots were analyzed at anatomical and ultrastructural levels. The study showed that syncytia of males were induced in the root pericycle. Proliferation and expansion of cambium and peridermal tissues were triggered by expanding the syncytium in a manner similar to secondary growth, which resulted in the formation of additional xylem and phloem elements in the roots of the *Arabidopsis* plants. Moreover, the syncytia associated with males was less hypertrophied and composed of more cells than syncytia associated with females (Sobczak et al., 1997).

Several transcriptome studies have been carried out to analyze the BCN-host interaction. The transcriptome studies were performed for both the host plant and the nematode. A transcriptome study was recently carried out for the resistance sugar beet cultivar 'Nemakill' to reveal the molecular insights of the compatible and incompatible interactions between sugar beet and *H. schachtii* (Ghaemi et al., 2020). Ghaemi et al. (2020) showed that most defense-related genes were induced at 4 days after infection while suppressed 10 days after infection in the susceptible cultivar. Whereas the genes involved in plant defense response were observed at both time points in the resistant cultivar Nemakill. In addition, an exogenous application of Methyl Jasmonate on susceptible plants revealed a significant reduction in plant susceptibility. The genes involved in the phenylpropanoid pathway and genes encoding CYSTM domain-containing proteins, F-box proteins, chitinase, galactono-1,4-lactone dehydrogenase, and CASP-like protein were identified in the resistant Nemakill cultivar (Ghaemi et al., 2020). In another study, cDNA-AFLP technique was used to isolate genes that were up-regulated upon infection (Samuelian et al., 2004). In this study out of 8000 transcript-derived fragments (TDF), one TDF of interest was further examined. Upon overexpression of this TDF in sugar beet hairy roots, 12 out of 15 clones showed significant reduction in number of developing females.

In order to unravel the set of secretory proteins (effectors) secreted by the nematodes for the induction and maintenance of their syncytial feeding sites in plant roots, a transcriptome and parasitome analysis of the BCN *H. schachtii* was performed (Elashry et al., 2020). A whole animal pre-infective J2-stage transcriptome together with pre- and post-infective J2 gland cell transcriptome was sequenced using NGS. Out of the 200 putative effectors identified, expressions of six putative effectors were quantified using qPCR. Functional analysis using RNAi of three putative effectors indicated that the level of nematodes pathogenicity and/or the average female size was reduced in all tested genes, indicating the role of these genes in the cyst nematode parasitism.

R genes have been shown to be involved in plant defense mechanisms against cyst nematodes (Table 2). In potato, the resistance gene, *Gpa2*, encodes a leucine-zipper-Nucleotide Binding Site-Leucine Rich Repeats (NBS-LRR) that confers resistance against the potato cyst nematode *Globodera pallida* (Van Der Vossen et al., 2000). Another potato *R* gene, *Gro1-4*, encodes a Toll-interleukin 1 receptor-NBS-LRR protein that confers resistance to the endoparasitic nematode *G. rostochiensis* pathotype Ro1 (Paal et al., 2004). *Gro1-4* is expressed in most tissues of the potato plant, including non-infected roots (Paal et al., 2004). In tomato (*Solanum pimpinellifolium*), the *Hero* gene is an NBS-LRR gene shown to confer resistance to all pathotypes of *G. rostochiensis* and *G. pallida* (Ernst et al., 2002). The extracellular immune receptor protein Cf-2 of the red currant tomato (*S. pimpinellifolium*) has been shown to confer dual resistance against the fungus *Cladosporium fulvum* and the root parasitic nematode *G. rostochiensis* (Lozano-Torres et al., 2012). Cf-2 mediates resistance via sensing perturbations caused by binding of the venom allergen-like effector protein Gr-VAP1 of *G. rostochiensis* to the active site of a papain-like cysteine protease Rcr3. Nematode infection in tomato plants harboring Cf-2 and Rcr3 induces defense related programmed cell death in the infected plant cells (Lozano-Torres et al., 2012). Another NBS-LRR class of *R* genes, *Mi* genes, are among the most important resistance genes. The gene *Mi-1.2* was cloned from the *Mi* locus of tomato that confers resistance to the most damaging species of root-knot nematodes (*Meloidogyne* spp) (Milligan et al., 1998b). In addition to root-knot nematodes, the *Mi-1.2* gene was shown to confer resistance to two major sucking type pests of tomato, whitefly, and aphids, making it an essential source in integrated pest management programs (Nombela et al., 2003). A homolog of *Mi-1.2*, *Mi-9*, was cloned from *S. arcanum* and was shown to confer a heat-stable resistance to root-knot nematodes (Jablonska et al., 2007). Another *Mi* homolog, *CaMi*, was isolated from a root-knot nematode-resistant pepper line (*Capsicum annuum* L.) (Chen et al., 2007). *CaMi* was highly expressed in the roots, leaves, and flowers of the resistant plants. RKN, susceptible tomato plants, expressing *CaMi* gene, induced a hypersensitive response and necrotic cells around the nematode instead of galls or root knots as in the wild type (Chen et al., 2007).

Apart from *R* genes, *RESISTANCE TO HETERODERA GLYCINES 4* (*Rhg4*) gene, which encodes a serine hydroxymethyl transferase enzyme, was shown to confer resistance against the soybean cyst nematode *H. glycines* (Liu et al., 2012). *Rhg4* was cloned by map-based cloning from the soybean cyst nematode resistant soybean cultivar cv. Forrest. In plants carrying the *Rhg4* gene, the roots are penetrated by the nematodes but the feeding site degenerates causing nematodes to die before reaching the adult stage. Interaction between three prominent resistance proteins, effectors, and their role in plant-defense pathway is illustrated in Figure 3.

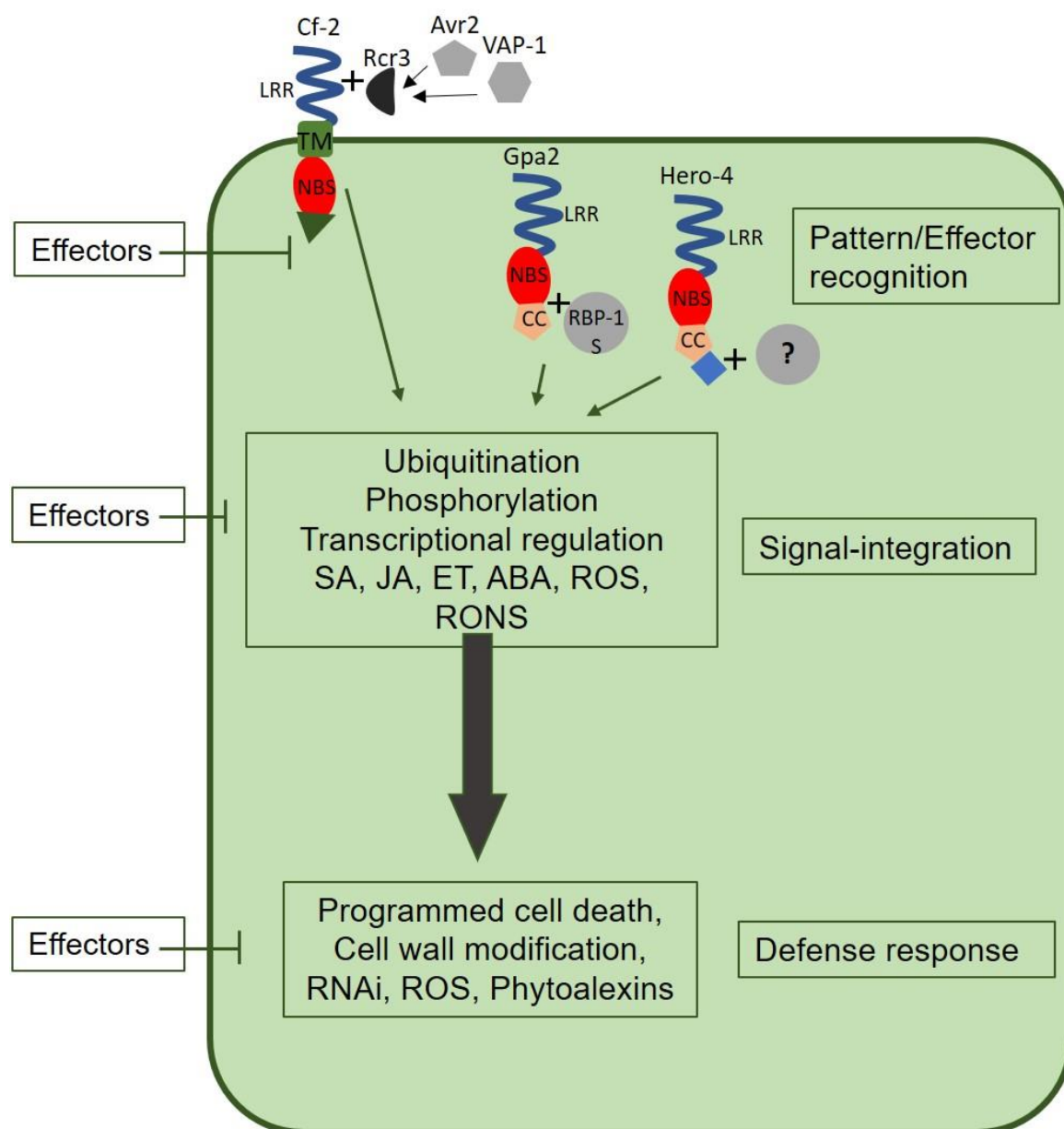


Figure 3: Schematic representation of different stages of plant defense response along with some well-studied resistance genes. Recognition of effectors by RLK/RLP such as *Cf-2*, *Gpa2*, *Hero-4*. Signal integration via salicylic acid (SA), jasmonic acid (JA), ethylene (ET), abscisic acid (ABA), reactive oxygen species (ROS), reactive oxygen-nitrogen species (RONS) signaling. The last stage, defense response, includes programmed cell death (PCD), cell wall modification, RNAi, ROS, Phytoalexins. NBS: Nucleotide-binding site domain; TM: transmembrane domain; LRR: Leucine-rich repeat.

Table 2: Cloned nematode resistance genes.

Gene	Host species	Nematode species	Functional domain	Reference
<i>Hs1^{pro-1}</i>	<i>Patellifolia procumbens</i>	<i>Heterodera schachtii</i>	N-terminal extra-cellular LRR, TM	(Cai et al., 1997)
<i>Mi-1.2</i>	<i>Solanum peruvianum</i>	<i>Meloidogyne Incognita</i> , <i>M. javanica</i> , <i>M. arenaria</i>	CCc-NBSd- LRR	(Milligan et al., 1998a)
<i>Gpa-2</i>	<i>Solanum tuberosum ssp.</i> <i>Andigena</i>	<i>Globodera pallida</i>	CC-NBS-LRR	(Van Der Vossen et al., 2000)
<i>Hero A</i>	<i>Solanum</i> <i>pimpinellifolium</i>	<i>G. pallida</i> , <i>G. rostochiensis</i>	CC-NBS-LRR	(Ernst et al., 2002)
<i>Rhg4</i>	<i>Glycine max</i>	<i>H. glycines</i>	Serine hydroxymethyl transferase	(Liu et al., 2012)
<i>Gro1-4</i>	<i>Solanum spegazzinii</i>	<i>G. rostochiensis</i>	TIR-NBS-LRR	(Paal et al., 2004)
<i>Rhg1</i> locus	<i>Glycine max</i>	<i>H. glycines</i>	multiple genes*	(Cook et al., 2012)
<i>Mi-9</i>	<i>Solanum arcanum</i>	<i>M. incognita</i>	CC-NBS- LRR	(Jablonska et al., 2007)
<i>CaMi</i>	<i>Capsicum annuum</i>	<i>M. incognita</i>	CC-NBS- LRR	(Chen et al., 2007)
<i>Ma</i>	<i>Prunus cerasifera</i>	<i>M. incognita</i> , <i>M. arenaria</i> , <i>M. javanica</i>	TIR-NBS-LRR	(Claverie et al., 2011)

*Three genes are positioned in tandem within the *Rhg1* locus conferring resistance against *H. glycines*

1.7 Genetic resources for beet cyst nematode resistance and tolerance

The sources of resistance against the BCN are mainly *B. vulgaris* spp. *maritima*, *P. procumbens*, *P. patellaris* and *P. webbiana* (R. Viglierchio, 1960). The resistance genes originating from *P. procumbens* and *P. webbiana* are located on chromosomes 1, 7, and 8 were named as *Hs1^{pro-1}*, *Hs1^{web-1}*, *Hs2^{pro-7}*, *Hs2^{web-7}*, *Hs3^{pro-8}*, and *Hs3^{web-8}*. Whereas, the resistance gene from *P. patellaris* is located on chromosome 1 and was named as *Hs1^{pat-1}* (Lange et al., 1993; Löptien, 1984a; Löptien, 1984b; Reamon-Ramos and Wricke, 1992). Several attempts have been made to transfer resistance from these species to sugar beet via conventional breeding (Coons, 1975; Heijbroek et al., 1988; Löptien, 1984b; Savitsky, 1975). However, in most of the attempts, the interspecific hybrids were not fertile, therefore successful breeding for BCN resistance was never possible.

In a unique, ground-breaking attempt to transfer resistance from *P. procumbens* to *B. vulgaris*, Savitsky (1975) crossed a tetraploid *B. vulgaris* with diploid *P. procumbens* producing triploid offspring ($2n=3x=18+9$). Savitsky, then, backcrossed the triploid offspring with diploid sugar beet. The progeny of these crosses, harboring one extra chromosome from *P. procumbens*, termed as monosomic addition lines (MAL) ($2n=18+1$) were produced. Out of 6,750 plants, Savitsky selected four nematode-resistant trisomics. In a later study, Savitsky (1975) further backcrossed those resistant trisomics. Out of 8,834 plants, two diploid plants were found to be resistant against BCN, demonstrating that a *P. procumbens* chromosome segment carrying resistance was transferred to *B. vulgaris*.

1.8 Mapping and characterization of the translocation segment from *P. procumbens* to sugar beet

Further analysis of MALs revealed that three chromosomes of *P. procumbens* and one chromosome of *P. patellaris* bear resistance against BCN (Jung et al., 1986). Jung and Wricke (1987) crossed resistant MAL with susceptible *B. vulgaris*. The study found that 21% of the resulting progenies were resistant. Thereby, three diploid-resistant sugar beet plants, termed PRO1, PRO3, and PRO4, carrying a chromosomal segment from *P. procumbens*, were selected. Later, it was shown that PRO1 and PRO4 carry a translocation from chromosome 1 while PRO3 carries a translocation from *P. procumbens* chromosome 7 (Jung and Wricke, 1987; Van Geyt et al., 1988). Heller et al. (1996b) investigated the genetic locations of different translocation fragments originating from the wild beet genome in translocation lines using a segregating F₂ population and RFLP markers. Three resistance genes *Hs1^{pro-1}*, *Hs1^{web-1}*, and *Hs2^{web-7}* in four different translocation lines A906001, PRO4, WEB6, and WEB11 were analyzed (Heller et al. 1996). The results suggested that the translocation fragment was located at the end of chromosome 9 in the translocation lines (Heller et al., 1996b). A yeast artificial chromosome (YAC) library was developed to investigate the nematode resistance derived from the wild beet *P. procumbens* in the sugar beet translocation line A906001 (Kleine et al., 1995). The YAC library was later used to clone the nematode resistance gene *Hs1^{pro-1}* (Cai et al., 1997). The multi-color FISH technique was performed for the physical mapping of two differentially labeled YACs and the nematode resistance gene *Hs1^{pro-1}* originating from the translocation (Desel et al., 2001). The results from the comparative chromosomal mapping of the 684 bp *Hs1^{pro-1}* probe revealed that the translocation from wild beet was conferring resistance against BCN to a monosomic addition line 950039, the translocation line A906001 and *P. procumbens*. In addition, Desel et al. (2001)

suggested that the *P. procumbens* chromosome fragment carrying the resistance gene was translocated at the distal region of *B. vulgaris* chromosome 9 in the translocation lines.

Although cultivars carrying the translocations are commercially available, they are not fully exploited in the sugar beet breeding programs because of a substantial yield penalty which could be attributed to linkage drag of the genes within the translocation segment (Schulte et al., 2006). Therefore, partial resistance originating from *B. vulgaris* spp. *maritima* was exploited for breeding tolerant sugar beet varieties. Seven genetic loci were observed to provide resistance against BCN in *P. procumbens*, *P. patellaris*, and *P. webbiana*, as listed in Table 3. Out of these seven loci, only one gene, *HsI^{pro-1}*, was successfully cloned (Cai et al., 1997).

Table 3: Wild beet chromosomes carrying genes for beet cyst nematode resistance or tolerance.

Species	Chromosome	Gene	Resistance/Tolerance	Reference
<i>P. patellaris</i>	1	<i>HsI^{pat-1}</i>	Resistance	(Löptien, 1984b)
<i>P. webbiana</i>	1	<i>HsI^{web-1}</i>	Resistance	(Salentijn et al., 1992)
<i>P. webbiana</i>	1	<i>HsI^{web-1}</i>	Resistance	(Heller et al., 1996b)
<i>B. vulgaris</i> ssp. <i>maritima</i> WB242 ¹	7	<i>Hs^{Bvm-1}</i>	Tolerance	(Lange et al., 1993; Stevanato et al., 2015)
<i>P. procumbens</i>	7	<i>Hs2^{pro-7}</i>	Resistance	(Lange et al., 1993)
<i>P. procumbens</i>	1	<i>HsI^{pro-1}</i>	Resistance	(Cai et al., 1997)
<i>P. procumbens</i>	8	<i>Hs3^{pro-8}</i>	Resistance	(Jung et al., 1998)

1.9 Resistance genes and candidate genes from the *P. procumbens* translocation

The first resistance gene to be cloned from the translocation line TR520 (A90600), *HsI^{pro-1}*, was believed to be the gene conferring complete resistance against BCN (Cai et al., 1997). TR520 carries a translocation from *P. procumbens* chromosome 1. TR363 is an independent translocation line that also carries a chromosomal segment from *P. procumbens* chromosome 1 and is resistant to BCN (Jung and Wricke, 1987). Using differential display technique, a putative cation transporter gene linked to *HsI^{pro-1}* was found and proposed to play a role in the defense response and/or signal transduction (Oberschmidt et al., 2003). However, Schulte et al. (2006) showed that TR363 does not carry the gene *HsI^{pro-1}*. Therefore, it was concluded that *HsI^{pro-1}* confers only partial resistance and that chromosome 1 from *P. procumbens* carries another resistance gene that was yet to be discovered. For simplicity, the second gene will be, thereafter, called *Hs4* (formerly known as *HsI-1* (Schulte et al., 2006) or *HsI-2* (Capistrano, 2010; Jaeger et al., 2008)).

Schulte et al. (2006) sequenced 5 BAC clones, BACs 57, 132, 137, 123, and 67, spanning 580 kb to cover the overlapping region between TR520 and TR363 translocation lines. Schulte et al. (2006) showed that the left end of BAC57 and the right end of BAC67 were missing on the TR363 translocation segment but were present in TR520. Moreover, she suggested that the BAC67 was the candidate clone that spans the translocation breakpoint. Schulte et al. (2006) claimed that TR363 was not the suitable line because of the proximity of the new candidate gene *Hs1-1* to the telomeric regions, thereby possibly compromising the cloning process due to the presence of repetitive sequences towards the distal end of the chromosome.

To fine map *Hs4*, the TR520 resistant translocation lines were gamma-irradiated (Hans Harloff, personal communication) at IAEA, Seibersdorf, Austria. Two lines harboring smaller translocations were selected and named TR659 and TR320. TR659 harbors longer translocation than TR320. Infection tests demonstrated that both TR659 and TR320 are susceptible to BCN (Hans Harloff, personal communication).

Studies comparing the two resistant lines, TR520 and TR363, to the two-gamma irradiated susceptible lines directed the discovery of *Hs4*. For the several candidate genes that have been proposed (Table 4), resistance genes analogs (RGA), i.e. NBS-LRR genes, were the topmost candidates. Based on conserved motifs between previously cloned NBS-LRR like genes from other crop species, such as *Cre3*, *Mi*, and *Gpa2*, degenerated primers were designed and 12 expressed RGAs were identified by Tian et al. (2004). Full-length sequences of four (*cZR-1*, *cZR-3*, *cZR-7*, *cZR-9*) out of those 12 RGAs were determined. However, only *cZR-3* and *cZR-7* were proposed as the candidates for *Hs4* because they were the only ones present in the translocation region and common between TR520 and TR363. In the over-expression study, 35S-*cZR-3*-expressing hairy roots from the susceptible diploid sugar beet, 093161, were observed to have a significantly fewer number of cysts/females 4 weeks after inoculation of BCN, however, complete resistance was not attained (Tian, 2003).

Capistrano (2010) proposed *ORF702* encoding a β -1,3-galactosyltransferase. *ORF702* with putative function as *Avr9* elicitor response like protein is present in the translocation segments of TR520 as well as TR363. It was proposed as a candidate gene because of its likely function and expression in both resistant translocation line TR520 as well as *P. procumbens*. *ORF702* was cloned and functionally characterized by (Jäger, 2012). However, sugar beet hairy root clones overexpressing *ORF702* were not resistant against BCN. Therefore, it was concluded that *ORF702* is not the anticipated *Hs4*. Few other genes were proposed by (Jäger, 2012), *ORF801*, *ORF802*, *ORF803*, as best candidates for *Hs4*. Further expression and functional analyses of *ORF801*, *ORF802*, and *ORF803* were carried out. In the expression analysis of the three candidate genes, it was observed that *ORF801* and *ORF802* were also expressed in the susceptible line TR659, hence they were excluded as the speculated *Hs4*. *ORF803*, encoding phosphatidylinositol kinase, was only expressed in resistant lines and *P. procumbens* was chosen for further functional analysis. *ORF803* was knocked down using the RNAi technique (Fen Qiao, unpublished) Fen. Transgenic hairy roots obtained from resistant plants were still resistant despite expressing *OR803*-sRNA. Therefore, *ORF803* was also disregarded as possible *Hs4*. Three other candidates, *ORF901* encoding cationic amino acid transporter, *ORF902* encoding for calcineurin B-like protein (CBL), *ORF906* encoding histone deacetylase 2, were proposed as possible *Hs4* gene, but the genes were not functionally characterized. The list of all the proposed candidate genes/ORF as *Hs4* is listed in Table 4.

Table 4: List of proposed candidate genes/ORF's from resistant sugar beet translocation line. RGA: Resistance Gene Analogue; NBS-LRR: Nucleotide Binding Site-Leucine Rich Repeats; MAP3K: Mitogen Activated Protein 3 kinase

Gene / Annotation code	Putative function	Position on translocation verified*	Functional characterization
<i>cZR-1</i> ¹	RGA (NBS-LRR)	No	No significant reduction in females
<i>cZR-3</i> ¹	RGA (NBS-LRR)	Yes	Significant reduction in females
<i>cZR-7</i> ¹	RGA analogue (NBS-LRR)	Yes	No significant reduction in females
<i>cZR-9</i> ¹	RGA (NBS-LRR)	No	No significant reduction in females
ORF702 ²	β-1,3-Galactosyltransferase	Yes	No significant reduction in females ²
ORF801 ³	Serine/Threonine kinase	No	Not done
ORF802 ³	Phosphatidylinositol kinase	No	Not done
ORF803 ³	MAP3K	Yes	No significant reduction in females
ORF901 ⁴	Cationic amino acid transporter	Yes	Not done
ORF902 ⁴	Calcineurin B-like protein	Yes	Not done
ORF906 ⁴	Histone deacetylase 2	Yes	Not done

¹ Tian et al. (2004)

² Capistrano (2010)

³ Jäger (2012)

⁴ (Fen Qiao, Unpublished data)

*As per the original study

1.10 Plant proteases and their role in plant defense mechanisms

Proteases are crucial enzymes that can perform protein catabolism by hydrolyzing the peptide bonds of a protein and ultimately releasing simpler peptides or amino acids. Therefore, proteases degrade proteins and terminate their functions. Proteases are required in a variety of biological processes, including plant defense mechanisms (Thomas and van der Hoorn, 2018). In plants, the space outside the plasma membrane, known as an apoplast, acts as an important site of communication with its outer environment. Early interactions between a pathogen(s) and plant occur in the apoplast. The apoplast acts as the site for host colonization in many pathogens including bacteria, fungi, oomycetes, and nematodes, where they secrete effector molecules. Plants

constitutively or inducibly accumulate defense-related proteins including proteases in the apoplast (Hou et al., 2018). Extracellular defense signaling starts immediately after the recognition of the pathogen-effector molecules by the host proteins in the apoplast or plasma membrane.

Plants have evolved different sophisticated ways to interact and respond to different environmental stimuli. Plants can retain the signals/memory of the pathogenic attack, a process known as priming. This retention of the memory of previous attacks enables the plant to launch more rapid and amplified defense responses upon future attacks (Ramírez et al., 2013). In *Arabidopsis*, it has been shown that the serine protease SBT3.3 (Subtilase 3.3), a subtilase member of the S8 family, regulates the defense priming. *Arabidopsis sbt3.3* mutants were found to be impaired in priming of both gene expression and signaling activity and were hypersusceptible to both *Pseudomonas syringae* and the oomycete *Hyaloperonospora arabidopsidis*. Moreover, the activity of mitogen-associated protein kinase (MAPK) was enhanced upon the overexpression of *SBT3.3* gene. Interestingly, upon overexpression of *SBT3.3*, transcriptional activating epigenetic marks at SA-regulated genes, including the promoters of WRKY transcription factors were also increased (Meyer et al., 2016). Another study in *Arabidopsis* showed that an apoplastic protease known as the papain-like cysteine protease Cathepsin B (CathB), a member of the C1 family, acts as a positive regulator of hypersensitive response. In addition, multiple *CathB* genes contribute redundantly to basal resistance (Gilroy et al., 2007). In a close relative of tobacco, *Nicotiana benthamiana*, it was shown that CathB protein is secreted into the plant apoplast, where it gets activated (Gilroy et al., 2007).

CDR1 (Constitutive Disease Resistance-1), another apoplastic protease of the A1 family in *Arabidopsis*, has been also shown to be implicated in local and systemic defense signaling. Activation tagging of the *CDR1* gene resulted in the constitutive expression of pathogenesis-related (PR) genes in a salicylic acid-dependent manner and enhanced resistance to multiple *P. syringae* strains. An abolished PR gene expression was observed in *CDR1* active site mutants. This indicated that CDR1 generates an extracellular mobile signal capable of inducing defense responses (Xia et al., 2004). Similar apoplastic fluids that induce systemic defenses, were induced by *Oryza sativa CDR1 (OsCDR1)* in *A. thaliana*, which indicated that CDR1 activity might be conserved between species (Prasad et al., 2009).

After the successful invasion of the apoplast, the pathogen is still not successful in establishing itself in the plant cell. These responses include changes in gene regulation, metabolite biosynthesis, and induction of PCD. Two antagonistically acting cytosolic metacaspases, *MCI (Metacaspase-1)* and *MC2 (Metacaspase-2)* belonging to family C14 showed the regulation of HR in *Arabidopsis* (Coll et al., 2010). Induction of HR cell death induced by *P. syringae* carrying avrRPM1 has been shown to be positively regulated by *AtMCI* (Coll et al., 2010).

Protein homeostasis, a regulatory network for controlling the amount and concentration of different proteins is essentially controlled by the host plant proteome. Multiple subunits forming heptameric rings constituting three catalytic β subunits with distinct proteolytic activities composes the core particle of the proteome (Murata et al., 2009). The threonine protease PBA1/ β 1 of the T1 family, one of the catalytic subunits has been studied in the context of HR because of its caspase-3-like activity (Hatsugai et al., 2009). In tobacco, it has been shown that *PBA1* expression is induced following treatment with the fungal elicitor, cryptogein (Suty et al., 2003). Moreover, the *PBA1*-silenced plants showed a reduced expression of the two other catalytic subunits, PBB and PBE. Suppression of HR

upon the silencing of these subunits indicated a role of the proteasome in HR induction (Hatsugai et al., 2009).

The third type of proteases identified so far in plants is vacuolar proteases. Rupturing of the vacuole during HR results in the alteration of cytoplasm because of acidification, the release of lytic enzymes, and potential cell-death mediators (Suty et al., 2003). The vacuolar processing enzymes (VPEs/Asparaginyl endopeptidases/Legumains), members of the C13 family, regulate the integrity of the tonoplast during PCD. C14 and *RD21* (*Response to Dehydration 21*), the orthologous papain-like proteases from tomato and *Arabidopsis*, respectively, have been detected in the vacuole. It has been shown that in *N. benthamiana*, the silencing or overexpression of a C14 homolog causes an increase or decrease in the susceptibility to *P. infestans* (Bozkurt et al., 2011). However, *Arabidopsis rd21* knock-out lines lost the susceptibility towards the oomycete *H. arabidopsidis*. Nevertheless, these knock-out lines showed more susceptibility to *Botrytis cinerea* when whole plants were infected (Shindo et al., 2012).

The two other main components of the plant cell, the endoplasmic reticulum, and the Golgi network, play essential roles in the adaptability of the cell under stress. AtCEP1 is a specific papain-like cysteine endopeptidase that belongs to the C1A family. A C-terminal 'KDEL' sequence sequesters the protease within ER-derived compartments. It has been shown that the expression of *AtCEP1* was induced upon the infection with the fungus *Erysiphe cruciferarum*. In these plants, the cells undergo PCD upon penetration by fungal haustoria. On the other hand, the *Arabidopsis cep1* mutants showed hyper-susceptibility to *E. cruciferarum* with characteristically reduced PCD as compared to *CEP1* functional plants (Höwing et al., 2018).

All the above-mentioned proteases lack the knowledge of verified substrate or known role/ function for their identified substrate. The Golgi-localized SBT6.1 belonging to the S8 family presents a rare example of a protease with its known substrate in plant immunity. The transmembrane malectin-like receptor kinase FERONIA extracellularly perceives the rapid alkalization factor 23 (RALF23), converted into a mature signaling peptide via SBT6.1 (Srivastava et al., 2009). Upon recognition, RALF23 weakens the immune signaling via inhibition of PRR complex formation, leading to the restriction of excessive defense responses. It has been shown in *Arabidopsis* that the activity of SBT6.1 and RALF23 increases rapidly with the infection by *P. syringae*. It was suggested that the regulation of SBT6.1 could be a mechanism to fine-tune the immune response by rapidly controlling the abundance of mature RALF23 (Stegmann et al., 2017). Some of the essential plant proteases with their role in plant defense mechanisms are listed in Table 5.

Table 5: Plant proteases with known functions in plant defense response.

Gene	Plant species	Pathogen	Protease family	Sub-cellular localization	Reference
<i>SBT3-S1</i>	<i>S. lycopersicum</i>	<i>Manduca sexta</i>	Serine protease	Apoplast	(Meyer et al., 2016)
<i>CathB</i>	<i>N. benthamiana</i> , <i>A. thaliana</i>	<i>Erwinia amylovora</i>	Cysteine protease	Apoplast	(Gilroy et al., 2007)
<i>CDR1</i>	<i>A. thaliana</i>	<i>Pseudomonas syringae</i>	Aspartic protease	Apoplast	(Simões et al., 2007)
<i>Rcr3</i>	<i>S. lycopersicum</i>	<i>C. fulvum</i> , <i>Phytophthora infestans</i> , <i>Globodera rostochiensis</i>	Cysteine protease	Apoplast	(Krüger et al., 2002; Lozano-Torres et al., 2012; Rooney et al., 2005)
<i>AtMC1</i> , <i>AtMC2</i>	<i>A. thaliana</i>	<i>P. syringae</i>	Metalloprotease	Apoplast	(Coll et al., 2010)
<i>PBA1</i>	<i>A. thaliana</i>	<i>P. syringae</i>	DEVDase	Cytosol and nucleus	(Hatsugai et al., 2009)
<i>VPE</i>	<i>N. benthamiana</i> , <i>A. thaliana</i>	<i>P. syringae</i>	Cysteine protease	Vacuole	(Zhang et al., 2010)
<i>CEP1</i>	<i>A. thaliana</i>	<i>Erysiphe cruciferarum</i>	Cysteine protease	Endoplasmic reticulum	(Höwing et al., 2014)
<i>SIP</i>	<i>A. thaliana</i>	Lepidoptera	Serine protease	Thick-walled sieve elements	(Coursol et al., 2003)
<i>C14/RD21</i>	<i>S. lycopersicum</i> , <i>A. thaliana</i>	<i>P. infestans</i> , <i>Hyaloperonospora arabidopsidis</i>	Cysteine protease	Apoplast	(Boex-Fontvieille et al., 2015; Pogorelko et al., 2019a; Shindo et al., 2012)
<i>Mir1-CP/Mir1</i>	<i>Z. mays</i>	<i>Spodoptera frugiperda</i>	Cysteine protease	-	(Mohan et al., 2006; Pechan et al., 2002)
<i>PIP1</i>	<i>S. lycopersicum</i>	<i>C. fulvum</i> , <i>P. infestans</i> , <i>P. syringae</i>	Cysteine protease	Apoplast	(Ilyas et al., 2015; Tian et al., 2007)
<i>HvPap-1 CIA</i>	<i>H. vulgare</i>	<i>Magnaporthe oryzae</i> , <i>Tetranychus urticae</i>	Cysteine protease	-	(Díaz-Mendoza et al., 2014)

1.11 Aims, expectations, and objectives

The beet cyst nematode *H. schachtii* causes severe losses in sugar beet production worldwide. Sugar beet lines, TR520 and TR363, carrying a chromosomal translocation segment from the wild relative *P. procumbens* are resistant to BCN. Gamma irradiation of the resistance translocation line TR520 resulted in the lines TR659 and TR320 that are susceptible to BCN. Therefore, I expected that the nematode resistance gene against the sugar beet cyst nematode *H. schachtii* is located on the region shared by the resistant translocation lines TR520 and TR363 while absent in the susceptible lines TR659 and TR320.

The aims of my study are as follows:

1. to generate a complete sequence assembly of the translocation segment on the translocation line TR520,
2. to identify the translocation breakpoint,
3. to identify and characterize the candidate genes found in the translocation regions that are present in the resistance lines but absent in the susceptible lines (critical-region),
4. to identify the candidate genes, analyze their structure and expression profiles, and functionally characterize them through knockout and overexpression in the hairy root system.

2 Materials and Methods

2.1 Plant materials and growth conditions

The plant material used in this study is described in Table 6. Translocation line TR320 and TR659 were identified after a marker-based screening of 578 M₁ offspring of 400 Gy gamma-irradiated seeds of line 950631 are susceptible to the BCN (Hans Harloff, personal communication). The resistant hybrid NEMATA and the sensitive line 093161 were used for CRISPR-Cas and overexpression experiments, respectively. For DNA isolation, plants were grown in the greenhouse under long-day conditions (16h light/8h dark) at 22°C.

Plant material used for all experiments including their seed codes and phenotypes are shown in Table 6.

Table 6: Seed code information for all the seed material used in this project.

Name	Seed code	Species	Phenotype ^a	Source
TR520	100043	<i>B. vulgaris</i>	Resistant	Heijbroek et al. (1988)
TR363	930363	<i>B. vulgaris</i>	Resistant	C. Jung & Wricke, (1987)
TR659	121490	<i>B. vulgaris</i>	Susceptible	Hans Harloff (personal communication)
TR320	071320	<i>B. vulgaris</i>	Susceptible	Hans Harloff (personal communication)
093161	930176	<i>B. vulgaris</i>	Susceptible	Dieckmann GmbH & Co. KG
<i>P. procumbens</i>	950056	<i>P. procumbens</i>	Resistant	Seed bank, IPK
NEMATA	180819	<i>B. vulgaris</i>	Resistant	Syngenta Seeds GmbH

^aResistance against BCN

2.2 Nematode resistance tests

The beet cyst nematodes were propagated in the greenhouse under non-sterile conditions on susceptible sugar beet plants (093161). Fully developed brown cysts were harvested from the roots onto 50µm sieves. A 3 mM ZnCl₂ was used to stimulate the hatching of juveniles under dark. Nematodes were examined under a binocular microscope. Only suspensions with >90% mobile nematodes were taken as inoculum. For in vitro tests, nematodes were surface sterilized by soaking them in 0.05% HgCl₂ solution for 30 seconds, followed by four times washing with sterile water and resuspension in 0.2% (w/v) Gelrite (Duchefa Biochemie B.V., Harlem, Netherlands). Two hundred fifty sterile nematodes were used to inoculate the roots.

For greenhouse resistance tests, plants were grown in 20 ml tubes filled with sterile sand (grain size 0.1 - 1.5 mm), sterilized at 80°C for 3 hours. Six hundred freshly hatched second-stage juveniles (J2-larvae) were added to each plant with a syringe. After 4 weeks, plants were harvested and washed, and roots were examined under a binocular microscope. For further analysis, root samples were collected 3, 6, 9, and 12-days after inoculation, with three biological replicates per sample, and were stored at -80°C.

For in-vitro resistance tests, hairy roots of size 1 cm in length were transferred to fresh Petri-plates containing B5 media. Petri-plates with hairy roots were kept in the climate chamber (22°C at dark). After two to three weeks of growth, the hairy roots were inoculated with 250 sterilized *H. schachtii* Schach-0 J2 larvae. Cysts/L4-females were counted under the stereomicroscope four weeks after inoculation.

2.3 DNA isolation

Genomic DNA was isolated from leaves and hairy roots using the CTAB extraction method (Rogers and Bendich, 1985). Leaves and hairy root tissues were freeze-dried for 48hrs. Lyophilized tissues were ground using Geno/Grinder® (SPEX SamplePrep, New Jersey, United States). 65°C 2X CTAB buffer was added to the ground tissues. After incubation for 30 min at 65°C, an equal amount of chloroform/isoamyl alcohol (24:1) was added. After centrifugation, the supernatant was transferred to a new tube and mixed with an equal amount of isopropanol. After overnight incubation at -20°C, samples were centrifuged at 11000g. DNA precipitation was followed by two washing steps. DNA was resuspended in 100 µl ddH₂O containing RNase A.

2.4 RNA isolation

For RT-qPCR, RNA was isolated from root and leaf tissues, and hairy root clones were generated in the over-expression experiment. Root samples were harvested from inoculated plants 3, 6, and 9 days after inoculation (dpi) along with samples from non-inoculated plants. Roots were briefly washed with water to remove sand. Samples from RFP-positive hairy roots were collected before infection tests to measure relative gene expression. Leaves, roots, and hairy roots were shock-frozen in liquid nitrogen and stored at -70°C until further use. Frozen tissues were homogenized in four cycles for root tissues and two cycles for leaf tissues (2 min each cycle) using a Geno/Grinder® (SPEX SamplePrep, New Jersey, United States) according to the manufacturer's instruction. Samples were submerged into liquid nitrogen after each cycle to prevent them from thawing. RNA isolation and DNase treatment were carried out according to the instructions provided with the PeqGold Total RNA Kit (PEQLAB Biotechnologie GmbH, Erlangen, Germany). The quality of the RNA was checked by agarose gel electrophoresis and a NanoDrop2000 spectrophotometer (ThermoFisher Scientific, Waltham, United States) was used to calculate the concentration of the isolated RNA.

2.5 PCR and RT-qPCR

For the analysis of the breakpoint sequence, primers H203 and H208 were used. For subcloning of the AtU6-26-sgRNA cassette from pChimera to p201G vector, the primers rbd_gRNA_F and rbd_gRNA_R were used. PCR experiments were also carried out to confirm the presence of the transgene in genomic DNA from hairy root clones. For the CRISPR-Cas knock-out experiment, Cas1_f and Cas1_r primers were used. For the overexpression experiment, F_OEX and R_OEX primers were used. For genotyping of CRISPR-Cas mutagenized hairy root clones, rbd_10F and rbd_10R primers flanking the two target sites were used.

For RT-qPCR, one µg of RNA was used for cDNA synthesis using the First Strand cDNA Kit (Thermo Fisher Scientific, Waltham, United States). cDNA was produced by adding one µl of Oligo (dT)₁₈ to the RNA. The mixture was centrifuged and incubated at 65 °C for 5 min. Then, 5X reaction buffer, RiboLock RNase inhibitor, dNTP mix, RevertAid M-MuLV RT (reverse transcriptase) were added. The mixture was incubated at 42 °C for 60 min. The reaction was terminated by heating at 70 °C for 5 min.

RT-qPCR was carried out using Platinum™ SYBR™ Green qPCR SuperMix (ThermoFischer Scientific, Waltham, United States) using the primers hs4_F and hs4_R. Expression levels were calculated with the comparative $\Delta\Delta C_t$ method (Livak & Schmittgen, 2001). *Bv.GAPDH* was used as endogenous controls to normalize gene expression levels. Three biological replicates were taken for each sample.

The sequences of the primers used in addition to the annealing temperature are mentioned in Supplementary Table 1.

2.6 DNA sequencing

DNA isolation for Sanger sequencing was done using CTAB extraction method (Rogers and Bendich, 1985). Sanger sequencing was carried out for the breakpoint sequence analysis, in addition to identification of CRISPR-Cas induced mutations. Sanger sequencing was performed at Institute of Clinical Molecular Biology (IKMB, Kiel University, Germany).

2.7 DNA sequence libraries

For this study, the following sequence libraries were available. Two paired-end (PE) libraries, one mate-pair (MP) library and a RNA sequencing library for TR520, in addition to two PE libraries for TR363 and one PE library for TR659 (Jäger, 2012). I downloaded the whole genome sequencing (WGS) reads for the commercial sugar beet line KWS2320 from NCBI-SRA (SRR869754). Read quality was assessed using the quality control tool FastQC (version 0.11.5) (Andrews, 2010). The WGS data and transcriptome data that were used in this study are listed in Table 7.

Table 7: Whole genome sequencing/transcriptome sequencing data used in this study. PE: Paired-end, MP: Mate-pair, RNA-Seq: Transcriptome sequencing. The *P. procumbens* sequences were kindly provided by Dr. Chris Richards, (ARS USDA, USA).

Genotype	Number of sequence libraries	Type of NGS libraries	Read length (bp)	Number of reads (millions)	Reference	NCBI #
TR520	2	PE	101	214	(Jäger, 2012)	-
TR520	1	MP	101	139.8	Unpublished (Jäger, 2012)	-
TR520	1	RNA-Seq	101	81.3	(Jäger, 2012)	-
TR363	2	PE	101	213	(Jäger, 2012)	-
TR659	1	PE	101	178.5	Unpublished	-
KWS2320	1	PE	101	20.9	(Jäger, 2012)	SRR869754
<i>P. procumbens</i>	2	PE	150	367.9	Chris Richards, (personal communication)	-

2.8 Bioinformatic analysis

TR520 paired-end and mate-pair genomic DNA read libraries were *de-novo* assembled using SOAPdenovo2 assembler (version 2.04) (Luo et al., 2012) on a hybrid NEC high performance system at the Kiel University Computing Centre. The assembly was performed using a computing node with 384 GB of RAM. KmerGenie was used to predict the best k-mer length (Chikhi and Medvedev, 2013). Two *de-novo* assemblies, one with k-mer size 91 (TR520v1) and another one with k-mer size 61 (TR520v2), were generated. To fill the gaps within the scaffolds, the seed-and-extend local assembler GapFiller was used, using default parameters (Boetzer and Pirovano, 2012).

RepeatModeler based repeat annotation workflow was followed, using NCBI BLASTDB as input to the repeat modelling pipeline, to identify and classify the repetitive sequences in the translocation line TR520 (<http://www.repeatmasker.org>). NCBI standalone blast+ (version 2.9.0) was used for sequence search and comparison (Camacho et al., 2009). Prediction of protein coding gene structures was performed using the MAKER gene-annotation pipeline (blast_type = ncbi+, est = 1, protein = 1, cpu = 32), with Illumina RNA-seq reads as transcript evidence and published protein sequence from RefBeetv1.2.2 assembly as protein evidences (Holt and Yandell, 2011). Small and non-coding RNA were identified based on homology searches. Functional annotation was done using *de-novo* transcriptome assembly strategy based on Trinotate *de-novo* annotation pipeline (Bryant et al., 2017).

Configuration files including the parameters used by SOAPdenovo2 *de-novo* assembler and MAKER pipeline are available as supplementary data.

The subcellular localization was predicted using DeepLoc-1.0 webserver (<http://www.cbs.dtu.dk/services/DeepLoc-1.0/>) (parameter = protein encoding profiles), which uses convolutional neural network to predict the most probable sub-cellular localization of a protein based on its amino acid sequence.

Sequence similarity of the protein sequence to entries from the peptidases database MEROPS was assessed using the online BLASTP (default settings) service of EMBL-EBI (<https://www.ebi.ac.uk/Tools/sss/ncbiblast/>). BLASTP hits from MEROPS database were aligned using Clustal Omega (Madeira et al., 2019) (default parameters). The multiple sequence alignment was used to construct the phylogenetic tree which was then reformatted for better visibility using iTOL webserver (Letunic and Bork, 2019). Finally, the multiple sequence alignment was visualized in the multiple sequence alignment and analysis workbench Jalview (Waterhouse et al., 2009).

2.9 CRISPR-Cas mediated gene knockout and overexpression studies

For targeted (CRISPR-Cas) mutagenesis, two vectors, pChimera (Fauser et al., 2014) and p201G (Jacobs et al., 2015), were used. I selected two 20-bp target sequences (T1 and T2) located within the first exon of the *ORF1* gene and next to a 5'-NGG PAM site. For cloning into the pChimera vector, DNA oligonucleotides of the designed target sequences were synthesized having overhangs of the BbsI restriction site (Eurofins Genomics, Ebersberg, Germany). To avoid using antibiotics, I used the p201G vector (Addgene, Massachusetts, United States), containing the green fluorescence protein (GFP) encoding gene under the transcriptional control of the enhanced CaMV35S promoter as a reporter gene. The AtU6-26(P)-sgRNA cassette was amplified from the pChimera vector using the rbd_gRNA_F and rbd_gRNA_R primers. The PCR product and the p201G vector were digested with the I-PpoI restriction enzyme. Then, the digested PCR product was ligated

into the p201G plasmid by incubating the digested plasmid and PCR product with T4 ligase at 22°C for 1 hour (Figure 4).

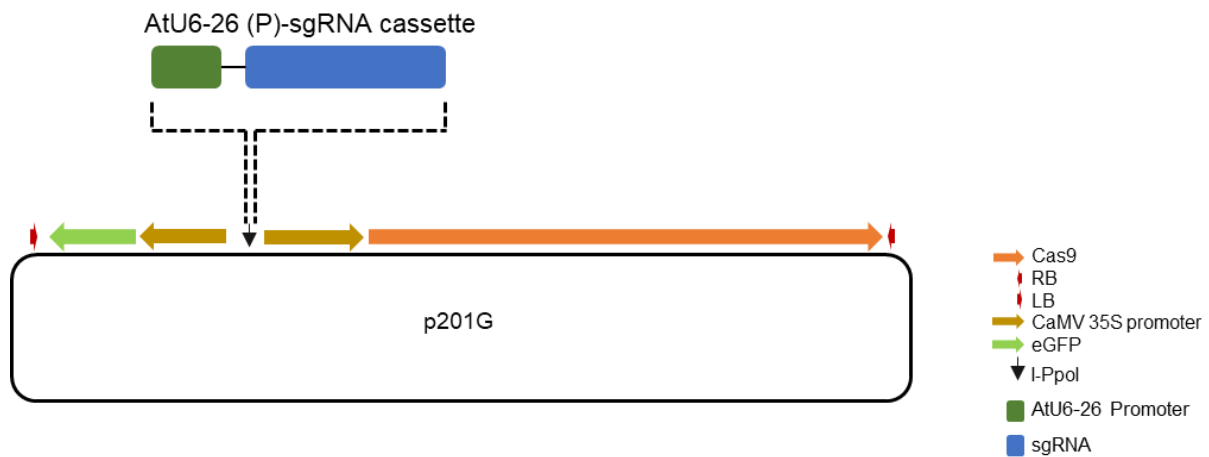


Figure 4: Sub-cloning of AtU6-26(P)-sgRNA cassette into p201G vector. The p201G plasmid carries Cas9 and eGFP encoding genes, each driven by a 35SCaMV promoter and an I-PpoI restriction site to clone the AtU6-26 (P)-sgRNA cassette. The cassette was amplified from the pChimera vector using rbd_gRNA_F and rbd_gRNA_R primers with I-PpoI restriction site overhangs and ligated into the p201G vector.

For overexpression of the *ORF1* coding sequence, the vector pBin35SRed was used (Pidkowich et al., 2007). pBin35SRed carries the gene encoding Red Fluorescent Protein (RFP) under the transcriptional control of the cassava vein mosaic virus (CsVMV) promoter as a selectable marker. The coding sequence of the *ORF1* gene was amplified by PCR using primers F_OEX and R_OEX with XbaI and EcoRI restriction sites. The PCR product and the pBin35SRed vector were digested with the EcoRI and XbaI restriction enzyme. Then, the digested PCR product was ligated into the pBin35SRed plasmid (Figure 5). The recombinant plasmids obtained (p201G-Cas-9-gRNA and pBin35SRed-*ORF1*) were transformed into *A. rhizogenes* using heat-shock approach (Quandt, 1993)

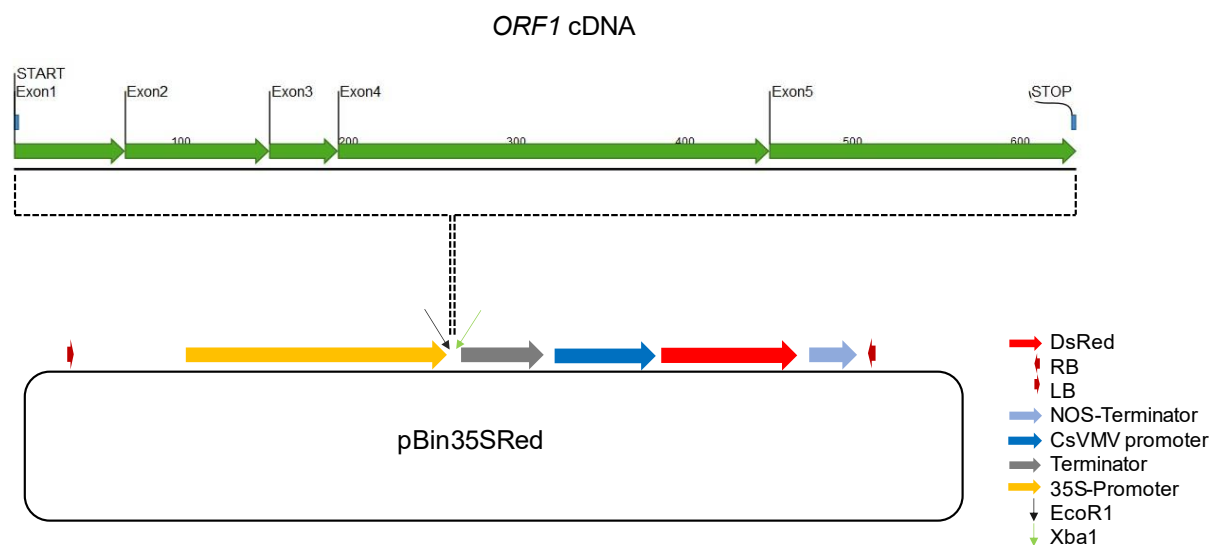


Figure 5: Cloning of *ORF1* cDNA into the overexpression vector plasmid pBin35SRed. The *ORF1* gene was cloned into pBin35SRed plasmid carrying the RFP encoding gene driven by the CsVMV (Cassava vein mosaic virus) promoter. *ORF1* cDNA was amplified using F_OEX and R_OEX primers with EcoRI and XbaI restriction sites overhangs. The PCR amplicon was ligated to the pBin35SRed vector behind the 35S promoter at the EcoRI/XbaI restriction sites.

2.10 Agrobacterium mediated hairy root transformation

Agrobacterium transformation and hairy roots production were carried out according to the protocol by Kifle et al. (1999). In brief, 1-2 cm leaf stalks from the resistant line NEMATA were cut and co-cultivated with *A. rhizogenes* containing the recombinant vector p201G-Cas-9-gRNA placed on B5 plates where they were kept in a climate chamber in the dark at 22°C for 48 hours. Afterward, the leaf stalks were kept at 20°C under long-day conditions until hairy roots were observed. Hairy roots 1-2 cm in size were cut from the leaf stalks and placed on fresh B5 plates containing cefotaxime (40 mg/l) (Duchefa Biochemie B.V., Harlem, Netherlands) and kept at 20°C under long-day conditions until hairy roots were observed. Hairy roots 1-2 cm in size were cut from the leaf stalks and placed on fresh B5 plates containing cefotaxime (100 mg/l). The same procedure was carried out for the over-expression of *ORF1*, where leaf stalks from the susceptible line 093161 were transformed by *A. rhizogenes* carrying the recombinant vector pBin35SRed-*Hs4*.

2.11 Microscopic studies

For fluorescence microscopy analyses, transgenic hairy root clones in sealed Petri plates were observed under a fluorescence stereomicroscope (Nikon SMZ25, Nikon Instruments Europe B.V., Netherlands). NIS-Elements BR microscope imaging software (Nikon Instruments Europe B.V., Netherlands) was used to capture images. Nikon GFP-B bandpass filter cube (λ Excitation = 470 nm) was used to detect hairy root clones expressing GFP. Nikon Texas Red bandpass filter cube (λ Excitation = 560 nm) was used to detect hairy root clones expressing DsRed (RFP). Images were taken with a 100 and 900 ms exposure time for detecting GFP and RFP, respectively. Further processing of the raw images was performed using the FIJI image processing package (Schindelin et al., 2012).

3 A complete sequence of the *P. procumbens* translocation in sugar beet

3.1 Genome assembly and translocation characterization

The genome of the translocation line TR520 was assembled using Illumina HiSeq2000 reads with depth of coverage around 110x and mate-pair reads library. The initial assembly (TR520v0) was obtained using SOAPdenovo2 (Luo et al., 2012) assembler with k-mer size of 91. To reduce the number of gaps, I further improved the assembly using GapFiller (Boetzer & Pirovano, 2012b). The number of gaps was reduced from 510,137 to 192,258. A final assembly of 614 Mb with 87.82% non-gapped sequences was obtained (TR520v1). Knowing that assembling with different k-mer sizes can help in resolving different repetitive sequences, a second *de-novo* assembly (TR520v2) for the translocation line TR520 (TR520v2) with k-mer size of 61 was generated. In addition to WGS reads of TR520, high quality raw reads (mean phred score = 38 and average read length = 101 bp) of the resistant line TR363 and the susceptible line TR659 were also available (Jäger, 2012).

To find the translocation-specific sequences, I followed three search strategies. First, to find the initial set of translocation-specific sequences, I used previously published translocation-specific molecular markers in addition to BAC and YAC sequences (Schulte et al., 2006). Based on the presence of repetitive sequences that are specific to *P. procumbens* (Dechyeva et al., 2003; Schmidt et al., 1990), I identified more translocation-specific scaffolds. To further elucidate the translocation segment, WGS paired-end reads from *P. procumbens* and KWS2320 sugar beet lines were mapped to TR520 assemblies, and read mapping coverage for both mapping files was calculated using genomeCoverage of Bedtools (Quinlan and Hall, 2010). Scaffolds on which reads from *P. procumbens* but not from KWS2320 were mapping, were assigned as translocation-specific scaffolds. Employing these strategies, 19 scaffolds from the first assembly TR520v1 and 42 scaffolds from the second assembly TR520v2 were identified. Using an all to all local BLASTN (word size = 91) search, in addition to manual visualization of mate-pair read mapping on IGV, both sets of scaffolds were then arranged, joined, or combined to get a consensus translocation sequence. Finally, two super scaffolds for the whole translocation, super-scaffold1, and super-scaffold2 of size 1.109 and 2.121 Mb, respectively, were generated. There is still a gap between super-scaffold1 and super-scaffold2. However, as indicated in the Desel et al. (2001) the translocation is at the very end of the chromosome. It indicates that gap should not be very big. However due to highly repetitive sequences either side of the gaps, it is not feasible to fill the gap using only short read sequencing technique. Detailed summary statistics of the *de-novo* assemblies are presented in Table 8.

To determine the position, size and characteristics of the translocation region, I used four BAC and four YAC clone sequences published in Schulte et al. (2006) to select the initial set of scaffolds that are specific for the translocation segment. Previously published *P. procumbens* specific repetitive sequences, pTS4.1, pRK643, pAp4 and pAp22, were localized onto 19 translocation specific sequences from TR520v1 using BLASTN alignment with e-value < 7.0E-4 and word size = 30 (Figure 6). Except for the highly repetitive regions of the translocation segments in TR363 and TR520, high sequence homology (> 95%) between the two translocations was observed since the translocation in both TR520 and TR363 originates from *P. procumbens*. The sequence analysis also

showed that certain regions of the translocation segment are absent in TR363 compared to TR520. This was expected as the two translocation lines originated from two independent translocation events. I used mate-pair mapping data to determine the adjacent scaffolds. Then, I manually curated and joined the translocation-specific scaffolds by inspecting the mate-pair library mapping data in IGV. A final translocation region assembly of 3.2 Mb, spanning 2 super scaffolds of 1,109 kb and 2,121 kb was generated (Figure 7A, Table 9).

Table 8: Whole genome assembly statistics for TR520.

Statistics	<i>de-novo</i> assembly k- mer size 91	Gap-filled assembly
Total bases	598,022 kbp	614,682 kbp
Total scaffolds/contigs	24,288	24,288
Largest scaffold	820,831 bp	838,629 bp
N50	97,583 bp	99,863 bp
N's	137,815,733 bp	74,832,912
#Gaps	510,137	192,258

A complete sequence of the *P. procumbens* translocation in sugar beet

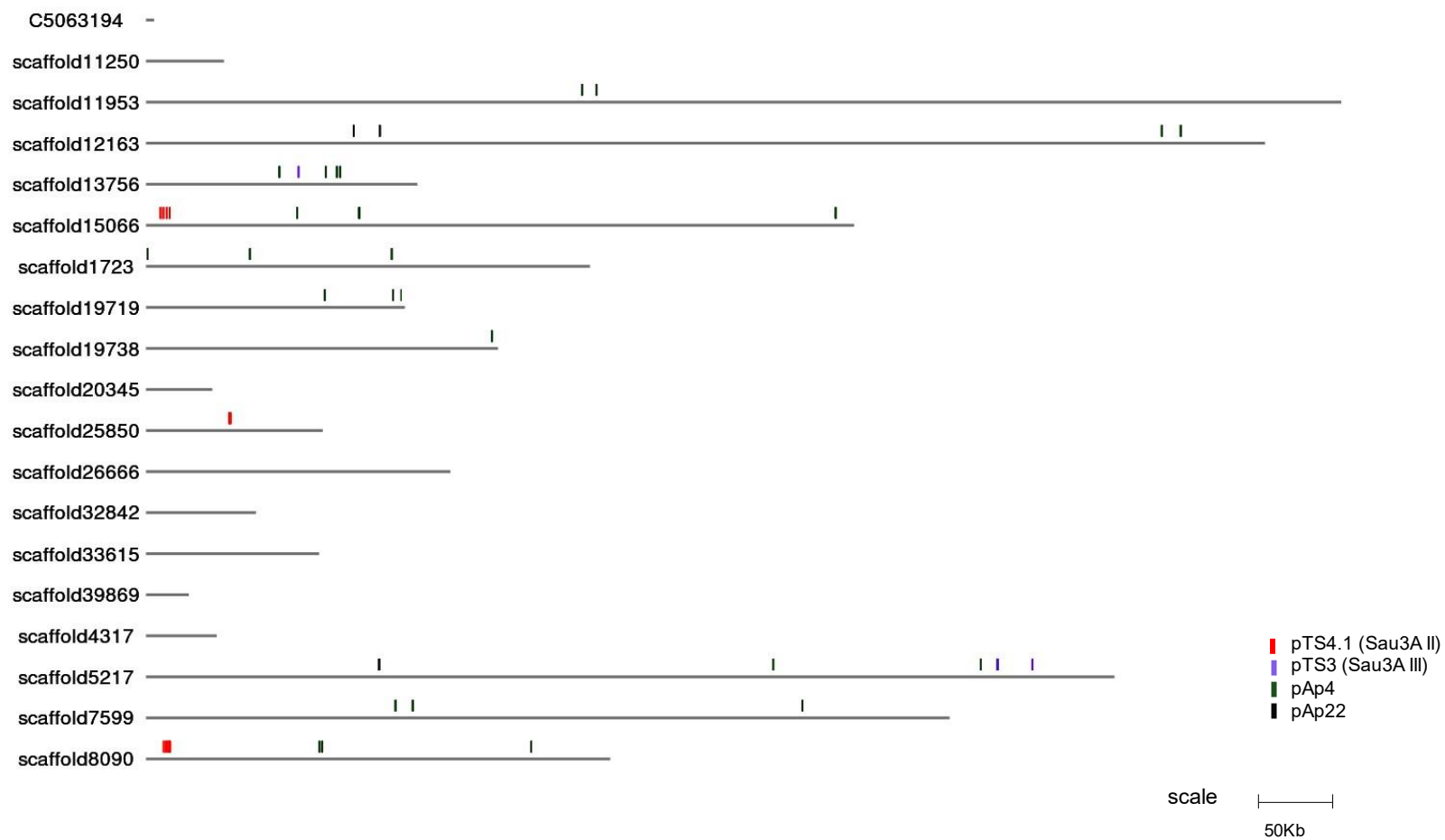


Figure 6: Localization of previously published *P. procumbens* specific repetitive elements on translocation specific scaffolds. pTS4.1 (Sau3A II) (red vertical bars), pTS3 (Sau3A III) (purple vertical bars), pAp4 (green vertical bars) and pAp22 (black vertical bars). Scale size is 50 kb. Each horizontal bar represents a scaffold.

A complete sequence of the *P. procumbens* translocation in sugar beet

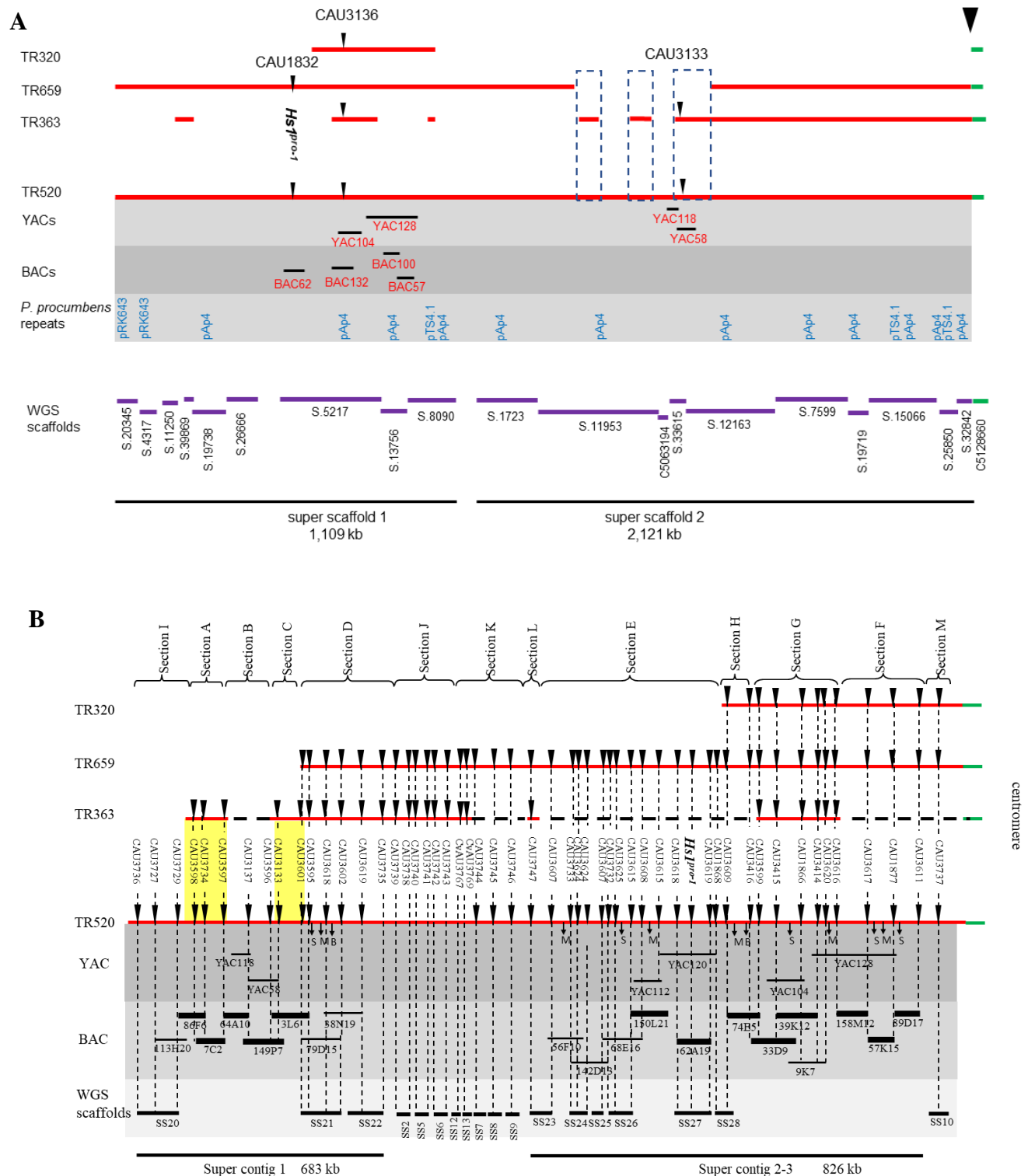


Figure 7: Comparison of the physical map of the *P. procumbens* translocation based on (A) this study, and (B) previous work of Jäger (2012). (A) Physical map of the *P. procumbens* translocation from the resistant sugar beet line TR520 based on WGS data, *P. procumbens* specific repetitive sequences and YAC and BAC clones, in comparison to translocation lines TR363 (resistant), TR659 and TR320 (both susceptible). Green bar depicts sugar beet, red depicts the *P. procumbens* translocation. The dotted lines depict the sequences only present in resistant lines. Four sequences each for YAC and BAC clones are shown as black horizontal bars. Blue labels are published *P. procumbens* specific repeats. Individual scaffolds originating from WGS are shown as purple bars along with the name of each scaffold. The gaps between the red bars for TR320,

A complete sequence of the *P. procumbens* translocation in sugar beet

TR659, and TR363 represent absent translocation regions. The breakpoint of the translocation is marked with a black inverted triangle. Translocation line TR520 possesses a translocation segment of size 3.2 Mb composed of two super scaffolds bearing sizes 1.109 kb and 2.121 kb. **(B)** A sequence-based physical map of the *P. procumbens* translocation of line TR520. The map was aligned with molecular markers (CAU numbers) derived from the bacterial artificial chromosome (BAC), and yeast artificial chromosome (YAC) ends and scaffold sequences obtained by WGS sequencing (see Figure 1). A minimal tiling path of YACs and BACs and the scaffolds from WGS sequencing are integrated into this map. ‘Sections’ denote translocation-specific sequences present or absent on different translocation lines identified by marker analysis. The term ‘super scaffold’ (SS) is used for assembled contigs and scaffolds of different genome assemblies of *P. procumbens* and the translocation lines TR520 and TR363 of the WGS data. ‘Super contigs’ are the largest types of sequence assemblies on the physical map and incorporate BACs and super scaffolds of the WGS sequencing, whereas the contigs of the physical map presented by Capistrano (2009) are named ‘BAC contigs’ in the following (compare Figure 1). Scaffold mapping is shown in Figure 14. Shaded areas highlight the different sequence resources mapped to the translocation: YACs, BACs and WGS scaffolds. Regions in common between both resistant translocations and absent from the susceptible translocation are highlighted in yellow.

Table 9: Scaffolds from the translocation region (TR520), their size and the number of genes/ORFs present in each. The presence of a sequence was concluded based on the read mapping coverage.

#	WGS scaffold (TR520)	Super contig	Size (bp)	Present in (coverage)			#ORFs
				TR659	TR363	P. <i>procumbens</i>	
1	Scaffold20345	1	27,744	100	0	100	0
2	Scaffold4317	1	29,534	100	0	100	0
3	Scaffold11250	1	31,664	100	0	100	0
4	Scaffold39869	1	17,587	100	100	100	1
5	Scaffold19738	1	148,538	100	40	100	16
6	Scaffold26666	1	129,237	100	0	100	13
7	Scaffold5217	1	413,137	100	55	100	34
8	Scaffold13756	1	115,661	100	0	100	11
9	Scaffold8090	1	196,039	100	10	100	10
10	Scaffold1723	2	189,950	100	12	100	10
11	Scaffold11953	2	505,473	10	38	100	29
12	C5063194	2	3,447	0	100	100	0
13	Scaffold33615	2	74,133	0	46	100	8
14	Scaffold12163	2	472,883	60	67	100	30
15	Scaffold7599	2	341,655	100	100	100	26
16	Scaffold19719	2	109,697	100	100	100	5
17	Scaffold15066	2	302,074	100	100	100	29
18	Scaffold25850	2	75088	100	100	100	3
19	Scaffold32842	2	47,221	100	100	100	4
Total:			3,230,762				229

3.2 Identifying the translocation breakpoint at the end of super scaffold 2

To identify the translocation breakpoint, I had two possible scenerios. The first scenario was that the breakpoint should lie within a scaffold whose one side shows high sequence similarity to *B. vulgaris* chromosome while the other side shows high sequence similarity to *P. procumbens* chromosome. The second scenario was that some mate-pair reads, will have one read from the pair mapping to a translocation-specific scaffold and the second read mapping to a sugar beet-specific scaffold. The first scenerio assumes that the DNA fragment containing breakpoint was not lost during sequencing library preparation and was not discarded during de-novo assembly. From a practical point of view, there are plenty of reasons that the first hypothesis can fail. However, in the second scenerio, such

limitations do not apply because the second scenario doesn't detect breakpoints directly but instead identifies the regions near breakpoints.

After analyzing the mapping coverage of WGS reads from both *P. procumbens* and KWS2320, none of the translocation-specific scaffolds of TR520 *de novo* assembly had sequences matching to both *P. procumbens* and KWS2320. Therefore, I disregarded the first hypothesis. The used mate-pair reads have an average distance of 5 kb; therefore, the probability that the breakpoint lies between two mate-pair reads was high. In principle, there can be a single read that includes the breakpoint, but it is almost impossible to find such a single read due to very high computational requirements. Therefore searching for such a read was not a feasible option. I then used the mate-pair reads and identified the breakpoint by scaffold walking. In the process of finding adjacent scaffolds, mate-pair scaffold joining pairs were identified by looking at both ends of the scaffold in IGV. Certain mate-pair reads which mapped to the translocation specific scaffold, super-scaffold2, had mates mapping to the contig C5128660. As no reads from *P. procumbens* were mapped to contig C5128660 while reads from KWS2320 did, I concluded that this scaffold does not belong to the translocation region and specific to sugar beet. The fact that the translocation-specific super-scaffold2 joins the sugar beet-specific contig C5128660, was the first indication that the genomic region joining these two scaffolds harbors the translocation breakpoint. Therefore, PCR primers were designed to amplify the region from position 2,117,983 bp on super-scaffold2 to 195 on contig C5128660. I expected that the size of the PCR amplicon would be below 5 kb, which is the average read-pair distance. The PCR was carried out using genomic DNA from TR520, TR363, *P. procumbens* and sugar beet (093161). A band of 567 bp was observed for TR520. While a smaller band of size ~500 bp was observed for TR363, instead of expected band of size 567 bp. No band was observed for sugar beet nor for *P. procumbens* (Figure 8). The PCR product from TR520 was sequenced by Sanger sequencing. A BLASTN search showed that nucleotides 343 bp to 543 bp displayed complete homology to sugar beet (Bvchr9_un.sca002) whereas 1 bp to 342 bp displayed complete homology to super scaffold 2 from the translocation. This experiment showed the exact position of the translocation breakpoint between super-scaffold2 and contig C5128660 via Sanger sequencing PCR product at position 341 bp from forward primer (Figure 9).

The sudden drop in mapping coverage of WGS reads from TR520 mapped to RefBeetv1.2.2, at around 420 kb in Bvchr9_un.sca002 (Supplementary Figure 1), further confirmed that this was the translocation breakpoint. Interestingly the mapping coverage did not drop to zero, but it was almost half of the initial coverage. It shows that half of the genetic material of this region is still present in the genome represented by these reads. Therefore, I concluded that the translocation is hemizygous.

From this experiment, I could locate the proximal end of the translocation segment. However, it was not possible to resolve the assembly as I move towards the telomere. Therefore, the current size of the translocation segment does not represent the complete translocation. However, as shown by Desel et al. (2001), using FISH, the translocation is present at the end of chromosome 9. Furthermore, it is well known that close to telomere are highly repetitive sequences. Therefore it is safe to say that no genes are located close to the telomere. For further experiments, I used the identified translocation sequence of 3.12 Mb.

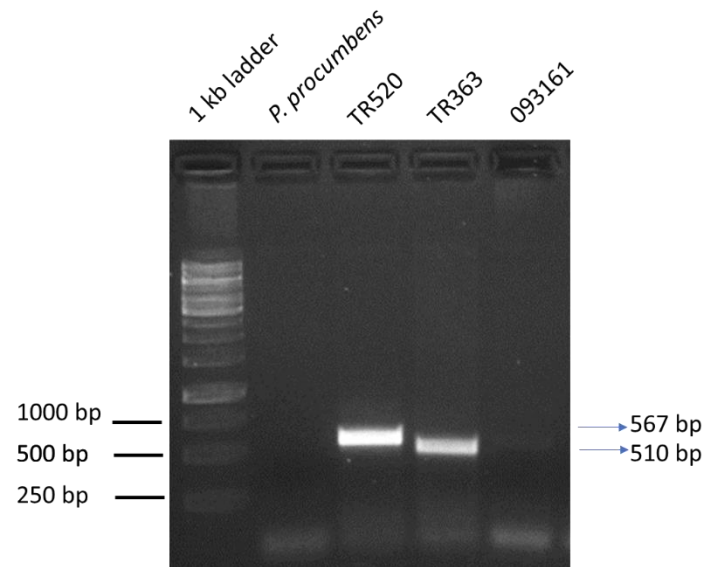


Figure 8: Gel electrophoresis with PCR products using a primer combination flanking translocation breakpoint. Genotyping the translocation breakpoint with two flanking markers (primer combination H208/H203). The primers were tested in four different genotypes, *P. procumbens*, translocation line TR520, translocation line TR363 and the sugar beet line 093161. A comparatively smaller band was observed in TR363 and no amplification was observed in *P. procumbens* and 093161. 1 % gel, 30 min., 90 V.

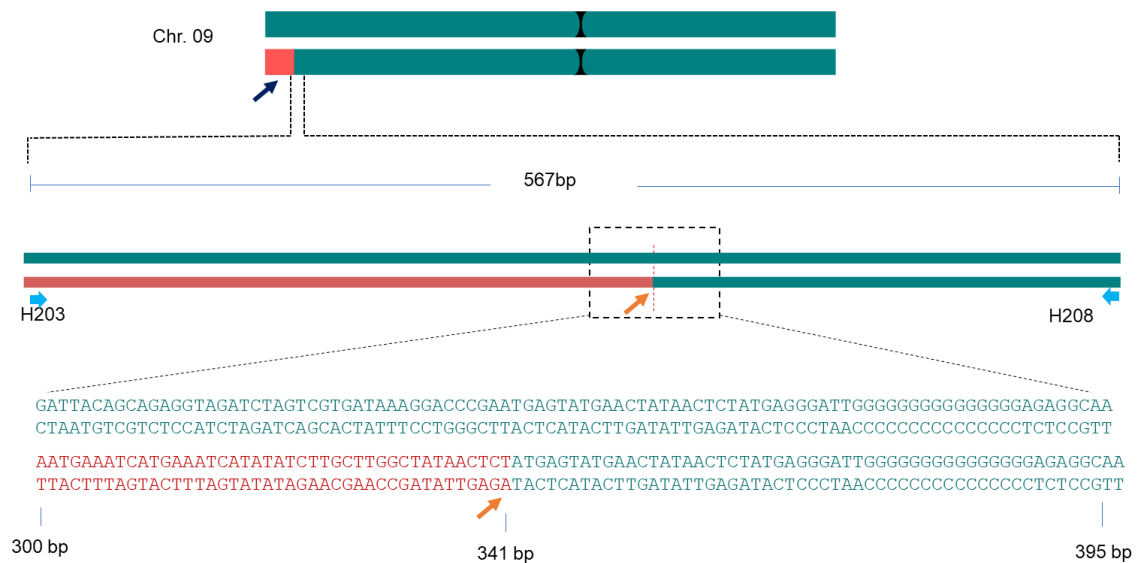


Figure 9: Schematic representation of the translocation breakpoint in TR520. The dotted blue box depicts the position of the breakpoint. Black arrow indicated the position of the translocation at the end of chromosome 9 of translocation line TR520. Black arrow indicates the presence of translocation segment at the end of one of the homologous chromosome 9 of sugar beet. Orange arrows indicate the position of translocation breakpoint. Blue arrows indicate the positions of the primers used to amplify the region.

4 Hs4 candidate gene identification

4.1 Identification of the critical regions

I aimed to identify the critical regions, which are defined as parts of the translocation segment present in both the resistant TR520 and TR363 lines but absent in the susceptible TR659 line. To determine the positions of the critical regions, Illumina paired-end reads from TR363 and TR659 were mapped to the TR520 genome assembly using BWA-MEM, with a mapping quality >20 (Li, 2013). The number of reads per base was calculated using genomeCoverage tool from Bedtools (Quinlan and Hall, 2010). As TR520 carries the longest translocation, critical regions were considered where reads originating from the gamma-irradiated TR659 did not map while reads from TR363 mapped, as shown by the blue dotted box in Figure 7A.

The average read-depth coverage of the TR659 and the TR363 libraries were 60x and 30x, respectively. However, in scaffold11953, at position ~105 kbp, the depth of coverage per base for TR659 decreased from ~60x to 0x. As I further analysed the coverage for TR659, I found that zero coverage for TR659 persisted throughout the next contig C5128660 and scaffold33615 and over a segment of ~110 kbp in scaffold12163 (Figure 10). Zero coverage for TR659 indicated that these regions are missing in TR659. I then analysed the read-depth coverage for TR363 in these regions and found that the average read-depth is 30x. On scaffold11953 from 267,279 bp to 308,200 bp and from 456,657 bp to 512,869 bp, on contig C5063194 from 0 bp to 3,447 bp, on scaffold33615 from 0 bp to 36,121 bp and on scaffold12163 from 7,615 bp to 44,847 bp and 156,162 bp to 185,888 bp, the coverage of TR363 was close to the average read-depth coverage. Therefore, I concluded that these regions are present in TR363 (Figure 10). Based on the coverage analysis from TR659 and TR363, I could identify the critical regions. I then used karyoploteR (Gel and Serra, 2017) to create a linear scaffold representation of the translocation region with the read-depth data of TR659 and TR363 (Figure 10). Usually, *de-novo* assemblers insert gaps (long stretches of N's) at unresolved places mainly due to repetitive sequences. Therefore, I assumed that it is highly improbable that a gap harbors a protein-encoding gene. Hence, gaps within assembled scaffolds were not considered as critical regions. Based on the read-depth coverage analysis, I identified critical regions encompassing a total size of 229 kb. I used the quality assessment tool for genome assembly, QUILT, to evaluate the sequences in the critical regions. QUILT identified 91 gaps spanning 17,502 bp within the critical region, 7.6% of the critical regions.

Table 10: Scaffolds with high sequence homology between TR520 and TR363. These sequences were named 'critical regions', because they were not present in susceptible TR659.

Scaffold ID	Start - End (bp)	Size (bp)	Sequence homology (%)
Scaffold11953	260,621 - 307,644	40,921	99%
Scaffold11953	449,496 – 505,473	56,212	98%
C5063194	0 – 3,447	3,447	99%
Scaffold33615	0-36,504	36,121	99%
Scaffold12163	7,665 – 41,750	37,232	99%
Scaffold12163	153,194 – 186,220	29,726	99%

Hs4 candidate gene identification

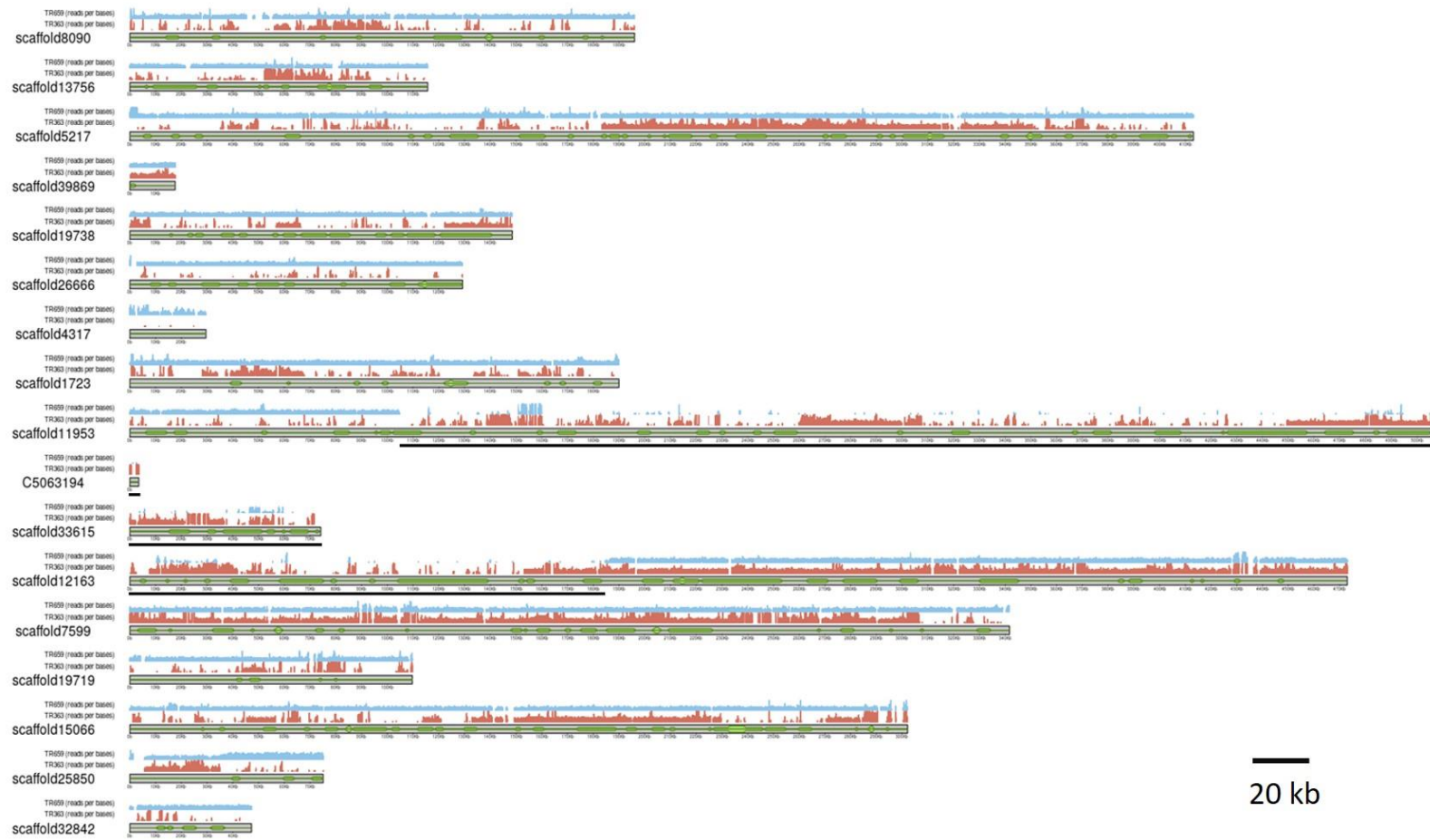


Figure 10: Identification of the critical regions. The ideogram shows sequence coverage of reads from TR363 (red) and TR659 (blue) mapped to the translocation region of TR520. The regions on scaffolds 11953, 33615, and 12163, where sequence reads from TR363 (red) are present while sequence reads from TR659 (blue) are absent, were considered to be the regions of interest. Critical regions are the regions within the black lines, where there is continuous presence of solid red vertical bars but no blue bars. Each horizontal bar represents one scaffold/contig of the translocation region. The green boxes within the bars represent the ORFs found in each scaffold. Scale bar = 20 kb.

4.2 Searching for ORFs within the critical regions

To find the genes in the critical region, I first annotated the whole translocation region. For structural annotations, I used the MAKER gene annotation pipeline (Holt and Yandell, 2011). RNA-seq reads were used by MAKER as transcript evidence and published protein sequences from RefBeet-1.2.2 assembly as protein evidence. Functional annotation was done using a *de-novo* transcriptome assembly strategy based on the Trinotate *de-novo* annotation pipeline (Bryant et al., 2017) using UniProt protein database, InterPro database, and Pfam database. In total, 229 protein-coding gene models were annotated within the translocation region.

4.3 Candidate gene identification and characterization

Out of the 229 gene models annotated from the translocation region, 33 were present in the critical regions, out of which 19 showed transcript evidence. None of the 33 genes showed homology to resistance gene analogues with NBS-LRR type structure or to any other known resistance gene. Genes present in the critical regions and have transcript evidence along with their location are listed in Table 6. Unexpressed ORFs were not considered as possible candidates. Putative functions were assigned based on sequence homology search against the UniProt, InterPro, and Pfam databases. Interestingly, none of the genes found in the critical regions were annotated as resistance gene analogues. Therefore, I concluded that the *Hs4* is not a resistance gene analogue.

Table 11: Genes present in the critical region. Gene ID is based on the most similar hit on the UniProt database. UniProt description is the description of Gene ID on the UniProt database. ORF number is given based on the position of ORF in the critical region. Previous ID, ID used in previous studies for the similar ORF.

Gene ID	UniProt description	InterPro ID	Pfam ID	Locus	ORF	Previous ID
rb11	Rhomboid-like protein 11	IPR022764	PF01694	scaffold12163:176733-182590	1	-
rh2b	Reticulocyte binding protein 2 homolog b	IPR012862	PF00795	scaffold12163:154933-156686	2	-
degp8	Protease Do-like 8	IPR001478 IPR009003	PF13180 PF13365	scaffold12163:104167-139066	3	-
anac094	Putative NAC domain-containing protein 94	IPR003441	PF02365	scaffold12163:151358-152934	4	-
kab1	Probable voltage-gated potassium channel subunit beta	IPR023210	PF00248	scaffold12163:78663-79887	5	-
ddi1	DNA damage-inducible protein 1	IPR000626 IPR001995 IPR015940	PF00240 PF09668 PF00627	scaffold12163:93151-95009	6	-

Hs4 candidate gene identification

at1g09760	U2 small nuclear ribonucleoprotein A'	IPR032675	PF14580	scaffold12163:58167-74947	7	-
topbp1	Topoisomerase (DNA) II binding protein	IPR001357	PF12738	scaffold12163:40881-45437	8	-
-	-	-	-	scaffold12163:39378-39780	9	-
ddi1	DNA damage-inducible protein 1	IPR000626 IPR001995 IPR015940	PF00240 PF09668 PF00627	scaffold12163:29298-30342	10	-
metap1	-	-	-	scaffold12163:21733-22102	11	-
fhy3	Protein FAR-RED Elongated hypocotyl 3	IPR018289	PF10551	scaffold12163:5057-5798	12	-
tah18	NADPH-dependent diflavin oxidoreductase 1	IPR002048 IPR011992	PF13202 PF13499	scaffold11953:488434-504646	13	-
b3galt12	Probable beta-1,3-galactosyltransferase 12	IPR013057	PF01490	scaffold11953:464428-474586	14	ORF702 (PhD. Thesis: Sarah Jager)
irt3	Fe(2+) transport protein 3	IPR005162	PF03732	scaffold11953:450591-456502	15	-
polx	Retrovirus-related Pol polyprotein from transposon TNT 1-94	IPR002659	PF01762	scaffold11953:483984-484606	16	-
log8	Cytokinin riboside 5'-monophosphate phosphoribohydrolase	IPR003689	PF02535	scaffold11953:298663-300030	17	-
slc38a6	Probable sodium-coupled neutral amino acid transporter 6	-	PF14223	scaffold33615:30649-32915	18	ORF901 (Fen Qiao, Unpublished data)
cb110	Calcineurin B-like protein 10	IPR031100	PF03641	scaffold33615:15329-23071	19	ORF902 (Fen Qiao, Unpublished data)

Among the 19 expressed genes found in the critical regions, none of the genes were shown to be involved in plant defense mechanisms. ORF14, earlier known as ORF702 (Capistrano, 2010), encoding beta-1,3-galactosyltransferase, was the only gene that was shown to be involved in defense mechanisms in *Vitis vinifera*. However, it was already

functionally characterized by Jäger (2012) and found not to confer resistance against BCN. Additionally, ORF803 speculated to be involved in BCN resistance was functionally investigated by Fen Qiao (Unpublished data) but did not show any phenotypic effect as per RNAi studies. Moreover, based on my work, I found that this gene is absent in any critical regions and was present in the susceptible line TR659.

As none of the genes in the critical regions showed homology to a resistance gene homologue, I carried out a careful literature survey to investigate whether any of the 19 ORFs are involved in plant defense mechanisms. Interestingly, proteases have been shown to be involved in plant defense mechanisms in several species. For instance, a 33 kDa cysteine protease was shown to confer resistance against Lepidopterans in maize (Pechan et al., 2000). In tomato, the papain like-cysteine protease Rcr3 was also shown to be involved in Cf-2 mediated resistance against *Globodera rostochiensis* (Lozano-Torres et al., 2012). In addition to the role of proteases in plant defense mechanisms, a serine protease Sep1 secreted by *Bacillus fermus* DS-1 has been shown to have a nematocidal activity. Sep1 could degrade multiple intestinal and cuticle-associated proteins in the nematodes, thus inducing digestive toxicity (Geng et al., 2016).

Two genes encode serine protease according to the UniProt database search. However, Protease Do-like8 (ORF3) was shown to be involved in the degradation of photo-damaged photosystem II reaction in Arabidopsis (Sun et al., 2010). The other protease was a rhomboid-like protease (ORF1). Rhomboid proteases are intramembrane serine proteases that hydrolyze peptide bonds within a cell membrane (Urban and Dickey, 2011). Although rhomboid-like proteases have not been shown to be involved in plant defense mechanisms, they are involved in developmental and innate immune responses in animals. Moreover, in the nematode parasitic fungus *Pochonia chlamydosporia*, rhomboid proteases have been expressed during the endophytic phase (Larriba et al., 2014).

To find other homologs of ORF1, I performed a BLASTP search of the ORF1 polypeptide sequence against the UniProt database. The results showed the highest similarity of 60.4% shared with two polypeptides. In addition, two proteins bearing identities A0A0J8FQU9 and A0A0K9RQE5 were observed as Rhomboid domain-containing proteins reported from *Beta vulgaris* (UniProt ID: BVRB_2g031850) and *Spinacia oleracea* (UniProt ID: SOVF_040520), respectively (Figure 11).

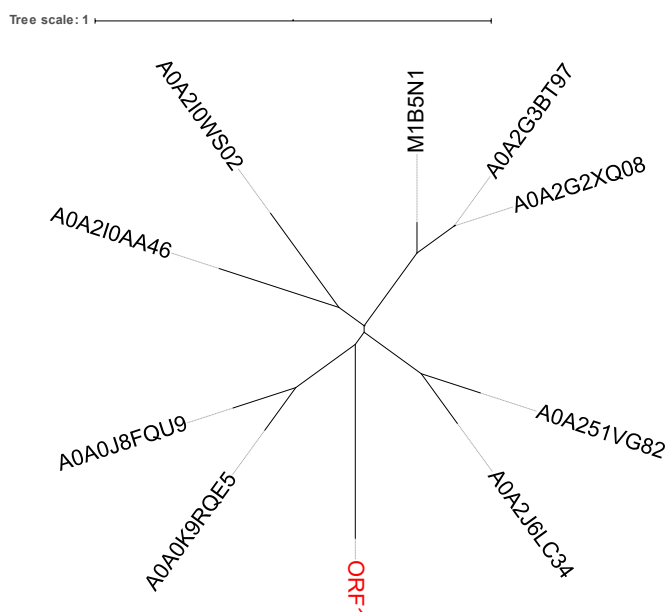


Figure 11: The unrooted phylogenetic tree of ORF1 with top nine BLASTP hits from UniProt database. The two closest homologs (60% identity) to ORF1 are SOVF_040520, a rhomboid-domain encoding gene from *Spinacia oleracea* (A0A0K9RQE5), and BVRB_2g031850, which is a rhomboid-domain encoding gene from *Beta vulgaris subsp. vulgaris* (A0A0J8FQU9). The phylogenetic tree was constructed using Clustal Omega and edited in iTOL. Branch lengths, corresponding to the number of substitutions per site, can be inferred by comparing to the given scale.

4.4 *In-silico* analysis of ORF1

According to the structural annotation by MAKER, the genomic sequence of *ORF1* gene covers 5786 bp, with 5 introns and 5 exons representing a coding sequence of 633 bp (Figure 12A, B). The polypeptide encoded is 210 aa in length. It carries the Rhomboid family domain between 112 bp to 582 bp (e-value = $1.63e-12$) (Figure 12C). To predict the protein features of ORF1, I used the fast and sensitive homology search webtool, HMMER (Potter et al., 2018). HMMER predicted the protein domains by comparing ORF1 protein to the Hidden Markov model (HMM) profile libraries. According to HMMER the ORF1 protein encompasses five transmembrane domains, one signal peptide, and two active sites (Figure 12C).

To further investigate the ORF1 protein in the context of other known proteases, I performed a BLASTP search of the ORF1 protein sequence against the MEROPS database. This database serves as an information resource for proteases and their corresponding inhibitors (Rawlings et al., 2017). Top 20 hits based on ascending e-values from MEROPS database search were used for multiple sequence alignment and phylogenetic analyses (Rawlings et al., 2010). None of the hits were observed to be involved in plant defense mechanisms. Using multiple sequence alignment, I identified a nine amino acid sequence (LLRDRCPDN), from 122 to 130 aa, specific to ORF1 protein (Supplementary Figure 2). Sequence alignment of ORF1 with its closest homologue in sugar beet (UniProt ID: A0A0J8FQU9) was carried out (Figure 13). ORF1 is 210 aa long, while A0A0J8FQU9 is 307 aa long. ORF1 protein and A0A0J8FQU9 shared 60.4% identity and shared the same active sites. The nine amino acid sequences that were shown to be specific to ORF1 were also absent in A0A0J8FQU9.

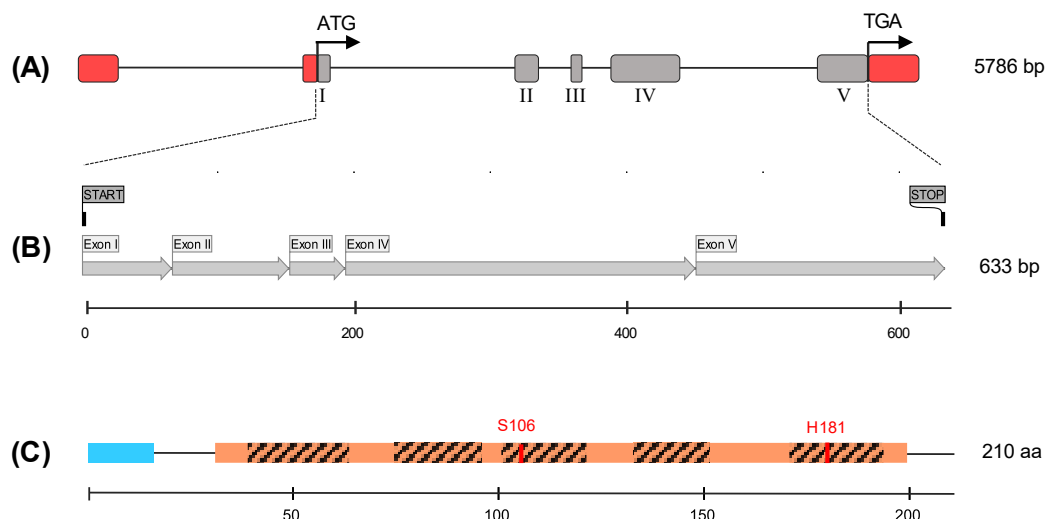


Figure 12: *In-silico* prediction of the *ORF1* gene structure and protein domains. **(A)** Structural analysis of *ORF1*. *ORF1* contains 5-exons (grey boxes) and 5-introns (black line). The 5' and 3'-UTRs are depicted as red boxes. ATG: Translation start site, TGA: Translation stop site. **(B)** *ORF1* coding sequence. *ORF1* coding sequence is a linear combination of 5 exons covering 633 bp. **(C)** *ORF1* polypeptide sequence. The signal peptide is represented by the blue box. Rhomboid family domain is represented by an orange box. Five transmembrane domains are represented by black stripes, and two active sites of *ORF1* protein are shown by red bars (S106 and H181).

To predict the sub-cellular localization of *ORF1* protein, I performed an *in-silico* analysis of *ORF1* using DeepLoc web server. The *ORF1* protein sequence was used as a query. DeepLoc predicted that the most probable sub-cellular location of *ORF1* protein is the endoplasmic reticulum (ER) membrane. The probability of *ORF1* being a membrane-bound protein was 0.99, whereas the probability of *ORF1* being localized to ER was 0.5825 (Table 12). Unlike *ORF1*, according to DeepLoc A0A0J8FQU9 is localized to the plastid membrane. In addition, A0A0J8FQU9 has only four transmembrane domains within the rhomboid family domain. The relatively low sequence homology and the different sub-cellular localization and the difference in the number of transmembrane domains suggested that *ORF1* and the closest homolog in the sugar beet genome are functionally different.

Table 12: Likelihood of *ORF1* sub-cellular localization as calculated by DeepLoc. DeepLoc predicts the subcellular localization of eukaryotic proteins.

Localization	Likelihood
Endoplasmic reticulum	0.5825
Golgi apparatus	0.2871
Lysosome/Vacuole	0.0881
Cell membrane	0.0288
Plastid	0.0065
Mitochondrion	0.006
Extracellular	0.0006
Cytoplasm	0.0002
Peroxisome	0.0001
Nucleus	0.0001



Figure 13: Sequence alignment of ORF1 protein with its closest homologue in the sugar beet genome (A0A0J8FQU9). ORF1 is 210 aa long, while A0A0J8FQU9 is 307 aa long. Two active sites of ORF1 protein are shown by red asterisks (S106 and H181). The signal peptide is marked by the green line. The rhomboid family domain is marked by the blue line. Brown line indicates the insertions

4.5 Transcriptional analysis of ORF1

Before I analyzed the quantitative expression via RT-qPCR in roots of NEMATA (3, 6, 9 and 12 dpi), I expected that *ORF1* is expressed only in response to nematode infection. Despite the slight increase of the transcriptional activity at 9 dpi, no significant difference was observed between infected and non-infected roots. This indicates that the *ORF1* gene activity is not altered upon nematode infection. Contrary to my expectations, I concluded that *ORF1* was stably expressed in roots before and after infection (Figure 14).

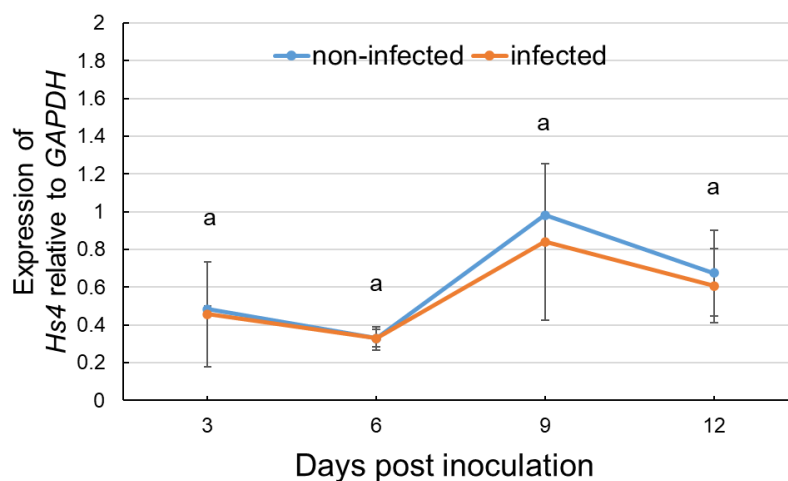


Figure 14: Expression analysis of *ORF1* gene in infected and non-infected roots. Quantitative expression analysis of *ORF1* gene in response to nematode infection in roots from the resistant translocation line 3, 6, 9 and 12 dpi. Gene expression was quantified relative to *BvGAPDH*. Error bars are defined by the SEM of three biological replicates. No statistical significance was observed between the stages. Statistical significance was analyzed applying ANOVA.

5 Functional analysis of the ORF1

To functionally characterize the gene activity of *ORF1*, I used two approaches: knock-out and overexpression studies. After *ORF1* was selected as the most promising candidate, I aimed to produce *ORF1* functional mutants by knockout via CRISPR-Cas induced targeted mutagenesis.

5.1 *ORF1* knockout mutants produced by CRISPR-Cas9 mutagenesis

To assure a complete knock-out of *ORF1*, I searched for target sites within the first exon of *ORF1*. I looked for sequences as per parameters described by Wang et al. (2019) to enable the most optimum nucleotide composition for cleavage by Cas9 endonuclease. I expected that hairy root clones generated from a resistant sugar beet line possessing a non-functional *ORF1* would be susceptible to BCN. The selected target sites, designated T1 and T2, were 20bp long with the upstream protospacer adjacent motif (PAM) 5'-'NGG'-3' (Figure 16A). Then, I looked for possible off-target binding sites using the BLASTN search conducted against the TR520 genome and the RefBeet-1.2.2 genome. The selected target sites were highly specific since no off-target hits were observed; with the exception of T1 and T2 that were highly specific to the intended target region. Target oligonucleotides T1 and T2 were first cloned into the pChimera vector, then the entire sgRNA sequence and the U6 promoter were sub-cloned into the p201G vector. The recombinant p201G, carrying the Cas9 encoding gene, sgRNA-T1/T2 oligonucleotide complex and GFP as a selectable marker, was used to transform a strain of *Agrobacterium* carrying the root inducing plasmid (Ri). *Agrobacterium* transformed with the final p201G vector was introduced to the BCN resistant sugar beet line NEMATA.

In total, 1,365 leaf stalks were transformed, out of which 64 stalks produced 184 independent hairy roots. On average, it took 3-4 weeks for the hairy roots to emerge from either end of the excised leaf stalks. I expected that hairy roots integrated with the CRISPR-Cas cassette should display a green fluorescence signal. Moreover, the successful activity of the Cas9 endonuclease would subsequently result in the targeted mutagenesis of *ORF1*.

Roots carried both T-DNA segments originating from the final Cas9 vector, and the Ri plasmid were detected based on the fluorescence signal under a binocular fluorescence microscope (Figure 15). The exposure time for image capture was calibrated to an extent where the negative control displays no signals and the putatively transformed hairy root clones showed a prominent signal. This enabled a reliable differentiation of genuine GFP signals emanating from the transformed hairy root clones while reducing the chance of false positives attributed to autofluorescence. An exposure time of 100 ms was optimum for this. Out of 66 hairy root clones, 35.8% were GFP positive, indicating a possibly stable expression of the Cas9 endonuclease (Table 13). The presence of the CRISPR-Cas cassette was further confirmed by standard PCR using Cas9 gene-specific primers (*cas1_f* and *cas1_r*) and gDNA from hairy root clones as a template.

For mutation screening, PCR using primers (*rbd_10F* and *rbd_10R*) flanking the target regions in the first exon of *ORF1* was carried out. The PCR products of all 66 roots were Sanger sequenced. Four out of the 66 GFP-positive hairy roots showed three different mutant alleles of *ORF1*, designated as *hs4_1*, *hs4_2*, *hs4_3*, and *hs4_4*. Allele *hs4_1* is a 9 bp deletion, thus not leading to a frame-shift mutation of *ORF1*. Two identical alleles,

Functional analysis of the ORF1

designated as *hs4_2* and *hs4_3*, were single base pair insertions, and one allele designated as *hs4_4* carried a 5 bp deletion (Figure 16B).

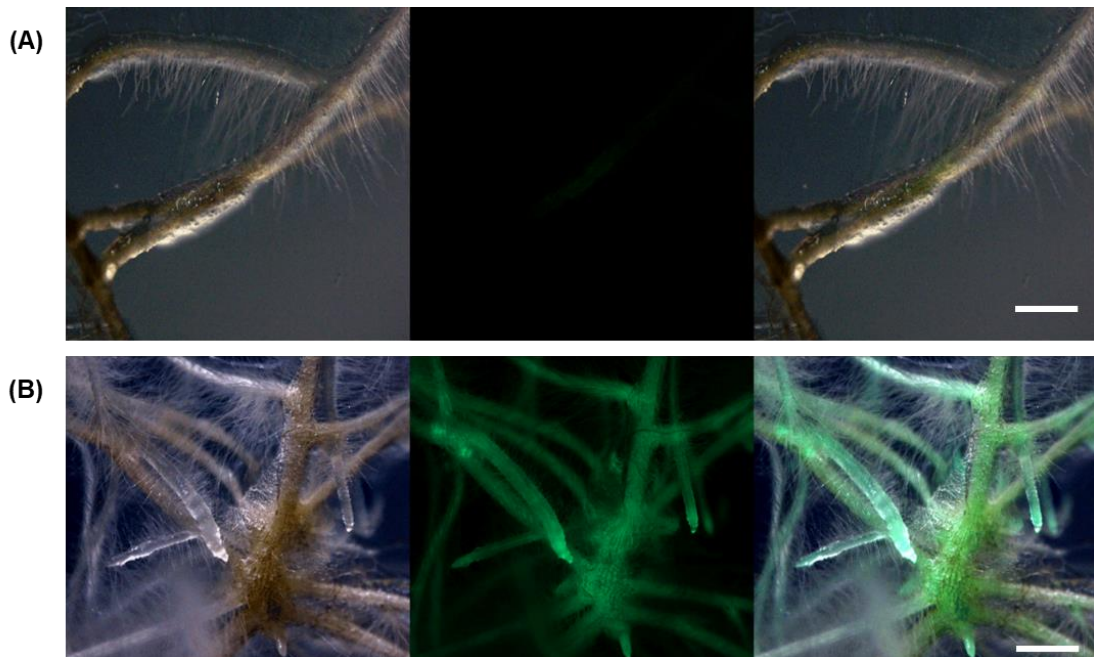


Figure 15: Detection of GFP in hairy root clones under fluorescence microscope. **(A)** Bright field, dark field and merged images of three weeks old hairy roots from control hairy root clone, showing no green fluorescence **(B)** Bright field, dark field and merged images of three weeks old hairy roots from NEMATA expressing the GFP gene from the p201G vector. Scale bar = 1000 µm.

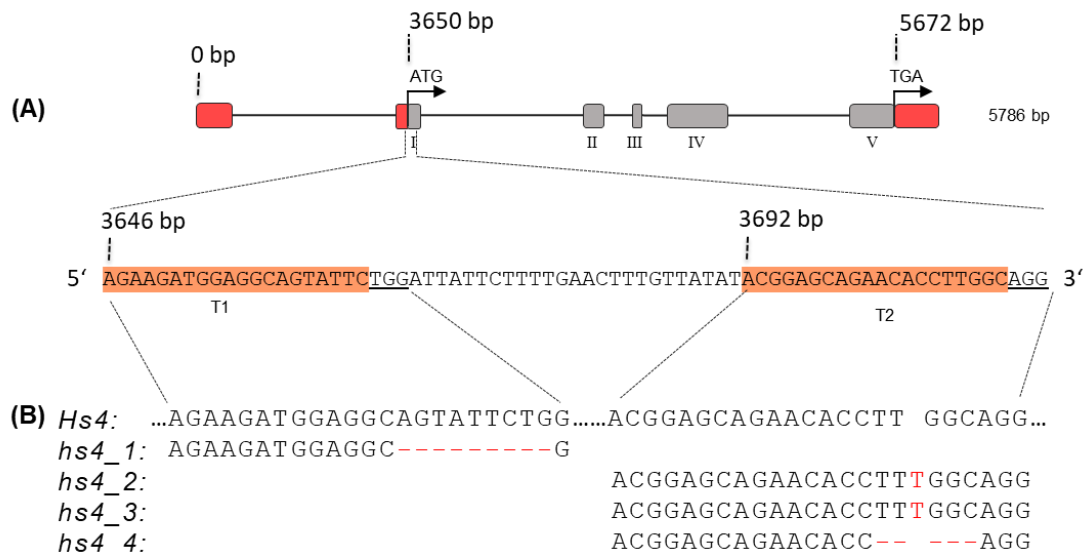


Figure 16: Results from the knockout experiments. **(A)** The sequences and positions of the two target sites in exon 1 of ORF1 are highlighted in orange. **(B)** Four independent CRISPR-Cas induced mutant alleles. *hs4_1* is a 9 bp deletion. *hs4_2* and *hs4_3* are single base pair insertion. *hs4_4* is a 5 bp deletion. Red letter indicates insertions; Red hyphens indicate deletion.

5.2 *ORF1* overexpression mutants

I reasoned that *ORF1* overexpression confers resistance against BCN. I hypothesized that the transcriptional activity varies for different transgenic events. Therefore, I carried out an overexpression experiment to study the effect of *ORF1* expression in hairy root clones derived from a susceptible sugar beet line. I first amplified the coding sequence of *ORF1* from cDNA. The complete coding sequence covering all 633 bp of *ORF1* was amplified by PCR using primers F_OEX and R_OEX. The fragment was cloned into the EcoRI and XbaI restriction sites of the overexpression vector pBin35SRed. The recombinant pBin35SRed vector carrying the *ORF1* coding sequence driven by a CsVMV (Cassava vein mosaic virus) promoter and DsRED (RFP) as a selectable marker was cloned into *Agrobacterium rhizogenes* strain ARqua1. *ORF1* overexpression cassette was introduced into susceptible sugar beet line 093161 using *Agrobacterium*-mediated transformation.

In total, 547 leaf stalks were transformed, out of which 37 hairy roots were obtained. As a result, 11 hairy roots were positive for RFP, indicating a double integration efficiency of ~30% (Table 13). The exposure time for image capture was carefully controlled to reduce the chances of false observations originating as false positives attributed to autofluorescence. An exposure time of 900 ms was found to be the most optimum.

To confirm the overexpression of *ORF1*, RNA was isolated from RFP-positive hairy roots, and RT-qPCR was carried out. The expression level of *ORF1* varied between the different overexpression hairy roots obtained. For example, hairy root clone OEX3 did not show any expression of *ORF1*, while OEX1 showed the highest expression level (Figure 17).

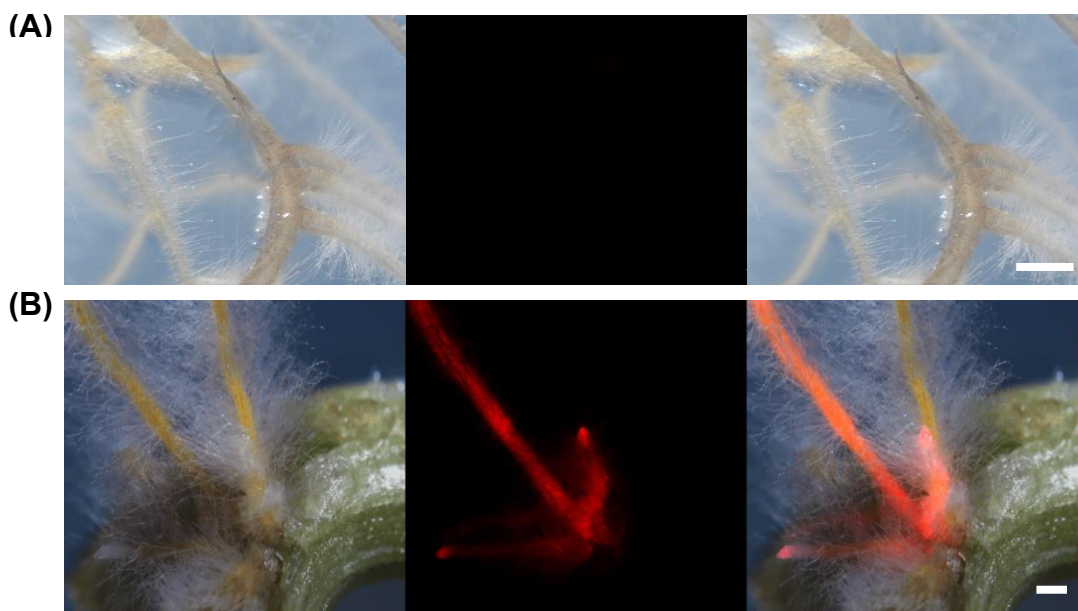


Figure 17: Detection of RFP in hairy root clones under the fluorescence microscope. (A) Bright field, dark field, and merged images of hairy roots from control hairy root clone, showing no red fluorescence (B) Bright field, dark field and merged images of 12 days old hairy roots from the susceptible line 093161 expressing the DsRed gene and carrying the *ORF1*-overexpression cassette from the pBin35SRed vector. Scale bar = 1000 µm.

Table 13: Summary statistics of hairy roots obtained from knock-out and overexpression experiments.

Seed code	Experiment	# Stalks transformed	Total number of hairy roots	# Fluorescing roots	# Edited roots	Roots expressing <i>Hs4</i>
180819	<i>Hs4</i> knock-out	1365	184	66 (35.8%)	4	-
93161	<i>Hs4</i> overexpression	547	37	11 (29.7%)	-	10

5.3 Infection tests with transgenic hairy roots

An experiment was started to measure the infection rates in the different transgenic hairy root clones generated. For infection tests, hairy root clones were transferred to the B5 media without antibiotics and were inoculated with 250 freshly hatched sterilized J2 larvae (Kifle et al., 1999). Inoculated roots were kept in the dark for four weeks; afterward, females were counted under a binocular microscope.

For *ORF1* knock-out hairy root clones, hairy root clones from the resistant line NEMATA carrying wild *ORF1* gene, and hairy root clones from susceptible sugar beet line 093161, lacking *ORF1*, were used as controls. Each clone was represented by 5 sub-clones (repeats). After four weeks of inoculation, females were counted in *hs4_1*, *hs4_2*, *hs4_3*, and *hs4_4* in addition to hairy root clones carrying the *ORF1* wild genotype, and hairy root clones from 093161. As expected, hairy root clones from NEMATA carrying the *ORF1* gene showed a complete resistance against BCN since no J4 females and cysts filled with eggs were observed. As expected, hairy roots carrying the *hs4_1* allele with a 9 bp deletion resulting in an in-frame mutation also showed complete resistance. The number of cysts developing on hairy roots from 093161 ranged from 15 to 22, while the number of cysts developing on hairy roots carrying *hs4_2*, *hs4_3*, and *hs4_4* alleles ranged from 34 to 51. This indicates that knock-out of *ORF1* gene results in loss of resistance against BCN in the otherwise resistant NEMATA (Figure 18)

To test the phenotypic effect of *ORF1* overexpression, infection tests were carried out with ten hairy roots that showed RFP signals and hairy roots from the susceptible 093161 as control. Four weeks after inoculation, I counted J4 females on all the infected hairy roots. As expected, J4 females developed on hairy roots from 093161. Interestingly, a relatively high number of J4 females (17 cysts) developed in OEX3 hairy roots, which showed very low expression levels of *ORF1*. No J4 females were observed on OEX1, OEX6, OEX7, OEX8, OEX9, and OEX10, which were the hairy roots that showed the highest expression levels of *ORF1* (3.7 to 1.46 relative to *BvGAPDH*). Hairy root clones OEX2, OEX4, and OEX5, developed between 1 and 12 J4-females/petri dishes. In these clones, the *ORF1* relative expression ranged from 0.9 to 3.6 (Figure 19).

Taken together, these results demonstrate that putative knock-out of the Rhomboid-like protease could be introduced via CRISPR-Cas mediated targeted mutagenesis. These loss-of-function mutations resulted in a complete abolishment of the resistance mechanism against BCN in *ORF1* gene-edited hairy root clones. In contrast, overexpression of the Rhomboid-like protease in susceptible hairy root clones was successful in inducing a resistance mechanism against BCN.

Functional analysis of the ORF1

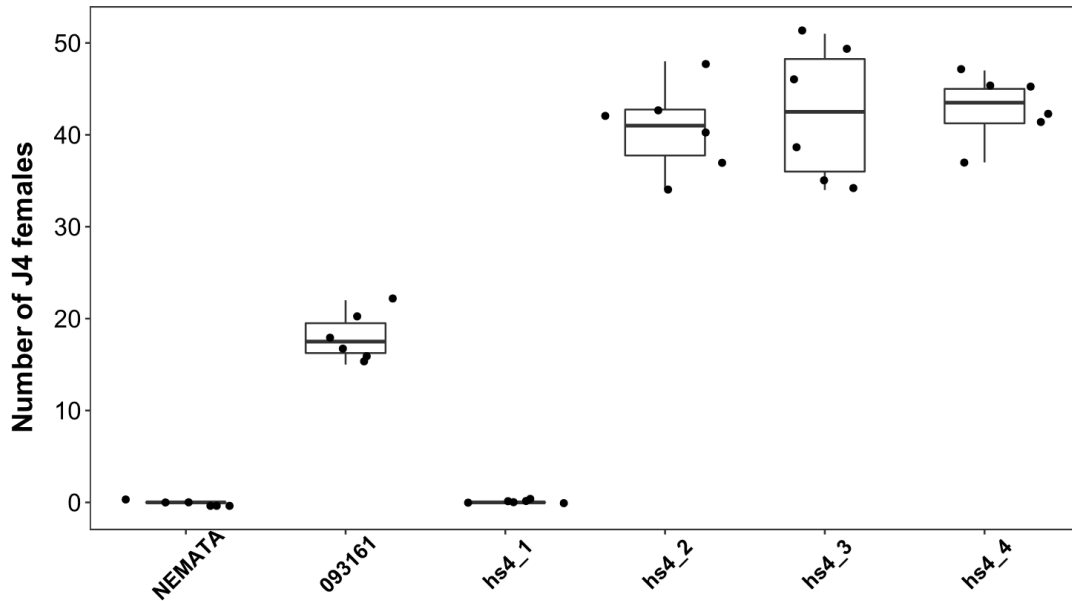


Figure 18: Infection tests of CRISPR-Cas mutated hairy root clones. Box-plot is showing the number of cysts 4-weeks after infection in resistant (NEMATOA), susceptible (093161), and 4 CRISPR-Cas mutated hairy root clones (*hs4_1*, *hs4_2*, *hs4_3*, *hs4_4*). Each box shows an interquartile range; horizontal line inside each box represents the median number of J4 females; vertical lines under the box show quartile 1; vertical lines above the box show quartile 3. Each dot represents an individual number of J4 females in the sample.

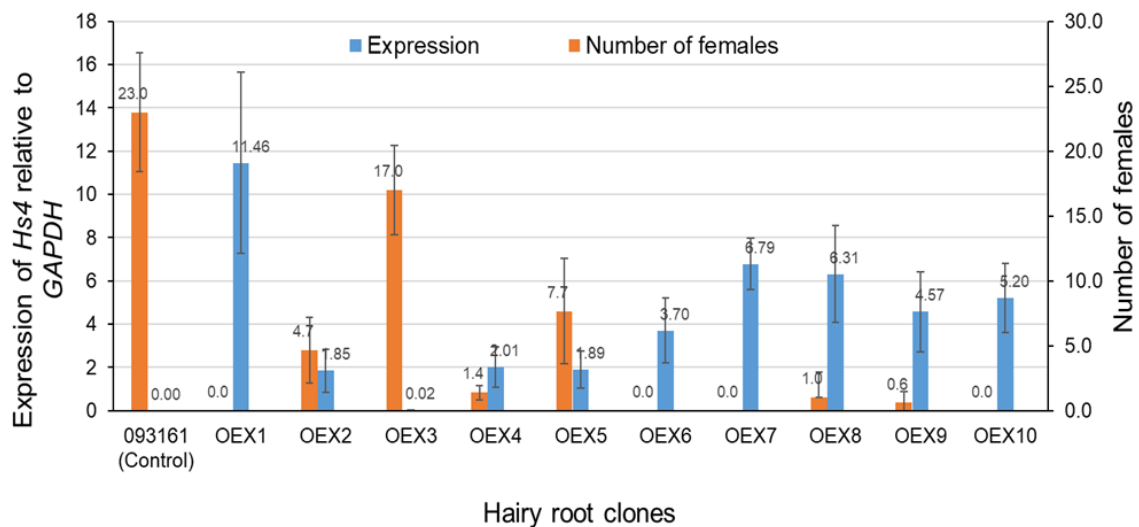


Figure 19: Number of cysts is negatively correlated with expression levels of *ORF1*. A bar-plot showing the expression levels of *ORF1* (blue bars, left y-axis) in ten different overexpression hairy root clones, and average number of cysts of five replicates 4-weeks after infection (yellow bars, right y-axis). Clone OEX3 showed no expression of *ORF1* and showed the highest number of cysts among all RFP-positive hairy root clones. Hairy root clones OEX1, OEX6, OEX7, and OEX10 had no developing J4 females. Error bars for the relative expression represent SEM of three biological replicates, and error bars for infection tests represent SEM of cysts in five replicates.

6 Closing Discussion

Plant-parasitic nematodes are detrimental pests in crop production worldwide. Sugar beet lines carrying a chromosomal translocation segment from the wild beet relative *P. procumbens* are resistant to the beet cyst nematode *Heterodera schachtii*. The focus of this work was to find and clone the beet cyst nematode resistance gene *Hs4*, from the resistant translocation lines TR520 and TR363. The aim was achieved through a series of milestones, which included identifying and characterization of the translocation segment, identifying the translocation breakpoint, identifying the candidate gene, and finally, functional characterization of the candidate gene by overexpression and CRISPR-Cas mediated knock-out.

This study has uncovered the *Hs4* gene, encoding a rhomboid-like protease and conferring complete resistance to beet cyst nematodes. The knock-out of rhomboid-like protease gene in sugar beet hairy roots generated from a resistant translocation line led to the loss of resistance against BCN. In contrast, overexpression of the gene *Hs4* in hairy roots generated from a susceptible sugar beet line conferred complete resistance to BCN. These results confirmed that a rhomboid-like protease is the long-sought *Hs4*. For the first time, a rhomboid-like protease is shown to be involved in plant defense mechanisms. Therefore, the results of this work could open new avenues for further understanding of plant-pathogen interactions.

6.1 A complete sequence of the *P. procumbens* translocation

Translocation is a change in the chromosomal segment's location. Interspecific chromosomal translocation can lead to increase genetic diversity. The importance of translocation in bringing essential genes from wild relatives has been shown in several species including wheat (Song et al., 2013). For instance, a well-known wheat-rye 1B/1R translocation line has been intensively utilized in breeding (Feldman and Levy, 2015; Ren et al., 2017; Reynolds et al., 2021). This chromosomal translocation lead to the transfer of important resistance genes of rye such as *Yr9*, *Pm8*, *Lr26*, and *Sr31* from rye into wheat (Mago et al., 2005; Ren et al., 2009). In addition to resistance, the 1B/1R translocation harbors genes for abiotic stress tolerance, high yield potential and adaptation to different environments (Lelley et al., 2004; McKendry et al., 1996; Villareal et al., 1991). Several hundred wheat cultivars have been released globally, which originated from 1B/1R translocation line (Ren et al., 2017). Similarly, a very useful chromosomal translocation in sugar beet harboring the beet cyst nematode resistance locus has been used by breeders. Earlier mapping studies have shown that the translocation from *P. procumbens* is attached to the end of chromosome 9 of sugar beet (Heller et al., 1996a). Later, multi-color *in situ* hybridizations of extended chromatin fibers had also demonstrated that the translocation is present at the distal end of chromosome 9 of sugar beet (Desel et al., 2001). Previously, the maximum size of the translocation was estimated to be around 1.5 Mb (Schulte et al., 2006) and was later extended to 2.22 Mb (Jäger, 2012). At the beginning of my work, I aimed to generate a complete annotated sequence of the *P. procumbens* translocation segment on chromosome 9 of the resistant translocation line TR520. As the translocation segment originated from *P. procumbens*, I hypothesized that this segment contains repetitive sequences specific to *P. procumbens* and that WGS reads originating from *P. procumbens* will map to this translocation segment, while WGS reads originating from a sugar beet line will not. Following up on this hypothesis, I used WGS data from *P. procumbens* and KWS2320 sugar beet line, *P. procumbens* specific

repetitive sequences (Dechyeva et al., 2003; Schmidt and Heslop-Harrison, 1996), in addition to previously developed translocation specific molecular markers, and BAC and YAC sequences (Schulte et al., 2006). Sequence analysis and assembly resulted in 19 scaffolds. I then carried out a manual curation and arranged the identified scaffolds into two super scaffolds of a final translocation size of 3.2 Mb.

Interestingly, as I carried out a manual curation of the scaffolds, certain regions of the translocation segment had different positions than the previously published translocation maps (Jäger, 2012). According to my analysis, super-scaffold 2 represents the proximal side of the translocation, while super-scaffold 1 represents the distal side of the translocation, which is significantly different from the previously published translocation map. Thus, the translocation regions, including the YAC clones YAC118 and YAC58 which were previously positioned at the distal side of the translocation segment, are now positioned at the proximal end. In contrast, regions including YAC128, YAC104, BAC62, BAC132, BAC57, and BAC100, which were positioned at the proximal side of the translocation segment (Jäger, 2012; Schulte et al., 2006) are now positioned at the distal side of the translocation segment (Figure 7).

In this work, as well as in the previous work from Jäger (2012), a similar hypothesis was proposed: translocation-specific sequences are found in the *P. procumbens* genome but are missing in the *B. vulgaris* genome. However, the approach followed in the previous study was different than this study. In Jäger (2012), the *de-novo* assembly of *P. procumbens*, based on short Illumina reads, was used in blast search to find the translocation-specific sequences. Knowing that the *P. procumbens* genome is highly heterozygous and repetitive, using only short reads resulted in a very poor-quality assembly that did not represent the whole *P. procumbens* genome. Therefore, a blast search against such an assembly did not lead to the identification of all the translocation specific sequences. In my study, I followed a different approach, where I first generated an improved TR520 *de-novo* assembly, then mapped WGS reads from *P. procumbens* and sugar beet (KWS2320) to the newly generated assembly. The strength of my strategy is that WGS reads represent sequences from the complete genome, unlike a subpar *de-novo* assembly such as of the *P. procumbens* genome assembly used by Jäger (2012).

To determine the complete sequences of the translocation segment, I carried out another experiment aiming to identify the proximal and distal ends of the translocation region.

6.2 Identification of the translocation breakpoint

Finding the translocation breakpoint is a critical step to understand how the translocation might have occurred. Therefore, in the second experiment, I aimed to identify the translocation breakpoint. Translocation, the movement of a chromosomal segment to a different chromosome, is one of the major classes of chromosomal rearrangements. Other major classes of chromosomal rearrangements include deletion, duplication, and inversion. DNA breakage plays a significant role in all these events (Griffiths AJF, 1999). The very first step in the generation of any chromosomal rearrangements is a double-strand break. To mitigate the lethal effect of double-strand break, the cell DNA repair mechanism repairs the double-stranded break by joining two broken ends. If by any chance, two different breaks are joined, it leads to chromosomal rearrangements. Another mechanism that can lead to chromosomal rearrangements is crossing over between repetitive DNA segments, termed nonallelic homologous recombination (NAHR) (Mani and Chinnaiyan, 2010; Shinohara and Shinohara, 2013).

In simple terms, the translocation breakpoint is where the chromosomal segment from *P. procumbens* got attached to the chromosome of the sugar beet genome. Therefore, I assumed that the breakpoint would be represented by a set of mate-pair reads, where one read mapped to translocation-specific scaffold and the second read mapped to a sugar beet-specific scaffold. I visually analyzed the mate-pair reads mapping data in IGV and found that in one of the mate-pair reads, one read mapped to superscaffold2 of the translocation region while the other read mapped to contig C5128660, which belongs to the sugar beet genome. Thereby, I concluded that the breakpoint lies between the end of superscaffold2 and the beginning of the contig C5128660. As expected, the contig C5128660 matched with the Bvchr9.unsca002 scaffold of sugar beet. Based on marker map, Bvchr9.unsca002 is placed on chromosome 9 of sugar beet (Dohm et al., 2014a). As I was inspecting the read mapping data of TR520 onto the sugar beet reference genome, I noticed a sudden drop in coverage from ~150 to ~70 reads per base. Drop in coverage of TR520 reads to half of the original coverage confirmed the translocation breakpoint at this position on chromosome 9 and further indicated that the translocation in TR520 is present in a hemizygous state.

To find the exact position of the translocation breakpoint, I performed a PCR followed by Sanger sequencing. Therefore I could identify the precise position of chromosome 9 at which the translocation took place. While I could locate and characterize the proximal end of the translocation region to the precision of a single base pair, it became cumbersome to resolve the genome assemblies as I progressed towards the telomeric regions of the chromosome. Therefore, the distal end of the translocation remains elusive, but it is of no great concern due to the presence of significantly fewer or no genes closer to the telomeres. Based on my analysis, the tentative length of the translocation was found to be ~3.2 Mb.

Contrary to my hypothesis, Jäger (2012) assumed that one of the assembled scaffolds harboured the translocation breakpoint. In Jäger (2012), blast search of translocation scaffolds to sugar beet and *P. procumbens* genome, with the expectation that one part of a scaffold aligns to a sugar beet specific sequence while the other part aligns to a *P. procumbens* specific sequence, did not assure the identification of translocation breakpoint. Theoretically, such a strategy should have resulted in the identification of the translocation breakpoint. However, in practice, not every genomic region is assembled into a scaffold. Therefore, relying on such a strategy, Jäger (2012) could not identify the position of the breakpoint.

Sequence analysis of the translocation segment indicated that the translocation sequence corresponded to a complete sequence from *P. procumbens* that was not interrupted by any sugar beet sequences. In addition, only one set of the mate-pair libraries that mapped to the end of super-scaffold 2 had pairs mapping to sugar beet-specific scaffolds. Moreover, none of the assembled translocation-specific scaffolds showed homology to sugar beet scaffolds, indicating only one translocation breakpoint that lies between super-scaffold 2 of the translocation and contig C5128660 of sugar beet. As the translocation segment is not interrupted by any sugar beet genomic sequences, and as I could find only one translocation breakpoint, I assume that the translocation occurred due to one chromosomal recombination event. Looking at the read mapping from *P. procumbens* outside the translocation region, I could infer that the two genomes are very distinct. Therefore, this translocation could not have occurred due to homologous recombination between chromosome 1 of *P. procumbens* and chromosome 9 of sugar beet.

Consequently, the recombination event that led to this translocation was most likely due to non-allelic homologous recombination between chromosome 9 of sugar beet and chromosome 1 of *P. procumbens*. This non-allelic recombination event could be due to microhomology between chromosomes 1 and 9 at the breakpoint site. Although I could fine map the breakpoint on chromosome 9 of sugar beet, sequence information about chromosome 1 from *P. procumbens* is still unknown. To fully understand how the translocation might have occurred, the *de-novo* sequence assembly of the *P. procumbens* genome must be generated and compared to the sugar beet genome to find micro-homologous regions between the two genomes.

6.3 Identification of candidate genes

My quest for the candidate gene began from a simple expectation that the gene conferring complete resistance against BCN is present in the resistant translocation lines but absent in the susceptible ones. Using whole-genome mapping coverage analysis, I identified the translocation regions present in the resistant lines TR520 and TR363 but absent in the susceptible line TR659. I found three regions, encompassing a total of 229 kb, and referred to them as “critical regions”. The next step was to identify the genes present in these critical regions. After structural and functional annotation of the translocation segment, I could identify 33 genes in the critical regions, out of which 19 had transcript evidences. Interestingly, none of the genes found in the critical regions were annotated as resistance gene analogues.

Before proposing my candidate gene, I first checked whether the previously proposed candidate genes were present in the critical regions (Capistrano, 2010; Jäger, 2012; Qiao, 2014; Tian, 2003). Tian (2003) proposed two genes, *cZr-3* and *cZR-7*. Jäger (2012) analysed the sequences of the proposed genes, *cZr-3* and *cZR-7*, and found that they originated from the sugar beet genome. As a proof of concept, I again analysed these genes and found that they match with 100% identity to regions on the sugar beet reference genome RefBeet1.2.2. Therefore, I excluded them from any further analysis. Capistrano (2010) proposed another gene, *ORF702*, that encodes a β -1,3-galactosyltransferase. Jäger (2012) showed that *ORF702* is present in the resistant translocation line TR520, while absent in the susceptible line TR659. However, further functional analysis showed that the hairy roots overexpressing *ORF702* were not resistant against BCN. According to my analysis, *ORF702* is indeed present in one of the critical regions but based on the results of the functional analysis by Jäger (2012), I excluded this gene from any further analysis. Jäger (2012) proposed three more genes, *ORF801*, *ORF802*, and *ORF803*, and found that *ORF801* and *ORF802* were expressed in the susceptible line TR659. Hence she excluded them as candidates for *Hs4*. Jäger (2012) performed an expression analysis of *ORF803* and reported that it was only expressed in resistant lines and *P. procumbens*. Fen Qiao (unpublished data) conducted RNAi-mediated silencing of *ORF803* and demonstrated that hairy roots with silenced *ORF803* remained resistant against *H. schachtii*. In my analysis, I found that *ORF803* belongs to the translocation segment but is not present in any critical regions. I, therefore, excluded this gene from the current study. Fen Qiao (unpublished data) had previously proposed three other candidate genes - *ORF901*, *ORF902*, and *ORF906*, which according to my analysis, are present in the critical regions, except *ORF906*. *ORF901* encodes amino acid transporter, *ORF902* encodes calcineurin B-like protein, and *ORF906* encodes histone deacetylase (Fen Qiao, Unpublished data). *ORF901* and *ORF906* RNAi-mediated gene silencing in hairy roots was conducted, however, due to poor regeneration of hairy roots, the tests were terminated (Birgit Defant, personal communication). Nevertheless, all these

three genes are coding for proteins involved in regulatory mechanism (Cho et al., 2016; Ding et al., 2012; Yang et al., 2014).

With all these previous failed attempts to find the nematode resistance gene, it was compelling to reassess the data and identify the gene that has not been studied so far and shows direct relation with nematode resistance. As Wyss and Grundler (1992) have explained, during the second stage of their lifecycle (J2 larvae), beet cyst nematodes use their stylet to penetrate the epidermis of the plant cell. This invasive thrusting of the stylet causes severe mechanical damage in the roots of a sugar beet plant. Hence, the initial response used against BCN could be comparable to the plants' defense mechanisms against pathogens, such as insects that cause mechanical damage in plant tissues. Maize lines resistant against lepidopteran species carry a unique 33-kDa cysteine protease (Pechan et al., 2000). Upon damage by Lepidopterans, the protease is mobilized to the mid-whorl, the preferred caterpillar-feeding site, and disrupts caterpillar nutrient utilization (Pechan et al., 2000). The red currant tomatoes (*Solanum pimpinellifolium*) harbouring the extracellular plant immune receptor protein Cf-2 are resistant to the leaf mold fungus *Cladosporium fulvum* and the plant-parasitic nematode *Globodera rostochiensis*. *S. pimpinellifolium* plants carry the apoplastic papain-like cysteine proteases Rcr3. The venom allergen-like effector Gr-VAP1, secreted by *G. rostochiensis* upon infection, binds to Rcr3. The binding of Gr-VAP1 to Rcr3 perturbs the active site of Rcr3, which is detected by the Cf-2 immune receptor leading to an immune response (Lozano-Torres et al. 2012). This is an example where proteases are involved in defence mechanisms, specifically against nematodes.

In addition to the role of proteases in the defence mechanisms against insects, another fascinating study showed that a serine protease, Sep1, secreted by *Bacillus firmus* DS-1 has a nematicidal activity (Geng et al., 2016). The study demonstrated that Sep1 degrades multiple intestinal and cuticle-associated proteins in the nematodes. In addition, in the nematode parasitic fungus *Pochonia chlamydosporia*, rhomboid proteases are expressed during the endophytic phase (Larriba et al., 2014).

Considering their role in plant defence mechanisms and their possible nematicidal activity, I decided to focus my analysis on proteases. Hence, my first candidate gene was *ORF1* which encodes a Rhomboid-like protease.

In-silico analysis predicted that the endoplasmic reticulum membrane is the most probable location of the ORF1 protein on a cellular level. This finding is supported by other studies where rhomboid proteases are shown to perform intramembrane proteolysis by localizing themselves on membranes of different cellular organelles (Kühnle et al., 2019). It has been shown that during their third feeding stage, nematode gland secretions are injected into the syncytium, where they form a feeding tube (Wyss and Grundler, 1992). One end of the feeding tube is attached to the cytoplasm of the host, while the other end is attached to the nematode stylet (Sobczak et al., 1997). In the cytoplasm, the outside surface of the feeding tube core has been connected to the endoplasmic reticulum's cisternae (Berg and Taylor, 2008). It has also been observed that nematodes could ingest molecules with a maximum Stokes radius of 3.2 - 4 nm (Böckenhoff and Grundler, 2009). ORF1 protease comprises 210 aa with an approximate Stokes radius of 2.5 nm, which is well below the threshold of Stokes radius ingestible by the nematodes.

In plants such as Arabidopsis and rice, there are around 20 rhomboid-domain-containing genes (Adamiec et al., 2017; Tripathi and Sowdhamini, 2006). In Arabidopsis, certain rhomboid-like proteins (RBL8-10) are located in the chloroplast envelope and were found to be involved in floral development (Knopf et al., 2012; Thompson et al., 2012). Other

RBL proteins are involved in regulating several cellular functions via regulated intramembrane proteolysis (Urban and Wolfe, 2005; Urban and Dickey, 2011). For instance, it has been shown that *RHOMBOID-LIKE 10 (RBL10)* mutants exhibited decreased fertility. *RBL10* was also suggested to take part in response to cold and photoprotective mechanisms (Thompson et al., 2012). Another *RBL* gene, *RBL14* was predicted to be involved in heat responses (Lee et al., 2014). However, the protein sequence comparison of ORF1 to other rhomboid-like proteases showed that ORF1 shares only 60% similarity with the closest homologue. Different rhomboid proteases can have very different functions based on the substrate they interact with, which is decided by the structure of the protease. As we know that ORF1 is only 60% similar to the rhomboid-protease from sugar beet, it is highly likely that both rhomboid-proteases interact with different substrates and hence attained different functions.

Considering the unique sequence of ORF1, its cellular localization, and the possibility that it has a nematicidal activity, in addition to the role of proteases in defence mechanisms, I, therefore, considered *ORF1* as the best candidate gene for further functional studies.

6.4 Functional analysis of the candidate gene

In this experiment, the main question was if *ORF1*, which encodes a rhomboid-like protease confers resistance against the beet cyst nematode *H. schachtii*. I hypothesized that hairy roots generated from a resistant translocation line carrying a non-functional *ORF1* are susceptible to BCN. In contrast, hairy roots generated from a susceptible sugar beet line overexpressing *ORF1* are resistant to BCN.

Unfortunately, transformation by *A. tumefaciens* to generate whole sugar beet plants was not feasible. This is because most sugar beet genotypes are recalcitrant to *Agrobacterium* transformation. Additionally, the available sugar beet transformation protocols are hard to replicate or are solely available as proprietary technologies offered by private companies and thus rather expensive (Jäger, 2012). Therefore, I used hairy roots for the functional analysis of *ORF1*.

I chose the commercial BCN resistant variety NEMATA to induce mutations because it was comparatively faster in generating hairy roots than translocation lines TR520 and TR363. Using two target sites in the first exon of *ORF1*, I successfully induced mutations within both target sites. Four CRISPR-Cas targeted mutagenized hairy root clones were obtained, with a mutation efficiency of 6.25%. Infection tests showed that all the hairy roots carrying frameshift mutations were susceptible to BCN, suggesting they lost the resistance against BCN. In contrast, hairy root clones with a 9 bp deletion, thus conferring no change in the reading frame of the *ORF1*, remained resistant. No change in phenotype after 9 bp deletion was another indication that *ORF1* confers resistance against BCN. I then overexpressed *ORF1* in the hairy roots of the susceptible 093161 sugar beet line. Infection tests showed that the number of developed cysts was negatively correlated with the expression level of *ORF1*. No cysts developed in the hairy root clones that showed the highest expression levels of *ORF1*. Notably, a hairy root clone with a very low expression level of *ORF1* had a similar number of cysts as the susceptible control hairy roots. The two experiments supported my hypothesis and indicated that *ORF1* does confer resistance against BCN.

The translocation lines TR520 and TR363 are completely resistant against BCN. However, none of the candidate genes proposed in previous studies provided complete resistance, not even *HsI^{pro-1}* (Cai et al., 1997). Therefore, based on this study, *ORF1*

knockout resulted in complete loss of resistance while overexpression caused resistance. Collectively, results from this study confirmed our hypothesis that ORF1 is the long-sought *Hs4* gene conveying complete resistance to BCN. Further functional analyses are needed to unravel the defence mechanism mediated by *ORF1*. *ORF1*-reporter gene fusion studies are necessary to confirm the cellular localization of ORF1 protein. In addition, overexpression of *ORF1* in other species that are also known to be susceptible to BCN, such as Arabidopsis, is essential to verify whether ORF1 will preserve its functional activity against BCN in a different species.

6.5 Possible mechanism of rhomboid-like protease-mediated resistance

Considering the possible nematicidal activity of serine proteases, one possibility could be that Hs4 is being ingested by the nematodes and causing toxicity. The fact that the feeding tube of nematodes is attached to the endoplasmic reticulum of the host plant cell, which is the predicted cellular location of Hs4, makes this hypothesis plausible. To study such a hypothesis, we first must study whether the J2 larvae are ingesting Hs4 during infection. Fluorescence microscopy studies of nematodes feeding on hairy roots overexpressing GFP-tagged Hs4 protein fusion proteins should be conducted. Interestingly, a GFP-Hs4 fusion protein Stokes radius is 3.19 nm, which is still below the maximum Stokes radius ingestible by the nematodes (Böckenhoff and Grundler, 1994). In addition, to study whether Hs4 leads to any damage in the digestive tract of nematodes, nematode toxicity bioassays should be carried out. The possibility of such a mechanism can also be studied in the model animal *C. elegans*. For instance, Geng et al. (2016) studied the nematicidal activity of the serine protease Sep1 through feeding nematodes *Sep1* expressing *E. coli*. Studies can also be conducted to find whether Hs4 retains its pathological activity in related species, such as in Arabidopsis. I also suggest testing the activity of the *Hs4* gene against other plant pathogenic nematodes.

It is known that BCN produces a feeding tube in front of its stylet through which it takes up nutrients and secretes effectors. The feeding tube is attached to the cisternae of the ER (Sobczak and Golinowski, 2011). The ER plays a central role in maintaining cell homeostasis. It facilitates intra- and intercellular communication (Hawes et al., 2015). One study showed that effectors secreted by a pathogen get attached to the ER membrane. These effectors interact with certain transcription factors and prevent their translocation to the host nuclei (McLellan et al., 2013). Another study has shown that effectors can manipulate components of the UPR (unfolded protein response), thus ER homeostasis. For example, it has been shown that *Phytophthora sojae* effector PsAvh262 directly interacts with ER luminal binding immunoglobulin proteins, increasing their stability and preventing ER stress-induced PCD (Fan et al., 2018). This is an example of how effectors attached to ER membrane can compromise plant's defence by disturbing its ER homeostasis. Although we do not know whether BCN-secreted effectors are attached to the ER of sugar beet plants, the localization of Hs4 in the ER membrane, in addition to the attachment of the feeding tube to the ER, suggests that a similar mechanism is being carried out in sugar beet. Therefore, Hs4 might be cleaving such membrane-bound effectors in resistant sugar beet lines, thus preventing the disturbance of ER homeostasis. In addition, syncytia from resistant plants showed ER membrane aggregates (Holtmann et al., 2000), which aligns with the idea that ER stress is induced as a resistance mechanism leading to plant cell death.

Post-translational modifications of proteins can determine protein localization, interactions, and stability. Such modifications are crucial in plant immunity (Withers and Dong, 2017). It has also been shown that post-translational modifications play a vital role

in the host-pathogen interface during ETI by activating PRRs (Roux et al., 2011). My expression analysis experiment showed that *Hs4* is constitutively expressed in roots, which indicates that *Hs4* activity might also be regulated post-transcriptionally. Upon infection, specific effectors secreted by the nematodes might directly or indirectly interact with *Hs4* protein. As *Hs4* protein becomes active after infection, it can cleave the effectors, thus preventing them from stabilizing the ER. As a result, ER-stress induced PCD can take place leading to the collapse of the syncytia, therefore, the death of nematodes. Due to the possibility of such regulation of *Hs4* activity, it would also be interesting to understand the role of different motifs and active sites of *Hs4*.

As reported earlier, BCN *H. schachtii* effector *Hs4E02* targets *A. thaliana* vacuolar papain-like cysteine protease (Pogorelko et al., 2019b). Such an interaction might also be possible between *Hs4* and BCN effector, leading to defense response. To find if the *Hs4* protein directly interacts with any of the effectors secreted by BCN upon infection. I suggest a yeast-two-hybrid screening of *Hs4* against a cDNA library of effectors from BCN. In addition, I suggest performing an electron microscopic study at the infection site of the resistant hairy roots overexpressing *Hs4* in the early stage of syncytium development to detect whether similar membrane aggregates are present as observed for resistant sugar beet lines in an earlier study (Holtmann et al., 2000). Finally, to unravel whether *Hs4* interacts with any plant cell proteins, I suggest yeast two-hybrid screens, co-immunoprecipitation, and pull-down assay of *Hs4* protein before and after infection.

This study has shown that the nematode resistance gene *Hs4* can provide complete resistance against beet cyst nematode *H. schachtii*. Thus, *Hs4* can become an essential source for plant breeders breeding nematode-resistant varieties. However, *P. procumbens* is genetically distant from sugar beet, and recombination between the translocation segment and other parts of the sugar beet genome has never been observed. Therefore, *Hs4* cannot be introgressed into susceptible sugar beet lines by conventional breeding methods. Instead, *Hs4* can be introduced to susceptible varieties by transgenic approaches. The transgenic varieties can be grown in countries like the USA and Canada, which are major sugar beet-producing countries, and where a vast majority of cultivated crops, such as soybean, cotton, and corn, are genetically engineered. In a similar fashion, if *Hs4* can provide resistance against other plant pathogenic nematodes, it can also be used to protect other plant species against nematodes.

7 Summary

Plant parasitic nematodes are important pests in crop production worldwide. Therefore, breeding nematode resistant crops is an important aim nowadays. The beet cyst nematode *Heterodera schachtii* causes severe losses in sugar beet production. The only source of resistance are sugar beet wild relatives from the genus *Patellifolia*. Sugar beet lines carrying a translocation stemming from chromosome 1 of *P. procumbens* are resistant to beet cyst nematode. The aim of this study was to clone the beet cyst nematode resistance gene *Hs4*.

In this study, I characterized the translocation segment in the resistant translocation line TR520. I then identified the translocation breakpoint to the precision of a single base pair. Finally, I identified a putative resistance gene. I functionally characterized my candidate gene by two approaches, CRISPR-Cas mediated knock-out and overexpression studies.

To characterize the translocation segment, I first generated a *de-novo* assembly of the resistant translocation line TR520. Then, using whole genome sequencing reads from *P. procumbens*, I identified the translocation specific scaffolds. Finally, two super scaffolds for the whole translocation segment, super scaffold 1 and super scaffold 2 were generated which are 1.109 and 2.121 Mb in size, respectively. I identified the translocation breakpoint in the following way. I inspected mate-pair read libraries for TR520 and found a set of reads joining super scaffold 2 with a scaffold from chromosome 9 of the sugar beet genome. Therefore, I concluded that the breakpoint lies between these two scaffolds. I amplified the breakpoint by PCR, followed by Sanger sequencing, and I could identify the exact nucleotide position at which the sequence of chromosome 9 from sugar beet ended and the translocation sequence started.

To find the resistance gene, I used whole genome sequencing data from two resistant translocation lines, TR520 and TR363, and one γ -irradiated susceptible translocation line TR659. I then identified the translocation regions that are present in the resistant lines but absent in the susceptible line. I found three regions, encompassing a total of 229 kb, and referred to them as “critical regions”. The next step was to identify the putative resistance gene. Using transcriptome data from the roots of the resistant plant after infection, I could narrow down the search to 19 gene-models within the critical regions. None of the genes in the critical regions was annotated as resistance gene analogue. Considering the role of proteases in plant-defence mechanisms, I chose a rhomboid-like protease encoding gene as a putative *Hs4*, which is predicted to be bound to the endoplasmic reticulum.

To functionally characterize my candidate gene, I performed CRISPR-Cas mediated knockout in hairy roots of a resistant translocation variety and overexpression study in hairy roots of a susceptible sugar beet line. Knock-out of my candidate gene resulted in complete loss of resistance, while overexpression led to complete resistance. This study has revealed the *Hs4* gene, encoding a rhomboid-like protease and conferring complete resistance to beet cyst nematodes. This gene is the first protease which alone causes resistance to a pest. Thus, it constitutes a previously unknown mechanism of plants to fight plant parasitic nematodes.

8 Zusammenfassung

Pflanzenparasitäre Nematoden sind wichtige Schädlinge in der weltweiten Pflanzenproduktion. Daher ist die Züchtung von nematodenresistenten Kultursorten heutzutage ein wichtiges Ziel. Der Rübenzystennematode *Heterodera schachtii* verursacht schwere Verluste in der Zuckerrübenproduktion. Die einzige Quelle für Resistenz sind wilde Verwandte der Zuckerrübe aus der Gattung *Patellifolia*. Zuckerrübenlinien, die eine Translokation tragen, die aus dem Chromosom 1 von *P. procumbens* stammt, sind resistent gegen Rübenzystennematoden. Ziel dieser Studie war es, das Rübenzysten-Nematodenresistenzgen *Hs4* zu identifizieren.

In meiner Arbeit habe ich das Translokationssegment in der resistenten Translokationslinie TR520 charakterisiert. Ich habe dann den Translokations-Bruchpunkt bis auf ein einzelnes Basenpaar genau bestimmt. Schließlich identifizierte ich das vermeintliche Resistenzgen. Ich charakterisierte mein Kandidatengen funktional durch zwei Ansätze, CRISPR-Cas vermittelte knock-out- und Überexpressionsstudien.

Um das Translokationssegment zu charakterisieren, habe ich zunächst eine *de-novo*-Sequenzierung der widerstandsfähigen Translokationslinie TR520 vorgenommen. Danach identifizierte ich die Translokations-spezifischen scaffolds mit Hilfe einer *P. procumbens* Genom-Sequenz. Schließlich wurden zwei super-scaffolds für die gesamte Translokation erzeugt, super-scaffold 1 und super-scaffold 2 mit 1109 bzw. 2121 Mbp Länge. Ich habe den Translokations-Bruchpunkt wie folgt identifiziert. Ich untersuchte mate-pair-read Bibliotheken von TR520 und fand eine Reihe von reads, die super-scaffold 2 mit einem scaffold aus Chromosom 9 der Zuckerrübe verbinden. So kam ich zu dem Schluss, dass der Bruchpunkt zwischen diesen beiden scaffolds liegt. Ich analysierte den Bruchpunkt im Detail, in dem ich die DNA mittel PCR amplifizierte und das PCR Produkt mittels Sanger-Sequenzierung sequenzierte. So konnte ich die genaue Nukleotid-Position identifizieren, an der die Sequenz des Chromosoms 9 aus Zuckerrüben endet und die Translokationssequenz beginnt.

Um das Resistenzgen zu finden, habe ich die Genomsequenzen von zwei resistenten Translokationslinien, TR520 und TR363, und einer γ -bestrahlten anfälligen Translokationslinie TR659 verwendet. Ich habe dann die Translokationsbereiche identifiziert, die in den resistenten Linien vorhanden sind, aber in der anfälligen Linie fehlen. Ich habe drei Regionen gefunden, die insgesamt 229 kb umfassen, und sie als „kritische Regionen“ bezeichnet. Der nächste Schritt war, das vermeintliche Resistenzgen zu identifizieren. Anhand von Transkriptomdaten aus infizierten Wurzeln einer resistenten Pflanze konnte ich die Suche auf 19 Genmodelle innerhalb der kritischen Regionen eingrenzen. Keines der Gene in den kritischen Regionen wurde als Resistenzgen-Analog geführt. In Anbetracht der Rolle von Proteasen bei pflanzlichen Abwehrmechanismen wählte ich ein Gen, welches für eine Rhomboid-ähnliche Protease kodiert, als vermeintliches *Hs4* Gen aus. Diese Protease verfügt über eine Bindedomäne für das endoplasmatische Retikulum.

Um mein Kandidatengen funktionell zu charakterisieren, schaltete ich das Kandidatengen mittels CRISPR-Cas vermittelter Mutagenese in hairy roots einer resistenten Translokationslinie aus. Danach wurde das Gen in hairy roots einer anfälligen Zuckerrübe überexprimiert. Der knock-out meines Kandidaten-Gens führte zu einem vollständigen Verlust der Resistenz, während die Überexpression zu vollständiger Resistenz führte. Diese Arbeit hat das Ziel vollständig erreicht und das *Hs4*-Gen identifiziert, welches für

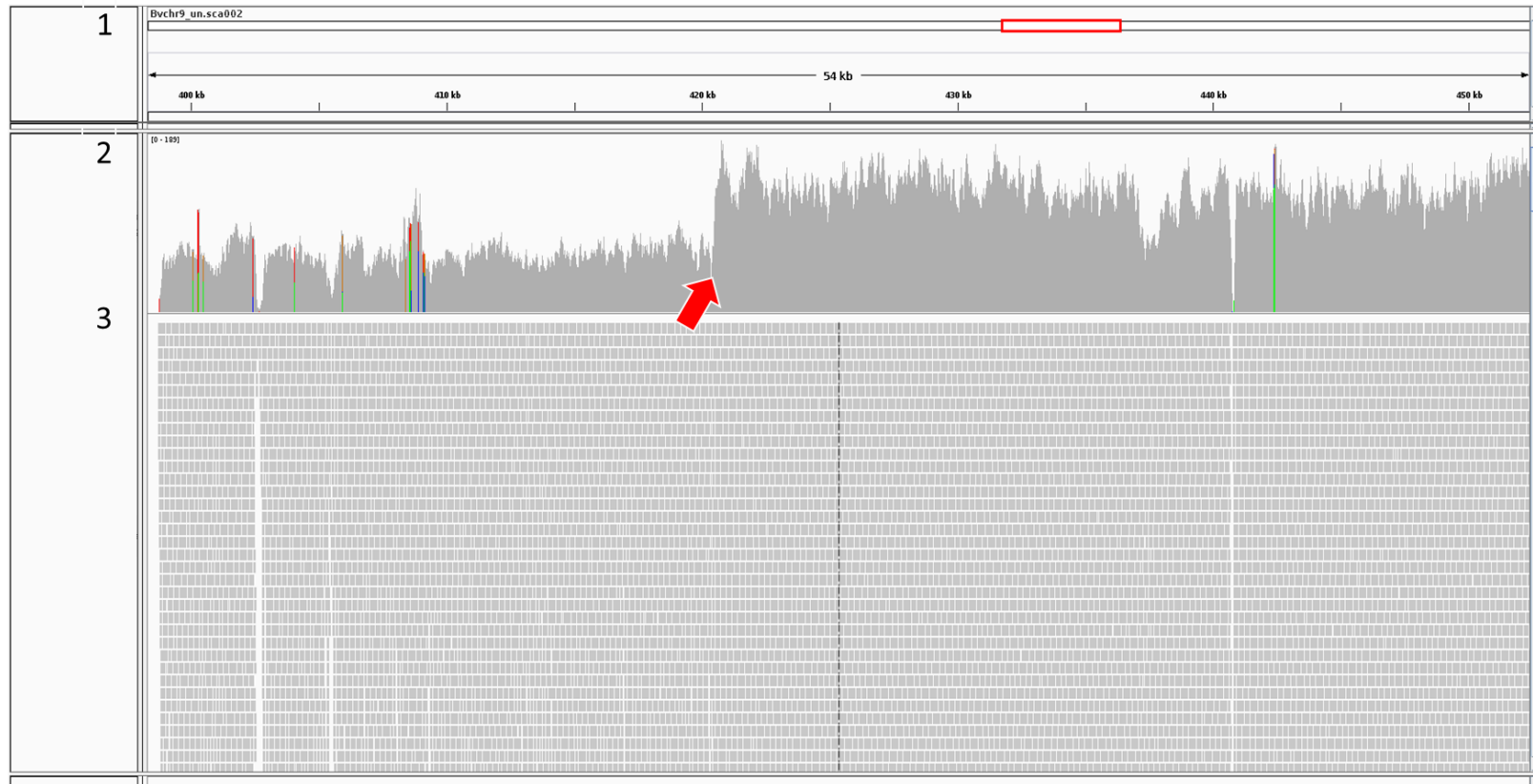
Zusammenfassung

eine Rhomboid-Protease kodiert und eine vollständige Resistenz gegen Rübenzystennematoden verleiht. Dieses Gen ist die erste Protease, die alleine Resistenz gegen einen Schädling verursacht. Somit stellt es einen bisher unbekanntem Mechanismus von Pflanzen dar, um parasitäre Nematoden zu bekämpfen.

9 Appendix

Supplementary Table 1: Primer sequences used in this study.

Gene, map position	Primer name	Primer sequence	Orientation	Annealing temperature (°C)
Translocation breakpoint	H203	AAGAACTGACCCGACCTGTT	forward	59
	H208	TACATACGGTGGGGTTTGCT	reverse	59
At-U6-sgRNA	rbd_gRNA_F	ATATACTCTCTTAAGGTAGCCTTTTTTCTTCTTTCGTTCA	forward	64
	rbd_gRNA_R	TGCTACTCTCTTAAGGTAGCCTAGAAAAAAGCACCGACTC	reverse	64
CRIPSR-Cas genotyping	rbd_10F	ACGCAATGTAGCATTCTTAAAAC	forward	60
	rbd_10R	GAGGCATGATAACCTGACCCT	reverse	60
<i>Hs4</i>	hs4_F	ATGGAGGCAGTATTCTGGAT	forward	59
	hs4_R	GCAGAGGTCACAAATTGGTACC	reverse	59
<i>Hs4</i> coding sequence	F_OEX	TAGCTGAATTCTTAAGAAGAAGATGGAGGCAGTA	forward	62
	R_OEX	TAGTTCTAGATCAGGGTGGCAAACAATTATTCA	reverse	62
<i>HsI^{pro-1}</i>	1832_f	CGAATAAGTGAGAGGATC	forward	58
	1832_r	GGCACCATCCAAACTCGG	reverse	58
BvGAPDH	GAPDH_F	GCTTTGAACGACCACTTCGC	forward	60
	GAPDH_R	ACGCCGAGAGCAACTTGAAC	reverse	60



Supplementary Figure 1: Visualization of the breakpoint on IGV. Screenshot from Integrative Genomics Viewer (IGV) showing read alignment results in a 54 kb region of Bvchr9_un.sca002:4,00,000-4,50,000 (track 1, top) around the translocation breakpoint (red arrow) on the sugar beet reference genome (RefBeet1.2). Track 2 is the coverage track. Track 3 is an alignment track showing individual paired-end reads mapped to this region of TR520.

Appendix

```

Hg4/1-210      20 LGKVEE IRTLYL IHDR PAWYQVTSFA CHYNNWHL CNLFFLY IFGKLV EEEVGG FYLWYY I LTA VSSNLVSWL LPRSGS SAGASGAVFGLFA ISFSVKL LLRDRCPDN KE DWRRF I EV I LGHFV LQRHMEAL HGR NAHVNANV I DPAL VPWNHIAHL AGAVVGMSLV I I PHD IRRLL SVNNCLPP - 210
MER060184/1-30: 122 LFLVDRDKFLYLHNO PAWYQVTA AF CHANWNHLSNLFFLY IFGKLV EEEGSA LWLSY I LTGAGANL VSWL VLP RNAVS VSGASGAVFGLFA ISVLVKM ----- SWDWRK ILEVL ILGQFV IEKVMEAAQASTGLSGF F --IGTNSLQSVNHI AHL SGAL VGVLVWLL SKVPSDASDEKGLPE - 301
MER0412756/1-28: 102 VFQVRAVRLALYLYHNR PAWYQV TATFCHP NWNHLSNLFFLY IFGKLV EEEGNFGLWLSY I LTGAGANL VSWL I LPRNAVS VSGASGAVFGLFA ISVLVKM ----- SWDWRK ILEVL ILGQFV IEKVMEAAQASTGLAEG I --HGG SALQSVNHI AHL SGAL IGVAL IHWL SGISSEPDQNNK - -- 279
MER0639291/1-29: 112 WLQ I R G I K A L Y L Y H N R P A W Y Q V T S F C H A N W N H L S S N L F F L Y I F G K L V E E E G N F A L W S Y I L T G A G A N L V S W L V L P R N A V S I G A S G A V F G L F A I S V L V K M I ----- SWDWRK ILEVL ILGQFV IEKVMEAAQASTGMSGS F --S --GGYMQSVNHI AHL SGAL VGVLVWLL SRVPSQFPDQKPHEN - 297
MER0659701/1-23: 117 L F Q V R I K S L Y L Y H N R P A W Y Q V T A T F C H A N W N H L S S N L F F L Y I F G K L V E E E G N F A L W S Y I L T G A G A N L V S W L L P R N A V S V G A S G A V F G L F A I S V L V K M ----- SWDWRK ILEVL ILGQFV I K R V M E A A Q A S A G P S A S L --QGNVALQSVNHI AHL SGAL IGVLVWLL SRVPSRSPNQKSTLR - 297
MER1084205/1-23: 45 WFQ I P D V K S L Y L Y H A W R P A W Y Q V T A A F C H A N W N H L S S N L F F L Y I F G K L V E E E G N F A L W S Y I L T G A G A N L V S W M V L P R T A V S V G A S G A V F G L F A I S V L V K M I ----- RWDWRK ILEVL ILGQFV I K R V M E A A Q A S A G M S G G -- --GFALQSVNHI AHL SGAL IGVLVWLL ISRIPSSKSKDQSSG - 222
MER1077849/1-28: 110 V F Q L R E I K A L Y L Y H N R P A W Y Q V T A T F C H A N W N H L S S N L F F L Y I F G K L V E E E G N F A L W S Y I L T G A G A N L V S W L I L P R C A V S V G A S G A V F G L F A I S V L V K M ----- SWDWRK ILEVL ILGQFV I D K V M E A A Q A S T G L A G -- --PYMQNVNHI AHL SGAL IGVALVLL ISRIPSPRT HDVKALR - 285
MER10113133/1-29: 106 L F Q V R A I K T L Y L Y H N R P A W Y Q V T A T F C H A N W N H L S S N L F F L Y I F G K L V E E E G N F G L W L S Y I L T G A G A N L V S W L V L P R N A V S V G A S G A V F G L F A I S V L V K I ----- SWDWRK ILEVL ILGQFV IEKVMEAAQASTGLSGTF F --LGGYS IQS INHIAHL SGAL IGVLVWLL SRVPSQPSNDEVSTLH - 286
MER0642474/1-28: 109 I L R L Q E I K S L Y L Y H I C P V W Y Q V T A T F C H A N W N H L S S N L F F L Y I F G K L V E E E G N F A L W S Y I L T G A G A N L V S W L L P K S A V S V G A S G A V F G L F V I S V L V K M ----- SWDWRK ILEVL ILGQFV IEKVMEAAQASTSLTGT Y -- --RGLMAVNHI AHL SGAL IGVALVLL VSRIPSPSGDDKNNK - 287
MER1076078/1-22: 44 L F K F Q G I Q S L Y L Y H N R P A W Y Q V T S T F C H A N W K H L S S N L F F L Y I F G K L V E E E G N F A L W L S Y I F T G A G A N I V S W L V L P R N A V S V G A S G A V F G L F A I S V L V K M ----- SWDWRK ILEVL ILGQFV I D K V M E A A Q A S T G L F G L -- --SPLQNVNHI AHL SGAL IGVLVWLL SKTSPSEPLDNEISSSP - 221
MER0436129/1-29: 114 L F Q V R I K S L Y L Y H N R P A W Y Q V T A T F C H A N W N H L S S N L F F L Y I F G K L V E E E G N L G L W L S Y L F T G V G A N L V S W L V L P R N A V S I G A S G A V F G L F A I S V L V K M ----- SWDWRK ILEVL ILGQFV I E R V M E A A Q A S A G V S G T V -- --YGGYSLQTVNHI AHL SGAL VGVLVWLL SRPSPSEFTDQKMKPH - 294
MER0650696/1-30: 115 F F Q V R S I K S L Y L Y H N R P A W Y Q V T A T F C H A S W C H L S S N L F F L Y I F G K L V E E E G N F G L W L S Y I L T G V G A N L V S W L V L P R N A V S V G A S G A V F G L F A I S V L V K I ----- TWDWRK ILEVL ILGQFV I E R V M E A A Q A S A L S G T F -- --RGGYPMQSVNHI AHL SGAL IGVLVWLL VSRIPSPRDMHESLRAK - 295
MER1078241/1-29: 116 F F Q V R S I K S L Y L Y H N R P A W Y Q V T A T F C H A S W N H L S S N L F F L Y I F G K L V E E E G N F G L W L S Y I L T G V G A N L V S W L V L P R N A V S V G A S G A V F G L F A I S V L V K M ----- SWDWRK ILEVL ILGQFV I E R V M E A A Q A S A G L S G T I -- --YGGYSLQTVNHI AHL SGAL VGVLVWLL SKFSETVDQDQSP - 296
MER0511706/1-29: 108 V F Q V R A I K M L Y L Y H N R P T W Y Q F L T A T F C H A N W N H L S S N L F F L Y I F G K L V E E E G N F A L W C Y I L T G V G A N L V S W L V L P R N A V S V G A S G A V F G L F T I S V L V K M ----- SWDWRK ILEVL ILGQFV I E K V M E A A Q A S T A L S G P F -- --GGGSLMNVNHI AHL SGAL VGVLVWLL SKVPSPPAQS - NFO - 287
MER0663899/1-29: 116 L F R L Q V K T L Y L Y H G W R P A W Y Q V T A T F C H A N W H L S S N L F F L Y I F G K L V E E E G N L A L W S Y I L T G A G A N L I S W L V L P R S A V S V G A S G A V F G L F A I S V L V K I ----- SWDWRK ILEVL ILGQFV I D K V L E A S T G L I D T W G T -- --GYAFP TVNHI AHL SGAL IGAAL I L V R H I T S Q R S G Q D I E R D K - 293
MER1075060/1-28: 106 F F H V N G A K T L Y L Y H R P A W Y Q V T A T F C H A N W K H L S S N L F F L Y I F G K L V E E E G N L A L W S Y I L T G V G A N L V S W L V L P R N T V S V G A S G A V F G L F T I S V L V K M ----- SWDWRK ILEVL ILGQFV IEKVMEAAQASTLSGF F -- --KGGYALQSVNHI AHL SGAL VGVLVWLL SRVPSDSSSYQ - - - - - 280
MER109413/1-30: 121 L L Q F V P I K W L Y L Y H D L P Q W Y Q F T A T F C H V N W S H L S S N L F F L Y I F G K L V E E E G G F A L W L S Y L T G A G A N L V S W L I L P R T V V S I G A S G A V F G L F A V S V L V K I ----- SWDWRK ILEVL ILGQFV IEKVMDAAQASTQFVGN N -- --SRILGVENI AHL SGAL VGVAL I W F L S K I P S G Q P K E G Q R K P - 301
MER0584988/1-29: 112 V F Q V R A I K M L Y L Y H N R P T W Y Q F L T A T F C H A N W S H L S S N L F F L Y I F G K L V E E E G N F A L W C Y I L T G V G A N L V S W L V L P R N A V S V G A S G A V F G L F T I S V L V K M ----- SWDWRK ILEVL ILGQFV IEKVMEAAQASTALSGTF F -- --GRNSLMNVNHI AHL SGAL VGVLVWLL SKVPSPPSDEGSNFO - 291
MER0551045/1-30: 114 V F Q V R A I K M L Y L Y H N R P T W Y Q F L T A T F C H A N W S H L S S N L F F L Y I F G K L V E E E G N F A L W C Y I L T G V G A N L V S W L V L P R N A V S V G A S G A V F G L F T I S V L V K M ----- SWDWRK ILEVL ILGQFV IEKVMEAAQASTALSGTF F -- --GRNSLMNVNHI AHL SGAL VGVLVWLL SKVPSPPSDEGSNFO - 293
MER0588401/1-30: 117 V F Q V R S I K S L Y L Y H N R P A W Y Q V T A A F C H A N W N H L S S N L F F L Y I F G K L V E E E G N F A L W S Y I F T G A A N F V S W L I L P K N A V S V G A S G A V F G L F A I S V L V K M ----- SWDWRK ILEVL ILGQFV IEKVMEAAQASAGMSGTF F -- --IGGYSVQS INHIAHL SGAL FGVFLVWLL SRVPSPPKKNKALNLR - 297
MER1075277/1-27: 108 V F K V S W I K T L Y L Y H S P A W Y Q V T A T F C H A S W H L S S N L F F L Y I F G K L V E E E G N F A L W I P Y I L T G V G A N L V S W L V L P R N A V S I G A S G A V F G L F A V S V L V K M ----- SWDWRK ILEVL ILGQFV IEKVMEAAQASANIGB F -- --DYAMQNVNHI AHL SGAL IGVL I W L L S S P P S T D B N -- - - - - 278
MER1083521/1-33: 145 L F Q V Q Q I K S L Y L Y H S W P T W Y Q F V T A T F C H A D W N H L S S N L F F L Y I F G K L V E E E G S F G V W S Y I T G V G A N L V S W L V L P R S A V S V G A S G A V F G L F A I S V L V K I ----- SWNRK ILEVL ILGQFV IEKVMEAAQASTRLP GTF -- --REVMQNVNHI AHL SGAL IGVAL I L L I A G V P S K S P E K D V K A L V - 323
MER1082771/1-27: 98 L L Q I R Q I K S L Y L Y H A F S W Y Q V T S T F C H A N W N H L S S N L F F Y I F G K L V E E E G N F A L W S Y I L T G A G A N L I S W L V L P T S S V S I G A S G A V F G L F T I S V L V K M ----- SWDWRK ILEVL ILGQFV D K V M E A A R A T I T -- - - - - - GA --ALQVNI AHL SGAL IGAALV F V I N R I R L S S N D N P K A S K - 271

```

Supplementary Figure 2: Multiple sequence alignment of the ORF1 protein sequence with top hits from the MEROPS database. The alignment shows extra 9 amino acids and two other amino acids that are present only in ORF1.

10 References

- Abad, P., J. Gouzy, J.M. Aury, P. Castagnone-Sereno, E.G.J. Danchin, E. Deleury, L. Perfus-Barbeoch, V. Anthouard, F. Artiguenave, V.C. Blok, M.C. Caillaud, P.M. Coutinho, C. Dasilva, F. De Luca, F. Deau, M. Esquibet, T. Flutre, J.V. Goldstone, N. Hamamouch, T. Hewezi, O. Jaillon, C. Jubin, P. Leonetti, M. Magliano, T.R. Maier, G.V. Markov, P. McVeigh, G. Pesole, J. Poulain, M. Robinson-Rechavi, E. Sallet, B. Ségurens, D. Steinbach, T. Tytgat, E. Ugarte, C. Van Ghelder, P. Veronico, T.J. Baum, M. Blaxter, T. Bleve-Zacheo, E.L. Davis, J.J. Ewbank, B. Favery, E. Grenier, B. Henrissat, J.T. Jones, V. Laudet, A.G. Maule, H. Quesneville, M.N. Rosso, T. Schiex, G. Smant, J. Weissenbach, and P. Wincker, 2008: Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*. *Nature Biotechnology* **26**, 909-915.
- Adamiec, M., M. Ciesielska, P. Zalaś, and R. Luciński, 2017: *Arabidopsis thaliana* intramembrane proteases, Vol. 39. Polish Academy of Sciences.
- Andrews, S., 2010: FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.
- Arnaud, J.F., F. Viard, M. Delescluse, and J. Cuguen, 2003: Evidence for gene flow via seed dispersal from crop to wild relatives in *Beta vulgaris* (Chenopodiaceae): consequences for the release of genetically modified crop species with weedy lineages. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**, 1565-1571.
- Arumuganathan, K., and E.D. Earle, 1991: Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter* **9**, 208-218.
- Bartsch, D., M. Lehnen, J. Clegg, M. Pohl-Orf, I.I. Schuphan, and N.C. Ellstrand, 1999: Impact of gene flow from cultivated beet on genetic diversity of wild sea beet populations. *Mol Ecol* **8**, 1733-41.
- Barzen, E., W. Mechelke, E. Ritter, E. Schulte-Kappert, and F. Salamini, 1995: An extended map of the sugar beet genome containing RFLP and RAPD loci. *Theoretical and Applied Genetics* **90**, 189-193.
- Berg, R.H., and C. Taylor, 2008: *Cell biology of plant nematode parasitism* Springer Science & Business Media.
- Blanc, C., M. Sy, D. Djigal, A. Brauman, P. Normand, and C. Villenave, 2006: Nutrition on bacteria by bacterial-feeding nematodes and consequences on the structure of soil bacterial community. *European Journal of Soil Biology* **42**, S70-S78.
- Blaxter, M.L., P. De Ley, J.R. Garey, L.X. Liu, P. Scheldeman, A. Vierstraete, J.R. Vanfleteren, L.Y. Mackey, M. Dorris, L.M. Frisse, J.T. Vida, and W.K. Thomas, 1998: A molecular evolutionary framework for the phylum Nematoda. *Nature* **392**, 71-75.

References

- Böckenhoff, A., and F.M.W. Grundler, 1994: Studies on the nutrient uptake by the beet cyst nematode *Heterodera schachtii* by in situ microinjection of fluorescent probes into the feeding structures in *Arabidopsis thaliana*. *Parasitology* **109**, 249-&.
- Böckenhoff, A., and F.M.W. Grundler, 2009: Studies on the nutrient uptake by the beet cyst nematode *Heterodera schachtii* by in situ microinjection of fluorescent probes into the feeding structures in *Arabidopsis thaliana*. *Parasitology* **109**, 249-255.
- Boetzer, M., and W. Pirovano, 2012: Toward almost closed genomes with GapFiller. *Genome Biology* **13**, R56-R56.
- Boex-Fontvieille, E., S. Rustgi, S. Reinbothe, and C. Reinbothe, 2015: A Kunitz-type protease inhibitor regulates programmed cell death during flower development in *Arabidopsis thaliana*. *Journal of Experimental Botany* **66**, 6119-6135.
- Bozkurt, T.O., S. Schornack, J. Win, T. Shindo, M. Ilyas, R. Oliva, L.M. Cano, A.M.E. Jones, E. Huitema, R.A.L. van der Hoorn, and S. Kamoun, 2011: *Phytophthora infestans* effector AVRblb2 prevents secretion of a plant immune protease at the haustorial interface. *Proceedings of the National Academy of Sciences* **108**, 20832-20837.
- Cai, D., M. Kleine, S. Kifle, H.J. Harloff, N.N. Sandal, K.A. Marcker, R.M. Klein-Lankhorst, E.M.J. Salentijn, W. Lange, W.J. Stiekema, U. Wyss, F.M.W. Grundler, and C. Jung, 1997: Positional cloning of a gene for nematode resistance in sugar beet. *Science* **275**, 832-834.
- Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, and T.L. Madden, 2009: BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421-421.
- Capistrano, G., 2010: A candidate sequence for the nematode resistance gene *Hs1-2* in sugar beet.
- Chen, R., H. Li, L. Zhang, J. Zhang, J. Xiao, and Z. Ye, 2007: *CaMi*, a root-knot nematode resistance gene from hot pepper (*Capsium annuum L.*) confers nematode resistance in tomato. *Plant Cell Reports* **26**, 895-905.
- Chikhi, R., and P. Medvedev, 2013: Informed and automated k-mer size selection for genome assembly. *Bioinformatics* **30**, 31-37.
- Cho, J.H., J.H. Lee, Y.K. Park, M.N. Choi, and K.-N. Kim, 2016: Calcineurin B-like protein CBL10 directly interacts with TOC34 (Translocon of the Outer membrane of the Chloroplasts) and decreases its GTPase activity in *Arabidopsis*. *Frontiers in Plant Science* **7**.
- Claverie, M., E. Dirlwanger, N. Bosselut, C. Van Ghelder, R. Voisin, M. Kleinhenz, B. Lafargue, P. Abad, M.-N. Rosso, B. Chalhoub, and D. Esmenjaud, 2011: The *Ma* Gene for Complete-Spectrum Resistance to *Meloidogyne* Species in *Prunus* Is a TNL with a Huge Repeated C-Terminal Post-LRR Region. *Plant Physiology* **156**, 779-792.
- Coll, N.S., D. Vercammen, A. Smidler, C. Clover, F. Van Breusegem, J.L. Dangl, and P. Epple, 2010: *Arabidopsis* type I metacaspases control cell death. *Science* **330**, 1393-1397.

References

- Cook, D.E., T.G. Lee, X. Guo, S. Melito, K. Wang, A.M. Bayless, J. Wang, T.J. Hughes, D.K. Willis, T.E. Clemente, B.W. Diers, J. Jiang, M.E. Hudson, and A.F. Bent, 2012: Copy Number Variation of Multiple Genes at *Rhg1* Mediates Nematode Resistance in Soybean. *Science* **338**, 1206-1209.
- Coons, G.H., 1975: Interspecific hybrids between *Beta vulgaris* L. and the wild species of Beta. *J Am Soc Sugar Beet Technol* **18**, 281-306.
- Coursol, S., L.M. Fan, H.L. Stunff, S. Splegel, S. Gilroy, and S.M. Assman, 2003: Sphingolipid signalling in Arabidopsis guard cells involves heterotrimeric G proteins. *Nature* **423**, 651-654.
- Dalke, L., 1977: Interspecific hybrids between sugar beet and *Beta corolliflora* of the Corollinae section Interspecific hybridization in plant breeding. Proceedings of the 8th Eucarpia Congress, Madrid, Spain, 113-118.
- De Almeida Engler, J., K. Van Poucke, M. Karimi, R. De Groot, G. Gheysen, G. Engler, and G. Gheysen, 2004: Dynamic cytoskeleton rearrangements in giant cells and syncytia of nematode-infected roots. *The Plant Journal* **38**, 12-26.
- Dechyeva, D., and T. Schmidt, 2006: Molecular organization of terminal repetitive DNA in Beta species. *Chromosome Research* **14**, 881-897.
- Dechyeva, D., F. Gindullis, and T. Schmidt, 2003: Divergence of satellite DNA and interspersion of dispersed repeats in the genome of the wild beet *Beta procumbens*. *Chromosome Research* **11**, 3-21.
- Decraemer, W., and D.J. Hunt, 2006: Structure and classification, In: R. N. Perry and M. Moens, (eds.) *Plant nematology*, 3-32. CABI Publishing: Wallingford, CA, USA.
- Desel, C., C. Jung, D. Cai, M. Kleine, and T. Schmidt, 2001: High-resolution mapping of YACs and the single-copy gene Hs1pro-1 on *Beta vulgaris* chromosomes by multi-colour fluorescence in situ hybridization, 113-122, Vol. 45.
- Díaz-Mendoza, M., B. Velasco-Arroyo, P. González-Melendi, M. Martínez, and I. Díaz, 2014: C1A cysteine protease–cystatin interactions in leaf senescence. *Journal of experimental botany* **65**, 3825-3833.
- Ding, B., M.d.R. Bellizzi, Y. Ning, B.C. Meyers, and G.-L. Wang, 2012: HDT701, a Histone H4 Deacetylase, Negatively Regulates Plant Innate Immunity by Modulating Histone H4 Acetylation of Defense-Related Genes in Rice *The Plant Cell* **24**, 3783-3794.
- Dohm, J.C., A.E. Minoche, D. Holtgräwe, S. Capella-Gutiérrez, F. Zakrzewski, H. Tafer, O. Rupp, T.R. Sørensen, R. Stracke, R. Reinhardt, A. Goesmann, T. Kraft, B. Schulz, P.F. Stadler, T. Schmidt, T. Gabaldón, H. Lehrach, B. Weisshaar, and H. Himmelbauer, 2014a: The genome of the recently domesticated crop plant sugar beet (*Beta vulgaris*). *Nature* **505**, 546-549.
- Dohm, J.C., A.E. Minoche, D. Holtgräwe, S. Capella-Gutiérrez, F. Zakrzewski, H. Tafer, O. Rupp, T.R. Sørensen, R. Stracke, R. Reinhardt, A. Goesmann, T. Kraft, B. Schulz, P.F. Stadler, T. Schmidt, T. Gabaldón, H. Lehrach, B. Weisshaar, and H. Himmelbauer,

References

- 2014b: The genome of the recently domesticated crop plant sugar beet (*Beta vulgaris*). *Nature* **505**.
- Elashry, A.M., S.S. Habash, P. Vijayapalani, N. Brocke-Ahmadinejad, R. Blumel, A. Seetharam, H. Schoof, and F.M.W. Grundler, 2020: Transcriptome and Parasitome Analysis of Beet Cyst Nematode *Heterodera schachtii*. *Scientific Reports* **10**, 12.
- Enrico Biancardi, L.W.P., Robert T. Lewellen, 2012: *Beta maritima: The origin of Beets* Springer.
- Enrico Biancardi, M.D.B., Larry G Campbell, 2005: *Genetics and Breeding of Sugar Beet*. 1 ed. CRC Press.
- Ernst, K., A. Kumar, D. Kriseleit, D.U. Kloos, M.S. Phillips, and M.W. Ganai, 2002: The broad-spectrum potato cyst nematode resistance gene (*Hero*) from tomato is the only member of a large gene family of NBS-LRR genes with an unusual amino acid repeat in the LRR region. *Plant Journal* **31**, 127-136.
- Fan, G., Y. Yang, T. Li, W. Lu, Y. Du, X. Qiang, Q. Wen, and W. Shan, 2018: A *Phytophthora capsici* RXLR Effector Targets and Inhibits a Plant PPIase to Suppress Endoplasmic Reticulum-Mediated Immunity. *Molecular Plant* **11**, 1067-1083.
- Fausser, F., S. Schiml, and H. Puchta, 2014: Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. *Plant J* **79**, 348-59.
- Feldman, M., and A.A. Levy, 2015: Origin and Evolution of Wheat and Related Triticeae Species, In: M. Molnár-Láng, C. Ceoloni and J. Doležel, (eds.) *Alien Introgression in Wheat: Cytogenetics, Molecular Biology, and Genomics*, 21-76. Springer International Publishing, Cham.
- Finkenstadt, V.L., 2014: A Review on the Complete Utilization of the Sugarbeet. *Sugar Tech* **16**, 339-346.
- Fischer, H.E., 1989: Origin of the 'Weisse Schlesische Rübe' (white Silesian beet) and resynthesis of sugar beet. *Euphytica* **41**, 75-80.
- Flavell, R.B., M.D. Bennett, J.B. Smith, and D.B. Smith, 1974: Genome size and the proportion of repeated nucleotide sequence DNA in plants. *Biochemical Genetics* **12**, 257-269.
- Frese, L., 2010: Conservation and access to sugarbeet germplasm. *Sugar Tech* **12**, 207-219.
- Funk, A., P. Galewski, and J.M. McGrath, 2018: Nucleotide-binding resistance gene signatures in sugar beet, insights from a new reference genome. *Plant Journal* **95**, 659-671.
- Gaafar, R.M., U. Hohmann, and C. Jung, 2005: Bacterial artificial chromosome-derived molecular markers for early bolting in sugar beet. *Theoretical and Applied Genetics* **110**, 1027-1037.

References

- Galewski, P., and J.M. McGrath, 2020: Genetic diversity among cultivated beets (*Beta vulgaris*) assessed via population-based whole genome sequences. *BMC Genomics* **21**, 189.
- Gao, D., T. Schmidt, and C. Jung, 2000: Molecular characterization and chromosomal distribution of species-specific repetitive DNA sequences from *Beta corolliflora*, a wild relative of sugar beet. *Genome* **43**, 1073-1080.
- Gel, B., and E. Serra, 2017: karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* **33**, 3088-3090.
- Geldner, N., J. Friml, Y.-D. Stierhof, G. Jürgens, and K. Palme, 2001: Auxin transport inhibitors block PIN1 cycling and vesicle trafficking. *Nature* **413**, 425-428.
- Geng, C., X. Nie, Z. Tang, Y. Zhang, J. Lin, M. Sun, and D. Peng, 2016: A novel serine protease, Sep1, from *Bacillus firmus* DS-1 has nematocidal activity and degrades multiple intestinal-associated nematode proteins. *Scientific Reports* **6**, 25012.
- Ghaemi, R., E. Pourjam, N. Safaie, B. Verstraeten, S.B. Mahmoudi, R. Mehrabi, T. De Meyer, and T. Kyndt, 2020: Molecular insights into the compatible and incompatible interactions between sugar beet and the beet cyst nematode. *BMC Plant Biology* **20**, 483.
- Gilroy, E.M., I. Hein, R. Van Der Hoorn, P.C. Boevink, E. Venter, H. McLellan, F. Kaffarnik, K. Hrubikova, J. Shaw, M. Holeva, E.C. López, O. Borrás-Hidalgo, L. Pritchard, G.J. Loake, C. Lacomme, and P.R.J. Birch, 2007: Involvement of cathepsin B in the plant disease resistance hypersensitive response. *Plant Journal* **52**, 1-13.
- Golinowski, W., F.M.W. Grundler, and M. Sobczak, 1996: Changes in the structure of *Arabidopsis thaliana* during female development of the plant-parasitic nematode *Heterodera schachtii*. *Protoplasma* **194**, 103-116.
- Goverse, A., H. Overmars, J. Engelbertink, A. Schots, J. Bakker, and J. Helder, 2000: Both Induction and Morphogenesis of Cyst Nematode Feeding Cells Are Mediated by Auxin. *Molecular Plant-Microbe Interactions*® **13**, 1121-1129.
- Griffiths AJF, G.W., Miller JH, et al., 1999: *Modern Genetic Analysis* W.H. Freeman & Co Ltd.
- Grundler, F.M.W., M. Sobczak, and S. Lange, 1997: Defence responses of *Arabidopsis thaliana* during invasion and feeding site induction by the plant-parasitic nematode *Heterodera glycines*. *Physiological and Molecular Plant Pathology* **50**, 419-429.
- Grundler, F.M.W., M. Sobczak, and W. Golinowski, 1998: Formation of wall openings in root cells of *Arabidopsis thaliana* following infection by the plant-parasitic nematode *Heterodera schachtii*. *European Journal of Plant Pathology* **104**, 545-551.
- Grunewald, W., B. Cannoot, J. Friml, and G. Gheysen, 2009: Parasitic Nematodes Modulate PIN-Mediated Auxin Transport to Facilitate Infection. *PLOS Pathogens* **5**, e1000266.
- Harlan, J.R., and J.M.J. de Wet, 1971: Toward a rational classification of cultivated plants. *TAXON* **20**, 509-517.

References

- Hatsugai, N., S. Iwasaki, K. Tamura, M. Kondo, K. Fuji, K. Ogasawara, M. Nishimura, and I. Hara-Nishimura, 2009: A novel membrane fusion-mediated plant immunity against bacterial pathogens. *Genes and Development* **23**, 2496-2506.
- Hawes, C., P. Kiviniemi, and V. Kriechbaumer, 2015: The endoplasmic reticulum: A dynamic and well-connected organelle. *Journal of Integrative Plant Biology* **57**, 50-62.
- Heijbroek, W., A.J. Roelands, J.H. de Jong, C. van Hulst, A.H.L. Schoone, and R.G. Munning, 1988: Sugar beets homozygous for resistance to beet cyst nematode (*Heterodera schachtii* Schm.), developed from monosomic additions of *Beta procumbens* to *B. vulgaris*. *Euphytica* **38**, 121-131.
- Heller, R., J. Schondelmaier, G. Steinrücken, and C. Jung, 1996a: Genetic localization of four genes for nematode (*Heterodera schachtii* Sch.) resistance in sugar beet (*Beta vulgaris* L.). *Theoretical and Applied Genetics* **92**, 991-997.
- Heller, R., J. Schondelmaier, G. Steinrücken, and C. Jung, 1996b: Genetic localization of four genes for nematode (*Heterodera schachtii* Sch.) resistance in sugar beet (*Beta vulgaris* L.). *Theoretical and Applied Genetics* **92**, 991-997.
- Hoffmann, C.M., 2010: Root Quality of Sugarbeet. *Sugar Tech* **12**, 276-287.
- Hoffmann, C.M., and B. Märlander, 2005: Composition of harmful nitrogen in sugar beet (*Beta vulgaris* L.)—amino acids, betaine, nitrate—as affected by genotype and environment. *European Journal of Agronomy* **22**, 255-265.
- Hohmann, U., G. Jacobs, A. Telgmann, R.M. Gaafar, S. Alam, and C. Jung, 2003: A bacterial artificial chromosome (BAC) library of sugar beet and a physical map of the region encompassing the bolting gene B. *Molecular Genetics and Genomics* **269**, 126-136.
- Holt, C., and M. Yandell, 2011: MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* **12**, 491-491.
- Holtmann, B., M. Kleine, and F.M.W. Grundler, 2000: Ultrastructure and anatomy of nematode-induced syncytia in roots of susceptible and resistant sugar beet. *Protoplasma* **211**, 39-50.
- Hou, S., P. Jamieson, and P. He, 2018: The cloak, dagger, and shield: proteases in plant–pathogen interactions, 2491-2509, Vol. 475. Portland Press Ltd.
- Höwing, T., C. Huesmann, C. Hoefle, M.K. Nagel, E. Isono, R. Hückelhoven, and C. Gietl, 2014: Endoplasmic reticulum KDEL-tailed cysteine endopeptidase 1 of *Arabidopsis* (AtCEP1) is involved in pathogen defense. *Frontiers in Plant Science* **5**.
- Höwing, T., M. Dann, B. Müller, M. Helm, S. Scholz, K. Schneitz, U.Z. Hammes, and C. Gietl, 2018: The role of KDEL-tailed cysteine endopeptidases of *Arabidopsis* (AtCEP2 and AtCEP1) in root development. *PLOS ONE* **13**, e0209407.
- Hussey, R.S., 1989: Disease-Inducing Secretions of Plant-Parasitic Nematodes. *Annual Review of Phytopathology* **27**, 123-141.

References

- Ilyas, M., A.C. Hörger, T.O. Bozkurt, H.A. Van Den Burg, F. Kaschani, M. Kaiser, K. Belhaj, M. Smoker, M.H.A.J. Joosten, S. Kamoun, and R.A.L. Van Der Hoorn, 2015: Functional Divergence of Two Secreted Immune Proteases of Tomato. *Current Biology* **25**, 2300-2306.
- Jablonska, B., J.S.S. Ammiraju, K.K. Bhattarai, S. Mantelin, O.M. de Ilarduya, P.A. Roberts, and I. Kaloshian, 2007: The *Mi-9* Gene from *Solanum arcanum* Conferring Heat-Stable Resistance to Root-Knot Nematodes Is a Homolog of *Mi-1*. *Plant Physiology* **143**, 1044-1054.
- Jacobs, T.B., P.R. LaFayette, R.J. Schmitz, and W.A. Parrott, 2015: Targeted genome modifications in soybean with CRISPR/Cas9. *BMC Biotechnology* **15**, 16-16.
- Jaeger, S.C., G.G.G. Capistrano, H.L. Harloff, and C. Jung, 2008: γ -irradiation of wild beet translocation lines and monosomic addition lines in sugar beet carrying nematode resistance genes.
- Jäger, S.C., 2012: Hybrid Assembly of Whole Genome Shotgun Sequences of Two Sugar Beet.
- Jones, J.T., A. Haegeman, E.G.J. Danchin, H.S. Gaur, J. Helder, M.G.K. Jones, T. Kikuchi, R. Manzanilla-López, J.E. Palomares-Rius, W.M.L. Wesemael, and R.N. Perry, 2013: Top 10 plant-parasitic nematodes in molecular plant pathology. *Molecular Plant Pathology* **14**, 946-961.
- Jung, C., and G. Wricke, 1987: Selection of Diploid Nematode-Resistant Sugar Beet from Monosomic Addition Lines. *Plant Breeding* **98**, 205-214.
- Jung, C., P. Wehling, and H. Löptien, 1986: Electrophoretic Investigations on Nematode Resistant Sugar Beets. *Plant Breeding* **97**, 39-45.
- Jung, C., D. Cai, and M. Kleine, 1998: Engineering nematode resistance in crop species, 266-271, Vol. 3. Elsevier.
- Jung, C., K. Pillen, L. Frese, S. Fähr, and A.E. Melchinger, 1993: Phylogenetic relationships between cultivated and wild species of the genus *Beta* revealed by DNA "fingerprinting". *Theoretical and Applied Genetics* **86**, 449-457.
- Kadereit, G., S. Hohmann, and J.W. Kadereit, 2006: A synopsis of *Chenopodiaceae* subfam. *Betoideae* and notes on the taxonomy of *Beta*. *Willdenowia* **36**, 9-19, 11.
- Kifle, S., M. Shao, C. Jung, and D. Cai, 1999: An improved transformation protocol for studying gene expression in hairy roots of sugar beet (*Beta vulgaris* L.). *Plant Cell Reports* **18**, 514-519.
- Kleine, M., D. Cai, C. Elbl, R.G. Herrmann, and C. Jung, 1995: Physical mapping and cloning of a translocation in sugar beet (*Beta vulgaris* L.) carrying a gene for nematode (*Heterodera schachtii*) resistance from *B. procumbens*. *Theoretical and Applied Genetics* **90**, 399-406.

- Knopf, R.R., A. Feder, K. Mayer, A. Lin, M. Rozenberg, A. Schaller, and Z. Adam, 2012: Rhomboid proteins in the chloroplast envelope affect the level of allene oxide synthase in *Arabidopsis thaliana*. *Plant J* **72**, 559-71.
- Krüger, J., C.M. Thomas, C. Golstein, M.S. Dixon, and M. Smoker, 2002: A tomato cysteine protease required for *Cf-2* -dependent disease resistance and suppression of autonecrosis. *Science* **296**.
- Kühnle, N., V. Dederer, and M.K. Lemberg, 2019: Intramembrane proteolysis at a glance: from signalling to protein degradation. *Journal of Cell Science* **132**.
- Kumar, V., M.R. Khan, and R.K. Walia, 2020: Crop Loss Estimations due to Plant-Parasitic Nematodes in Major Crops in India. *National Academy Science Letters* **43**, 409-412.
- Lange, W., J. Müller, and T. De Bock, 1993: Virulence in the beet cyst nematode (*Heterodera schachtii*) versus some alien genes for resistance in beet. *Fundamental and applied nematology* **16**, 447-454.
- Larriba, E., M.D.L.A. Jaime, J. Carbonell-Caballero, A. Conesa, J. Dopazo, C. Nislow, J. Martín-Nieto, and L.V. Lopez-Llorca, 2014: Sequencing and functional analysis of the genome of a nematode egg-parasitic fungus, *Pochonia chlamydosporia*. *Fungal Genetics and Biology* **65**, 69-80.
- Lee, T., S. Yang, E. Kim, Y. Ko, S. Hwang, J. Shin, J.E. Shim, H. Shim, H. Kim, C. Kim, and I. Lee, 2014: AraNet v2: an improved database of co-functional gene networks for the study of *Arabidopsis thaliana* and 27 other nonmodel plant species. *Nucleic Acids Research* **43**, D996-D1002.
- Lelley, T., C. Eder, and H. Grausgruber, 2004: Influence of 1BL.1RS wheat-rye chromosome translocation on genotype by environment interaction. *Journal of Cereal Science* **39**, 313-320.
- Letunic, I., and P. Bork, 2019: Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research* **47**, W256-W259.
- Li, H., 2013: Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
- Liu, S., P.K. Kandoth, S.D. Warren, G. Yeckel, R. Heinz, J. Alden, C. Yang, A. Jamai, T. El-Mellouki, P.S. Juvale, J. Hill, T.J. Baum, S. Cianzio, S.A. Whitham, D. Korkin, M.G. Mitchum, and K. Meksem, 2012: A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. *Nature* **492**, 256-260.
- Löptien, H., 1984a: Breeding nematode-resistant beets. II. Investigations into inheritance of resistance to *Heterodera schachtii* Schm. in wild species of the section Patellares. *Zeitschrift für Pflanzenzüchtung* **93**, 237-245.
- Löptien, H., 1984b: Breeding nematode-resistant beets - I. Development of resistant alien additions by crosses between *Beta vulgaris* and wild species of the section Patellares. *Journal of plant breeding* **92**.

References

- Lozano-Torres, J.L., R.H.P. Wilbers, P. Gawronski, J.C. Boshoven, A. Finkers-Tomczak, J.H.G. Cordewener, A.H.P. America, H.A. Overmars, J.W. Van 't Klooster, L. Baranowski, M. Sobczak, M. Ilyas, R.A.L. Van Der Hoorn, A. Schots, P.J.G.M. De Wit, J. Bakker, A. Goverse, and G. Smant, 2012: Dual disease resistance mediated by the immune receptor Cf-2 in tomato requires a common virulence target of a fungus and a nematode. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 10119-10124.
- Luo, R., B. Liu, Y. Xie, Z. Li, W. Huang, J. Yuan, G. He, Y. Chen, Q. Pan, Y. Liu, J. Tang, G. Wu, H. Zhang, Y. Shi, Y. Liu, C. Yu, B. Wang, Y. Lu, C. Han, D.W. Cheung, S.-M. Yiu, S. Peng, Z. Xiaoqian, G. Liu, X. Liao, Y. Li, H. Yang, J. Wang, T.-W. Lam, and J. Wang, 2012: SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience* **1**, 18-18.
- Madeira, F., Y.M. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A.R.N. Tivey, S.C. Potter, R.D. Finn, and R. Lopez, 2019: The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic acids research* **47**, W636-W641.
- Mago, R., H. Miah, G.J. Lawrence, C.R. Wellings, W. Spielmeier, H.S. Bariana, R.A. McIntosh, A.J. Pryor, and J.G. Ellis, 2005: High-resolution mapping and mutation analysis separate the rust resistance genes Sr31, Lr26 and Yr9 on the short arm of rye chromosome 1. *Theoretical and Applied Genetics* **112**, 41-50.
- Mani, R.-S., and A.M. Chinnaiyan, 2010: Triggers for genomic rearrangements: insights into genomic, cellular and environmental influences. *Nature Reviews Genetics* **11**, 819-829.
- Maung, T.A., and C.R. Gustafson, 2011: The economic feasibility of sugar beet biofuel production in central North Dakota. *Biomass and Bioenergy* **35**, 3737-3747.
- McCarter, J.P., 2009: Molecular approaches toward resistance to plant-parasitic nematodes. , In: R. H. Berg and C. E. Taylor, (eds.) *Cell biology of plant nematode parasitism series*, Vol. 15.
- McGrath, J.M., A. Funk, P. Galewski, S. Ou, B. Townsend, K. Davenport, H. Daligault, S. Johnson, J. Lee, A. Hastie, A. Darracq, G. Willems, S. Barnes, I. Liachko, S. Sullivan, S. Koren, A. Phillippy, J. Wang, T. Lu, J. Pulman, K. Childs, A. Yocum, D. Fermin, E. Mutasa-Göttgens, P. Stevanato, K. Taguchi, and K. Dorn, 2020: A contiguous *de novo* genome assembly of sugar beet EL10 (*Beta vulgaris* L.). *bioRxiv*, 2020.09.15.298315.
- McK Bird, D., 1996: Manipulation of host gene expression by root-knot nematodes. *The Journal of parasitology*, 881-888.
- McKendry, A.L., D.N. Tague, and K.E. Miskin, 1996: Effect of 1BL.1RS on Agronomic Performance of Soft Red Winter Wheat. *Crop Science* **36**, crops1996.0011183X003600040003x.
- McLellan, H., P.C. Boevink, M.R. Armstrong, L. Pritchard, S. Gomez, J. Morales, S.C. Whisson, J.L. Beynon, and P.R.J. Birch, 2013: An RxLR Effector from *Phytophthora infestans* Prevents Re-localisation of Two Plant NAC Transcription Factors from the Endoplasmic Reticulum to the Nucleus. *PLOS Pathogens* **9**, e1003670.

References

- Meyer, M., F. Huttenlocher, A. Cedzich, S. Procopio, J. Stroeder, C. Pau-Roblot, M. Lequart-Pillon, J. Pelloux, A. Stintzi, and A. Schaller, 2016: The subtilisin-like protease SBT3 contributes to insect resistance in tomato. *Journal of Experimental Botany* **67**, 4325-4338.
- Milligan, S.B., J. Bodeau, J. Yaghoobi, I. Kaloshian, P. Zabel, and V.M. Williamson, 1998a: The Root Knot Nematode Resistance Gene *Mi* from Tomato Is a Member of the Leucine Zipper, Nucleotide Binding, Leucine-Rich Repeat Family of Plant Genes, 1307-1319, Vol. 10.
- Milligan, S.B., J. Bodeau, J. Yaghoobi, I. Kaloshian, P. Zabel, and V.M. Williamson, 1998b: The Root Knot Nematode Resistance Gene *Mi* from Tomato Is a Member of the Leucine Zipper, Nucleotide Binding, Leucine-Rich Repeat Family of Plant Genes. *The Plant Cell* **10**, 1307-1319.
- Mita, G., M. Dani, P. Casciari, A. Pasquali, E. Selva, C. Minganti, and P. Piccardi, 1991: Assessment of the degree of genetic variation in beet based on RFLP analysis and the taxonomy of *Beta*. *Euphytica* **55**, 1-6.
- Mohan, S., P.W.K. Ma, T. Pechan, E.R. Bassford, W.P. Williams, and D.S. Luthe, 2006: Degradation of the *S. frugiperda* peritrophic matrix by an inducible maize cysteine protease. *Journal of Insect Physiology* **52**, 21-28.
- Murata, S., H. Yashiroda, and K. Tanaka, 2009: Molecular mechanisms of proteasome assembly. *Nature Reviews Molecular Cell Biology* **10**, 104-115.
- Nachtigall, M., L. Bülow, J. Schubert, and L. Frese, 2016: Development of SSR markers for the genus *Patellifolia* (Chenopodiaceae). *Applications in Plant Sciences* **4**, 1600040.
- Niebel, A., J. De Almeida Engler, A. Hemerly, P. Ferreira, D. Inzé, M. Van Montagu, and G. Gheysen, 1996: Induction of *cdc2a* and *cyc1At* expression in *Arabidopsis thaliana* during early phases of nematode-induced feeding cell formation. *The Plant Journal* **10**, 1037-1043.
- Nombela, G., V.M. Williamson, and M. Muñoz, 2003: The root-knot nematode resistance gene *Mi-1.2* of tomato is responsible for resistance against the whitefly *Bemisia tabaci*. *Mol Plant Microbe Interact* **16**, 645-9.
- Oberschmidt, O., F.M.W. Grundler, and M. Kleine, 2003: Identification of a putative cation transporter gene from sugar beet (*Beta vulgaris* L.) by DDRT-PCR closely linked to the beet cyst nematode resistance gene *Hs1pro-1*. *Plant Science* **165**, 777-784.
- Okada, H., H. Harada, and I. Kadota, 2005: Fungal-feeding habits of six nematode isolates in the genus *Filenchus*. *Soil Biology and Biochemistry* **37**, 1113-1120.
- Paal, J., H. Henselewski, J. Muth, K. Meksem, C.M. Menéndez, F. Salamini, A. Ballvora, and C. Gebhardt, 2004: Molecular cloning of the potato *Gro1-4* gene conferring resistance to pathotype *Ro1* of the root cyst nematode *Globodera rostochiensis*, based on a candidate gene approach. *Plant Journal* **38**, 285-297.
- Paesold, S., D. Borchardt, T. Schmidt, and D. Dechyeva, 2012: A sugar beet (*Beta vulgaris* L.) reference FISH karyotype for chromosome and chromosome-arm

References

- identification, integration of genetic linkage groups and analysis of major repeat family distribution. *Plant Journal* **72**, 600-611.
- Pechan, T., A. Cohen, W.P. Williams, and D.S. Luthe, 2002: Insect feeding mobilizes a unique plant defense protease that disrupts the peritrophic matrix of caterpillars. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 13319-13323.
- Pechan, T., L. Ye, Y.-m. Chang, A. Mitra, L. Lin, F.M. Davis, W.P. Williams, and D.S. Luthe, 2000: A Unique 33-kD Cysteine Proteinase Accumulates in Response to Larval Feeding in Maize Genotypes Resistant to Fall Armyworm and Other Lepidoptera. *The Plant Cell* **12**, 1031-1040.
- Pidkowich, M.S., H.T. Nguyen, I. Heilmann, T. Ischebeck, and J. Shanklin, 2007: Modulating seed beta-ketoacyl-acyl carrier protein synthase II level converts the composition of a temperate seed oil to that of a palm-like tropical oil. *Proc Natl Acad Sci U S A* **104**, 4742-7.
- Pillen, K., G. Sleinrücken, R.G. Herrmann, and C. Jung, 1993: An Extended Linkage Map of Sugar Beet (*Beta vulgaris* L.) Including Nine Putative Lethal Genes and the Restorer Gene X. *Plant Breeding* **111**, 265-272.
- Pogorelko, G.V., P.S. Juvale, W.B. Rutter, M. Hütten, T.R. Maier, T. Hewezi, J. Paulus, R.A.L. van der Hoorn, F.M.W. Grundler, S. Siddique, V. Lionetti, O.A. Zabolina, and T.J. Baum, 2019a: Re-targeting of a plant defense protease by a cyst nematode effector. *Plant Journal* **98**, 1000-1014.
- Pogorelko, G.V., P.S. Juvale, W.B. Rutter, M. Hütten, T.R. Maier, T. Hewezi, J. Paulus, R.A. van der Hoorn, F.M. Grundler, S. Siddique, V. Lionetti, O.A. Zabolina, and T.J. Baum, 2019b: Re-targeting of a plant defense protease by a cyst nematode effector. *The Plant Journal* **98**, 1000-1014.
- Potter, S.C., A. Luciani, S.R. Eddy, Y. Park, R. Lopez, and R.D. Finn, 2018: HMMER web server: 2018 update. *Nucleic Acids Research* **46**, W200-W204.
- Prasad, B.D., G. Creissen, C. Lamb, and B.B. Chattoo, 2009: Overexpression of Rice (*Oryza sativa* L.) OsCDR1 Leads to Constitutive Activation of Defense Responses in Rice and Arabidopsis. *Molecular Plant-Microbe Interactions*® **22**, 1635-1644.
- Qiao, F., 2014: Cloning of the *Hs1-2* gene for nematode resistance from sugar beet.
- Quandt, H.J., 1993: Transgenic Root Nodules of *Vicia hirsuta*: A Fast and Efficient System for the Study of Gene Expression in Indeterminate-Type Nodules. *Molecular Plant-Microbe Interactions* **6**, 699-699.
- Quinlan, A.R., and I.M. Hall, 2010: BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)* **26**, 841-2.
- R. Viglierchio, D., 1960: R. Viglierchio **395**, 389-395.
- Ramírez, V., A. López, B. Mauch-Mani, M.J. Gil, and P. Vera, 2013: An Extracellular Subtilase Switch for Immune Priming in Arabidopsis. *PLOS Pathogens* **9**, e1003445.

References

- Raski, D., 1950: The life history and morphology of the sugar-beet nematode, *Heterodera schachtii* Schmidt. *Phytopathology* **40**, 135-152.
- Rawlings, N.D., A.J. Barrett, and A. Bateman, 2010: MEROPS: the peptidase database. *Nucleic acids research* **38**, D227-33.
- Rawlings, N.D., A.J. Barrett, P.D. Thomas, X. Huang, A. Bateman, and R.D. Finn, 2017: The *MEROPS* database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Research* **46**, D624-D632.
- Reamon-Ramos, S.M., and G. Wricke, 1992: A full set of monosomic addition lines in *Beta vulgaris* from *Beta webbiana*: morphology and isozyme markers. *Theoretical and Applied Genetics* **84**, 411-418.
- Ren, T.-H., Z.-J. Yang, B.-J. Yan, H.-Q. Zhang, S.-L. Fu, and Z.-L. Ren, 2009: Development and characterization of a new 1BL.1RS translocation line with resistance to stripe rust and powdery mildew of wheat. *Euphytica* **169**, 207-213.
- Ren, T., Z. Tang, S. Fu, B. Yan, F. Tan, Z. Ren, and Z. Li, 2017: Molecular Cytogenetic Characterization of Novel Wheat-rye T1RS.1BL Translocation Lines with High Resistance to Diseases and Great Agronomic Traits. *Frontiers in Plant Science* **8**.
- Reynolds, M.P., J.M. Lewis, K. Ammar, B.R. Basnet, L. Crespo-Herrera, J. Crossa, K.S. Dhugga, S. Dreisigacker, P. Juliana, H. Karwat, M. Kishii, M.R. Krause, P. Langridge, A. Lashkari, S. Mondal, T. Payne, D. Pequeno, F. Pinto, C. Sansaloni, U. Schulthess, R.P. Singh, K. Sonder, S. Sukumaran, W. Xiong, and H.J. Braun, 2021: Harnessing translational research in wheat for climate resilience. *Journal of Experimental Botany* **72**, 5134-5157.
- Richards, E.J., and F.M. Ausubel, 1988: Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. *Cell* **53**, 127-136.
- Rodríguez del Río, Á., A.E. Minoche, N.F. Zwickl, A. Friedrich, S. Liedtke, T. Schmidt, H. Himmelbauer, and J.C. Dohm, 2019: Genomes of the wild beets *Beta patula* and *Beta vulgaris* ssp. *maritima*. *Plant Journal* **99**, 1242-1253.
- Rodríguez, L.A., M.E. Toro, F. Vazquez, M.L. Correa-Daneri, S.C. Gouiric, and M.D. Vallejo, 2010: Bioethanol production from grape and sugar beet pomaces by solid-state fermentation. *International Journal of Hydrogen Energy* **35**, 5914-5917.
- Rogers, S.O., and A.J. Bendich, 1985: Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues, 69-76, Vol. 5.
- Romeiras, M.M., A. Vieira, D.N. Silva, M. Moura, A. Santos-Guerra, D. Batista, M.C. Duarte, and O.S. Paulo, 2016: Evolutionary and Biogeographic Insights on the Macaronesian *Beta-Patellifolia* Species (Amaranthaceae) from a Time-Scaled Molecular Phylogeny. *PLOS ONE* **11**, e0152456.
- Rooney, H.C., J.W. Van't Klooster, R.A. van der Hoorn, M.H. Joosten, J.D. Jones, and P.J. de Wit, 2005: Cladosporium Avr2 inhibits tomato Rcr3 protease required for Cf-2-dependent disease resistance. *Science* **308**, 1783-1786.

References

- Ross-Ibarra, J., P.L. Morrell, and B.S. Gaut, 2007: Plant domestication, a unique opportunity to identify the genetic basis of adaptation. *Proceedings of the National Academy of Sciences* **104**, 8641-8648.
- Roux, M., B. Schwessinger, C. Albrecht, D. Chinchilla, A. Jones, N. Holton, F.G. Malinovsky, M. Tör, S. de Vries, and C. Zipfel, 2011: The Arabidopsis Leucine-Rich Repeat Receptor-Like Kinases BAK1/SERK3 and BKK1/SERK4 Are Required for Innate Immunity to Hemibiotrophic and Biotrophic Pathogens. *The Plant Cell* **23**, 2440-2455.
- Ruan, Y.-L., 2014: Sucrose Metabolism: Gateway to Diverse Carbon Use and Sugar Signaling. *Annual Review of Plant Biology* **65**, 33-67.
- Salentijn, E.M.J., N.N. Sandal, W. Lange, T.S.M. De Bock, F.A. Krens, K.A. Marcker, and W.J. Stiekema, 1992: Isolation of DNA markers linked to a beet cyst nematode resistance locus in *Beta patellaris* and *Beta procumbens*. *Molecular and General Genetics MGG* **235**, 432-440.
- Samuelian, S., M. Kleine, C.P. Ruyter-Spira, R.M. Klein-Lankhorst, and C. Jung, 2004: Cloning and functional analyses of a gene from sugar beet up-regulated upon cyst nematode infection. *Plant Molecular Biology* **54**, 147-156.
- Santoni, S., and A. Bervillé, 1992: Two different satellite DNAs in *Beta vulgaris* L.: evolution, quantification and distribution in the genus. *Theoretical and Applied Genetics* **84**, 1009-1016.
- Savitsky, H., 1975: Hybridization between *Beta vulgaris* and *B. procumbens* and transmission of nematode (*Heterodera schachtii*) resistance to sugarbeet. *Canadian Journal of Genetics and Cytology* **17**, 12.
- Schmidt, A., 1871: Über den rüben-nematoden (*Heterodera schachtii* AS). *Zeitschr. Ver. Rübenzucker-Ind. Zoolver* **21**, 1-19.
- Schmidt, T., and M. Metzloff, 1991: Cloning and characterization of a *Beta vulgaris* satellite DNA family. *Gene* **101**, 247-250.
- Schmidt, T., and J.S. Heslop-Harrison, 1996: High-resolution mapping of repetitive DNA by in situ hybridization: Molecular and chromosomal features of prominent dispersed and discretely localized DNA families from the wild beet species *Beta procumbens*. *Plant Molecular Biology* **30**, 1099-1113.
- Schmidt, T., H. Junghans, and M. Metzloff, 1990: Construction of *Beta procumbens*-specific DNA probes and their application for the screening of *B. vulgaris* x *B. procumbens* (2n = 19) addition lines. *Theoretical and Applied Genetics* **79**, 177-181.
- Schmidt, T., S. Kubis, A. Katsiotis, C. Jung, and J. Heslop-Harrison, 1998: Molecular and chromosomal organization of two repetitive DNA sequences with intercalary locations in sugar beet and other Beta species. *Theoretical and applied genetics* **97**, 696-704.
- Schneider, K., D. Kulosa, T.R. Soerensen, S. Möhring, M. Heine, G. Durstewitz, A. Polley, E. Weber, Jamsari, J. Lein, U. Hohmann, E. Tahiro, B. Weisshaar, B. Schulz, G. Koch, C. Jung, and M. Ganai, 2007: Analysis of DNA polymorphisms in sugar beet (*Beta*

References

- vulgaris* L.) and development of an SNP-based map of expressed genes. *Theoretical and Applied Genetics* **115**, 601.
- Schondelmaier, J., and C. Jung, 1997: Chromosomal assignment of the nine linkage groups of sugar beet (*Beta vulgaris* L.) using primary trisomics. *Theoretical and Applied Genetics* **95**, 590-596.
- Schulte, D., D. Cai, M. Kleine, L. Fan, S. Wang, and C. Jung, 2006: A complete physical map of a wild beet (*Beta procumbens*) translocation in sugar beet. *Molecular Genetics and Genomics* **275**, 504-511.
- Shen, Y., B.V. Ford-Iloyd, and H.J. Newbury, 1998: Genetic relationships within the genus *Beta* determined using both PCR-based marker and DNA sequencing techniques. *Heredity* **80**, 624-632.
- Shindo, T., J.C. Misas-Villamil, A.C. Hörger, J. Song, and R.A.L. van der Hoorn, 2012: A Role in Immunity for *Arabidopsis* Cysteine Protease RD21, the Ortholog of the Tomato Immune Protease C14. *PLoS ONE* **7**, e29317-e29317.
- Shinohara, M., and A. Shinohara, 2013: Multiple Pathways Suppress Non-Allelic Homologous Recombination during Meiosis in *Saccharomyces cerevisiae*. *PLOS ONE* **8**, e63144.
- Siddique, S., and F.M.W. Grundler, 2015: Chapter Five - Metabolism in Nematode Feeding Sites, In: C. Escobar and C. Fenoll, (eds.) *Advances in Botanical Research*, 119-138, Vol. 73. Academic Press.
- Sijmons, P.C., F.M.W. Grundler, N. von Mende, P.R. Burrows, and U. Wyss, 1991: *Arabidopsis thaliana* as a new model host for plant-parasitic nematodes. *The Plant Journal* **1**, 245-254.
- Simões, I., R. Faro, D. Bur, and C. Faro, 2007: Characterization of recombinant CDR1, an *Arabidopsis* aspartic proteinase involved in disease resistance. *Journal of Biological Chemistry* **282**, 31358-31365.
- Sobczak, M., and W. Golinowski, 2011: Cyst Nematodes and Syncytia Genomics and Molecular Genetics of Plant-Nematode Interactions, 61-82. Springer Netherlands.
- Sobczak, M., W. Golinowski, and F.M.W. Grundler, 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.
- Song, L., L. Jiang, H. Han, A. Gao, X. Yang, L. Li, and W. Liu, 2013: Efficient Induction of Wheat-Agropyron cristatum 6P Translocation Lines and GISH Detection. *PLOS ONE* **8**, e69501.
- Srivastava, R., J.X. Liu, H. Guo, Y. Yin, and S.H. Howell, 2009: Regulation and processing of a plant peptide hormone, AtRALF23, in *Arabidopsis*. *Plant J* **59**, 930-9.
- Staginnus, C., B. Huettel, C. Desel, T. Schmidt, and G. Kahl, 2001: A PCR-based assay to detect En/Spm-like transposon sequences in plants. *Chromosome Research* **9**, 591-605.

References

- Stegmann, M., J. Monaghan, E. Smakowska-Luzan, H. Rovenich, A. Lehner, N. Holton, Y. Belkhadir, and C. Zipfel, 2017: The receptor kinase FER is a RALF-regulated scaffold controlling plant immune signaling. *Science* **355**, 287-289.
- Stevanato, P., C. Broccanello, F. Biscarini, M. Del Corvo, G. Sablok, L. Panella, A. Stella, and G. Concheri, 2014: High-Throughput RAD-SNP Genotyping for Characterization of Sugar Beet Genotypes. *Plant Molecular Biology Reporter* **32**, 691-696.
- Stevanato, P., D. Trebbi, L. Panella, K. Richardson, C. Broccanello, L. Pakish, A.L. Fenwick, and M. Saccomani, 2015: Identification and Validation of a SNP Marker Linked to the Gene HsBvm-1 for Nematode Resistance in Sugar Beet. *Plant Molecular Biology Reporter* **33**, 474-479.
- Sun, X., T. Fu, N. Chen, J. Guo, J. Ma, M. Zou, C. Lu, and L. Zhang, 2010: The Stromal Chloroplast Deg7 Protease Participates in the Repair of Photosystem II after Photoinhibition in Arabidopsis. *Plant Physiology* **152**, 1263-1273.
- Suty, L., J. Lequeu, A. Lançon, P. Etienne, A.-S. Petitot, and J.-P. Blein, 2003: Preferential induction of 20S proteasome subunits during elicitation of plant defense reactions: towards the characterization of “plant defense proteasomes”. *The International Journal of Biochemistry & Cell Biology* **35**, 637-650.
- Szakasits, D., P. Heinen, K. Wieczorek, J. Hofmann, F. Wagner, D.P. Kreil, P. Sykacek, F.M.W. Grundler, and H. Bohlmann, 2009: The transcriptome of syncytia induced by the cyst nematode *Heterodera schachtii* in Arabidopsis roots. *Plant Journal* **57**, 771-784.
- Thomas, E.L., and R.A.L. van der Hoorn, 2018: Ten Prominent Host Proteases in Plant-Pathogen Interactions. *International journal of molecular sciences* **19**.
- Thompson, E.P., S.G. Llewellyn Smith, and B.J. Glover, 2012: An Arabidopsis rhomboid protease has roles in the chloroplast and in flower development. *Journal of Experimental Botany* **63**, 3559-3570.
- Tian, M., J. Win, R.V.d. Hoorn, E.V.d. Knaap, and S. Kamoun, 2007: A *Phytophthora infestans* cystatin-like protein targets a novel tomato papain-like apoplastic protease. *Plant Physiol* **143**.
- Tian, Y., 2003: PCR-based Cloning of the Second Nematode Resistance Gene *Hs1-1pro-1* and Resistance Gene Analogues from Sugar Beet (*Beta vulgaris* L.), Christian-Albrechts-Universität zu Kiel.
- Tian, Y., L. Fan, T. Thurau, C. Jung, and D. Cai, 2004: The Absence of TIR-Type Resistance Gene Analogues in the Sugar Beet (*Beta vulgaris* L.) Genome. *Journal of Molecular Evolution* **58**, 40-53.
- Tripathi, L.P., and R. Sowdhamini, 2006: Cross genome comparisons of serine proteases in Arabidopsis and rice. *BMC Genomics* **7**, 200.
- Urban, S., and M.S. Wolfe, 2005: Reconstitution of intramembrane proteolysis *in vitro* reveals that pure rhomboid is sufficient for catalysis and specificity. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 1883-1888.

- Urban, S., and S.W. Dickey, 2011: The rhomboid protease family: A decade of progress on function and mechanism, Vol. 12.
- Van Der Vossen, E.A.G., J.N.A.M.R. Van Der Voort, K. Kanyuka, A. Bendahmane, H. Sandbrink, D.C. Baulcombe, J. Bakker, W.J. Stiekema, and R.M. Klein-Lankhorst, 2000: Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: A virus and a nematode. *Plant Journal* **23**, 567-576.
- Van Geyt, J., W. Lange, M. Oleo, and T.S. De Bock, 1990: Natural variation within the genus *Beta* and its possible use for breeding sugar beet: a review. *Euphytica* **49**, 57-76.
- Van Geyt, J.P.C., M. Oléo, W. Lange, and T.S.M. De Bock, 1988: Monosomic additions in beet (*Beta vulgaris*) carrying extra chromosomes of *Beta procumbens*. *Theoretical and Applied Genetics* **76**, 577-586.
- Vanholme, B., P. Kast, A. Haegeman, J. Jacob, W.I.M. Grunewald, and G. Gheysen, 2009: Structural and functional investigation of a secreted chorismate mutase from the plant-parasitic nematode *Heterodera schachtii* in the context of related enzymes from diverse origins. *Molecular Plant Pathology* **10**, 189-200.
- Villareal, R.L., S. Rajaram, A. Mujeeb-Kazi, and E. Del Toro, 1991: The Effect of Chromosome 1B/1R Translocation on the Yield Potential of Certain Spring Wheats (*Triticum aestivum* L.). *Plant Breeding* **106**, 77-81.
- Wang, D., C. Zhang, B. Wang, B. Li, Q. Wang, D. Liu, H. Wang, Y. Zhou, L. Shi, F. Lan, and Y. Wang, 2019: Optimized CRISPR guide RNA design for two high-fidelity Cas9 variants by deep learning. *Nature Communications* **10**, 4284-4284.
- Wang, X., D. Meyers, Y. Yan, T. Baum, G. Smant, R. Hussey, and E. Davis, 1999: In planta localization of a beta-1,4-endoglucanase secreted by *Heterodera glycines*. *Mol Plant Microbe Interact* **12**, 64-7.
- Waterhouse, A.M., J.B. Procter, D.M.A. Martin, M. Clamp, and G.J. Barton, 2009: Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191.
- Withers, J., and X. Dong, 2017: Post-translational regulation of plant immunity. *Current Opinion in Plant Biology* **38**, 124-132.
- Wyss, P., T. Boller, and A. Wiemken, 1992: Testing the effect of biological control agents on the formation of vesicular arbuscular mycorrhiza. *Plant and Soil* **147**, 159-162.
- Wyss, U., and F.M.W. Grundler, 1992: Feeding behavior of sedentary plant parasitic nematodes. *Netherlands Journal of Plant Pathology* **98**, 165-173.
- Xia, Y., H. Suzuki, J. Borevitz, J. Blount, Z. Guo, K. Patel, R.A. Dixon, and C. Lamb, 2004: An extracellular aspartic protease functions in Arabidopsis disease resistance signaling. *The EMBO Journal* **23**, 980-988.
- Yang, H., S. Postel, B. Kemmerling, and U.W.E. Ludewig, 2014: Altered growth and improved resistance of Arabidopsis against *Pseudomonas syringae* by overexpression of the basic amino acid transporter AtCAT1. *Plant, Cell & Environment* **37**, 1404-1414.

References

Yeates, G.W., 2010: Nematodes in Ecological Webs eLS.

Yeates, G.W., T. Bongers, R.G. De Goede, D.W. Freckman, and S.S. Georgieva, 1993: Feeding habits in soil nematode families and genera-an outline for soil ecologists. *Journal of nematology* **25**, 315-331.

Zhang, H., S. Dong, M. Wang, W. Wang, W. Song, X. Dou, X. Zheng, and Z. Zhang, 2010: The role of vacuolar processing enzyme (VPE) from *Nicotiana benthamiana* in the elicitor-triggered hypersensitive response and stomatal closure. *Journal of Experimental Botany* **61**, 3799-3812.

11 Supplementary data on CD/DVD

The following supplemental data are available on a DVD and can be distributed upon request (contact: Prof. Dr. Christian Jung, c.jung@plantbreeding.uni-kiel.de).

File name	Content	Format
Hs4 sequences	<i>Hs4</i> annotated gene sequence	.clc
Hs4 protein	H4 protein sequence with motifs	.clc
Translocation scaffolds	Sequences of 19 translocation specific scaffolds	.fasta
Translocation specific gene-models	All gene models present on translocation	Microsoft Excel file
Translocation specific gene-models RPKM	RPKM expression value calculated using transcriptome data	Microsoft Excel file
Repeat blast	<i>P. procumbens</i> specific repeat blast search	Microsoft Excel file
SOAP config	Configuration file for SOAP de novo assembler	.txt
MAKER config	Configuration file for MAKER de novo assembler	.txt

12 Curriculum Vitae and Publications

12.1 Curriculum vitae

Personal Data

Name: Avneesh Kumar
Date of birth: 27th of July 1991
Place of birth: Meerut, India
Nationality: Indian

Education

09/2017 – present	Ph.D. student at the Plant Breeding Institute, Kiel University Topic: Cloning of the beet cyst nematode resistance gene <i>Hs4</i> from the wild beet <i>Patellifolia procumbens</i>
10/2014 - 06/2017	Master of AgriGenomics, Kiel University Master's thesis at the Environmental Genomics Lab, Kiel University Topic: Evolution of Promoter sequences in the fungal Pathogen <i>Zymoseptoria tritici</i> .
07/2010- 04/2014	Bachelor of Technology, School of Engineering, Department of Biotechnology, Sir Padampat Singhanian University, Udaipur, India

13 Declaration of own contribution

The chapters of this thesis are not published, and they include contributions of the doctoral candidate alone.

14 Acknowledgments

First, I would like to thank Prof. Dr. Christian Jung for giving me the opportunity to conduct my doctoral research at the Plant Breeding Institute and for your support and guidance. Your supervision, critical comments, and advice made me learn a lot about my research work and shape myself as a better scientist. Then, I would like to express my deepest gratitude to Birgit Defant and Maike Schneider for their continuous support throughout the project. Many thanks. I also would like to thank Sarah, Srijan, Amar and Annika for reviewing this thesis.

I am grateful to Moni for helping at the greenhouse and for your kind support outside the lab. I would also like to thank Hans for very fruitful discussions. I also would like to thank Siegbert for the discussions. I thank Gina for giving me initial start in the project.

I am incredibly thankful to Sarah, Srijan, Amar, Dilan, Nirosha, Tahmina, Smit, and Yixin not being just colleagues. You are really great friends. Many thanks for your valuable discussions and great support all the time. I would like to thank all the other colleagues at the institute for having a nice working atmosphere.

I appreciate the Federal State of Schleswig-Holstein's financial support and the Sanger sequencing service of the Institute of Clinical Molecular Biology Kiel.

Last but not least, special thanks go to my parents, wife, sisters. They are the ones who made me capable of being here; without them on my side every time, this wouldn't have been possible. I am delighted to thank my wife Nandini, who has supported me since the day we met. Encouragement and motivation, which I got from my family, kept me going.